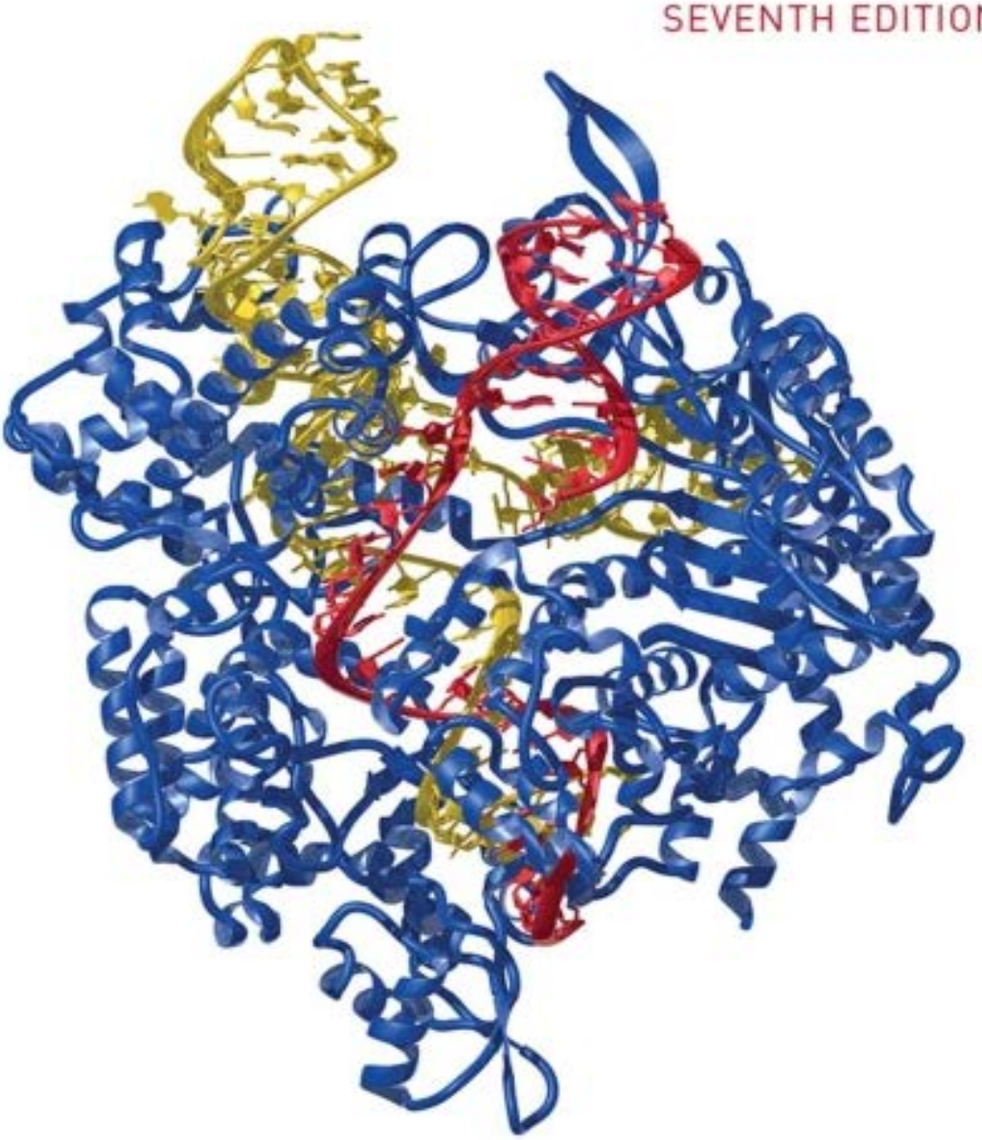
BIOCHEMISTRY

THE MOLECULAR BASIS OF LIFE

SEVENTH EDITION



TRUDY McKEE | JAMES R. McKEE

OXFORD
UNIVERSITY PRESS

Biochemistry

The Molecular Basis of Life

Biochemistry

The Molecular Basis of Life

Seventh Edition

Trudy McKee
James R. McKee

New York Oxford
OXFORD UNIVERSITY PRESS

Facts are stubborn things; and whatever may be our wishes, our inclinations, or the dictates of our passions, they cannot alter the state of facts and evidence.
—John Adams, second president of the United States (1797–1801)

The scientific endeavor seeks to understand the nature of reality. This edition is dedicated to those individuals who are willing to give up cherished ideas and assumptions when they are confronted by verified, proven facts.

Brief Contents

Chapter-Opening Vignettes
Biochemistry in Perspective Essays
Biochemistry in the Lab Boxes
Preface
General and Organic Chemistry Review Primer

1. Biochemistry: An Introduction
2. Living Cells
3. Water: The Matrix of Life
4. Energy
5. Amino Acids, Peptides, and Proteins
6. Enzymes
7. Carbohydrates
8. Carbohydrate Metabolism
9. Aerobic Metabolism I: The Citric Acid Cycle
10. Aerobic Metabolism II: Electron Transport and Oxidative Phosphorylation
11. Lipids and Membranes
12. Lipid Metabolism
13. Photosynthesis
14. Nitrogen Metabolism I: Synthesis
15. Nitrogen Metabolism II: Degradation
16. Integration of Metabolism
17. Nucleic Acids
18. Genetic Information
19. Protein Synthesis

Appendix: Solutions
Glossary
Credits
Index

Contents

Cover

Half Title

Title Page

Copyright

Dedication

Brief Content

Content

Chapter-Opening Vignettes

Biochemistry in Perspective Essays

Biochemistry in the Lab Boxes

Preface

Half Title 1

Common Abbreviations in Biochemistry

General and Organic Chemistry Review Primer

CHAPTER 1

Biochemistry: An Introduction

Why Study Biochemistry?

1.1 What Is Life?

1.2 Biomolecules

Functional Groups of Organic Biomolecules
Major Classes of Small Biomolecules

1.3 Is the Living Cell a Chemical Factory?

Biochemical Reactions
Energy
Overview of Metabolism
Biological Order

1.4 Systems Biology

Emergence
Robustness
Systems Biology Model Concepts
Living Organisms and Robustness

- **Biochemistry IN THE LAB**

  An Introduction

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Life: It Is a Mystery!

CHAPTER 2

Living Cells

Our Bodies, Our Selves

2.1 Core Biochemistry Concepts

Biochemistry and Water
Biological Membranes
Self-Assembly
Molecular Machines
Macromolecular Crowding
Proteostasis
Signal Transduction

2.2 Structure of Prokaryotic Cells

Cell Wall

Plasma Membrane
Cytoplasm
Pili and Flagella

2.3 Structure of Eukaryotic Cells

Plasma Membrane
Endoplasmic Reticulum
Golgi Apparatus
Vesicular Organelles and Lysosomes: The Endocytic Pathway
Nucleus
Mitochondria
Peroxisomes
Chloroplasts
Cytoskeleton

Biochemistry IN PERSPECTIVE
Primary Cilia and Human Disease

Biochemistry IN THE LAB
Cell Technology

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Organelles and Human Disease
Biochemistry IN PERSPECTIVE
Caveolar Endocytosis

CHAPTER 3

Water: The Matrix of Life

Water, Water, Everywhere

3.1 Molecular Structure of Water

3.2 Noncovalent Bonding

Ionic Interactions
Hydrogen Bonds
Van der Waals Forces

3.3 Thermal Properties of Water

3.4 Solvent Properties of Water

Hydrophilic Molecules, Cell Water Structuring, and Sol–Gel Transitions
Hydrophobic Molecules and the Hydrophobic Effect
Amphipathic Molecules
Osmotic Pressure

3.5 Ionization of Water

Acids, Bases, and pH
Buffers
Physiological Buffers

Biochemistry IN PERSPECTIVE

Cell Volume Regulation and Metabolism

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Water, Abiotic Stress, and Compatible Solutes
Biochemistry IN THE LAB
Dialysis

CHAPTER 4

Energy

Energy and Life's Deep, Dark Secrets

4.1 Thermodynamics

First Law of Thermodynamics
Second Law of Thermodynamics

4.2 Free Energy

Standard Free Energy Changes
Coupled Reactions
The Hydrophobic Effect Revisited

4.3 Energy Currency Molecules

Energy and Reducing Power

Biochemistry IN PERSPECTIVE

Nonequilibrium Thermodynamics and the Evolution of Life

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
The Extremophiles: Organisms That Make a Living in Hostile Environments

---

CHAPTER 5

Amino Acids, Peptides, and Proteins

Spider Silk: A Biosteel Protein

5.1 Amino Acids

Amino Acid Classes
Biologically Active Amino Acids
Modified Amino Acids in Proteins
Amino Acid Stereoisomers
Amino Acid Ionizable Side Chain Groups and pH
Amino Acid Reactions

5.2 Peptides

5.3 Proteins

Functions
Shape and Composition
Protein Structure
The Folding Problem
Fibrous Proteins
Globular Proteins

Biochemistry IN PERSPECTIVE
Spider Silk and Biomimetics

Biochemistry IN THE LAB
Protein Technology

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Protein Poisons

Biochemistry IN PERSPECTIVE
Protein Folding and Human Disease

Biochemistry IN PERSPECTIVE
Molecular Machines and Motor Proteins

Biochemistry IN PERSPECTIVE
Myosin: A Molecular Machine

CHAPTER 6

Enzymes

Pain, Aspirin, and the Cox Enzymes

6.1 Properties of Enzymes

Enzyme Catalysts: The Basics
Enzymes: Activation Energy and Reaction Equilibrium
Enzymes and Macromolecular Crowding Effects
Enzyme Specificity

6.2 Classification of Enzymes

6.3 Catalysis

Organic Reactions and the Transition State
Transition State Stabilization
Catalytic Mechanisms
The Role of Amino Acids in Enzyme Catalysis
The Role of Cofactors in Enzyme Catalysis
Effects of Temperature and pH on Enzyme-Catalyzed Reactions
Detailed Mechanisms of Enzyme Catalysis

6.4 Enzyme Kinetics

Michaelis–Menten Kinetics
Lineweaver–Burk Plots
Multisubstrate Reactions
Enzyme Inhibition
Enzyme Kinetics, Metabolism, and Macromolecular Crowding

6.5 Enzyme Regulation

Genetic Control
Covalent Modification
Allosteric Regulation
Compartmentation

Biochemistry IN PERSPECTIVE
Alcohol Dehydrogenase: A Tale of Two Species

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online

Biochemistry IN PERSPECTIVE
Enzymes and Clinical Medicine

Biochemistry IN PERSPECTIVE
Quantum Tunneling and Catalysis

Biochemistry IN PERSPECTIVE
Humans and Enzymes—A Brief History

CHAPTER 7

Carbohydrates

Sweet and Bitter Taste: The Roles of Sugar Molecules

7.1 Monosaccharides

Monosaccharide Stereoisomers
Cyclic Structure of Monosaccharides
Reactions of Monosaccharides
Important Monosaccharides
Monosaccharide Derivatives

7.2 Disaccharides

7.3 Polysaccharides

Homoglycans
Heteroglycans

7.4 Glycoconjugates

Proteoglycans
Glycoproteins

7.5 The Sugar Code

Lectins: Translators of the Sugar Code
The Glycome
Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Scurvy and Ascorbic Acid

Biochemistry IN PERSPECTIVE
Sweet Medicine

Biochemistry IN PERSPECTIVE
Conversion of Fischer Structures into Haworth Structures

Biochemistry IN THE LAB
Glycomics

CHAPTER 8

Carbohydrate Metabolism

Metabolism and Jet Engines

8.1 Glycolysis

The Reactions of the Glycolytic Pathway
The Fates of Pyruvate
The Energetics of Glycolysis
Regulation of Glycolysis

8.2 Gluconeogenesis

Gluconeogenesis Reactions
Gluconeogenesis Substrates
Gluconeogenesis Regulation

8.3 The Pentose Phosphate Pathway

8.4 Metabolism of Other Important Sugars

Fructose Metabolism

8.5 Glycogen Metabolism

Glycogenesis
Glycogenolysis
Regulation of Glycogen Metabolism

Biochemistry IN PERSPECTIVE
Saccharomyces cerevisiae and the Crabtree Effect

Biochemistry IN PERSPECTIVE
Turbo Design Can Be Dangerous

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Fermentation: An Ancient Heritage

CHAPTER 9

Aerobic Metabolism I: The Citric Acid Cycle

Oxygen and Evolution: Chance and Necessity

9.1 Oxidation-Reduction Reactions

Redox Coenzymes
Aerobic Metabolism

9.2 Citric Acid Cycle

Conversion of Pyruvate to Acetyl-CoA
Reactions of the Citric Acid Cycle
Fate of Carbon Atoms in the Citric Acid Cycle
The Amphibolic Citric Acid Cycle
Citric Acid Cycle Regulation
The Citric Acid Cycle and Human Disease
The Glyoxylate Cycle

Biochemistry IN PERSPECTIVE
Carcinogenesis: The Warburg Effect and Metabolic Reprogramming

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Hans Krebs and the Citric Acid Cycle
Biochemistry IN PERSPECTIVE
The Evolutionary History of the Citric Acid Cycle

CHAPTER 10

Aerobic Metabolism II: Electron Transport and Oxidative Phosphorylation

Oxygen: A Molecular Paradox

10.1 Electron Transport

Electron Transport and Its Components
Electron Transport: The Fluid State and Solid State Models
Electron Transport Inhibitors

10.2 Oxidative Phosphorylation

The Chemiosmotic Theory
ATP Synthesis
Control of Oxidative Phosphorylation
Creatine Kinase and Creatine
The Complete Oxidation of Glucose
Uncoupled Electron Transport

10.3 Oxygen, Cell Signaling, and Oxidative Stress

Reactive Oxygen Species
Antioxidant Enzyme Systems
The Redox Proteome
Antioxidant Nutrients

Biochemistry IN PERSPECTIVE
Myocardial Infarct: Ischemia and Reperfusion

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Glucose-6-Phosphate Dehydrogenase Deficiency

CHAPTER 11

Lipids and Membranes

The Low-Fat Diet

11.1 Lipid Classes

Fatty Acids
The Eicosanoids
Triacylglycerols
Wax Esters
Phospholipids
Phospholipases
Sphingolipids
Sphingolipid Storage Diseases
Isoprenoids
Lipoproteins

11.2 Membranes

Membrane Structure
Membrane Function

Biochemistry IN PERSPECTIVE
Alzheimer's Disease and Apolipoprotein E4

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Botulism and Membrane Fusion

CHAPTER 12

Lipid Metabolism

Abetalipoproteinemia

12.1 Fatty Acids, Triacylglerols, and the Lipoprotein Pathways

Dietary Fat: Digestion, Absorption, and Transport
Triacylglycerol Metabolism in Adipocytes
Fatty Acid Degradation
The Complete Oxidation of a Fatty Acid
Fatty Acid Oxidation: Double Bonds and Odd Chains
Fatty Acid Biosynthesis
Regulation of Fatty Acid Metabolism in Mammals
Lipoprotein Metabolism: The Endogenous Pathway

12.2 Membrane Lipid Metabolism

Phospholipid Metabolism

12.3 Isoprenoid Metabolism

Cholesterol Metabolism
The Cholesterol Biosynthetic Pathway and Drug Therapy

- **Biochemistry IN PERSPECTIVE**
  Atherosclerosis
- **Biochemistry IN PERSPECTIVE**
  Biotransformation

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Sphingolipid Metabolism

CHAPTER 13

Photosynthesis

Climate Change, Renewable Energy, and Photosynthesis

13.1 Chlorophyll and Chloroplasts

13.2 Light

13.3 Light Reactions

Photosystem II and Water Oxidation
Photosystem I and NADPH Synthesis
Photophosphorylation

13.4 The Light-Independent Reactions

The Calvin Cycle
Photorespiration
Alternatives to C3 Metabolism

13.5 Regulation of Photosynthesis

Light Control of Photosynthesis
Control of Ribulose-1,5-Bisphosphate Carboxylase

Biochemistry IN PERSPECTIVE
The Artificial Leaf: Biomimetic Photosynthesis

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Starch and Sucrose Metabolism
Biochemistry IN THE LAB
Photosynthetic Studies
Biochemistry IN PERSPECTIVE
Photosynthesis in the Deep
Biochemistry IN PERSPECTIVE
Crassulacean Acid Metabolism

CHAPTER 14

Nitrogen Metabolism I: Synthesis

Nitrogen and the Gulf of Mexico Dead Zone

14.1 Nitrogen Fixation

The Nitrogen Fixation Reaction
Nitrogen Assimilation

14.2 Amino Acid Biosynthesis

Amino Acid Metabolism Overview
Reactions of Amino Groups
Biosynthesis of the Amino Acids

14.3 Biosynthetic Reactions Involving Amino Acids

One-Carbon Metabolism
Glutathione
Neurotransmitters
Nucleotides
Heme

Biochemistry IN PERSPECTIVE
Gasotransmitters

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
The Amine Neurotransmitters
Biochemistry IN PERSPECTIVE
The Catecholamines
Biochemistry IN PERSPECTIVE
Parkinson's Disease and Dopamine
Biochemistry IN PERSPECTIVE
Heme and Chlorophyll Biosynthesis
Biochemistry IN PERSPECTIVE
The Essential Amino Acids
Biochemistry IN PERSPECTIVE
The Nucleotides: IMP Biosynthesis
Biochemistry IN PERSPECTIVE
Lead Poisoning

CHAPTER 15

Nitrogen Metabolism II: Degradation

Degradative Pathways and Human Disorders

15.1 Protein Turnover

Ubiquitin Proteasomal System
Autophagy-Lysosomal System

15.2 Amino Acid Catabolism

Deamination
Urea Synthesis

Control of the Urea Cycle
Catabolism of Amino Acid Carbon Skeletons

15.3 Nucleotide Degradation
Purine Catabolism
Pyrimidine Catabolism

Biochemistry IN PERSPECTIVE
Disorders of Amino Acid Catabolism

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
Hyperammonemia
Biochemistry IN PERSPECTIVE
Gout
Biochemistry IN PERSPECTIVE
Catecholamine Inactivation
Biochemistry IN PERSPECTIVE
Heme Biotransformation

CHAPTER 16

Integration of Metabolism

Hypertension and Uric Acid: A Diet Connection?

16.1 Overview of Metabolism

16.2 Intercellular Communication
Polypeptide Signal Molecules
Cell-Surface Receptors
Enzyme-Linked Receptors
Nuclear Receptors

16.3 Metabolism in the Mammalian Body: The Feeding–Fasting Cycle
The Feeding Phase
The Fasting Phase
Feeding Behavior

Biochemistry IN PERSPECTIVE
Diabetes Mellitus

Biochemistry IN PERSPECTIVE
Obesity and the Metabolic Syndrome

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online

Biochemistry IN PERSPECTIVE

Mammalian Hormones and the Hormone Cascade System

Biochemistry IN PERSPECTIVE

Metabolism in the Mammalian Body: Division of Labor

CHAPTER 17

Nucleic Acids

What Makes Us Human?

17.1 DNA

DNA Structure: The Nature of Mutation
DNA Structure: The Genetic Material
DNA Structure: Variations on a Theme
DNA Supercoiling
Chromosomes
Genome Structure

17.2 RNA

Transfer RNA
Ribosomal RNA
Messenger RNA
Noncoding RNA

17.3 Viruses

Bacteriophage T4: A Viral Lifestyle

- **Biochemistry IN PERSPECTIVE**
  Epigenetics and the Epigenome: Genetic Inheritance beyond DNA Base Sequences
- **Biochemistry IN THE LAB**
  Nucleic Acid Methods
- **Biochemistry IN PERSPECTIVE**
  Forensic Investigations
- **Biochemistry IN PERSPECTIVE**
  Ebola Virus

Chapter Summary
Suggested Readings
Key Words

Review Questions

Available Online
Biochemistry IN PERSPECTIVE
A Short History of DNA Research: The Early Years
Biochemistry IN PERSPECTIVE
HIV Infection

CHAPTER 18

Genetic Information

DNA and Chimeras: A Biological and Legal Mystery

18.1 Genetic Information: Replication, Repair, and Recombination

DNA Replication
DNA Repair
DNA Recombination

18.2 Transcription

Transcription in Prokaryotes
RNAP and the Prokaryotic Transcription Process
Transcription in Eukaryotes

18.3 Gene Expression

Gene Expression in Prokaryotes
Gene Expression in Eukaryotes

- **Biochemistry IN THE LAB**
  Genomics
- **Biochemistry IN THE LAB**
  CRISPR
- **Biochemistry IN PERSPECTIVE**
  Carcinogenesis

Chapter Summary
Suggested Readings
Key Words
Review Questions

Available Online
Biochemistry IN PERSPECTIVE
The Meselson–Stahl Experiment

CHAPTER 19

Protein Synthesis

MRSA: The Superbug

19.1 The Genetic Code

Codon–Anticodon Interactions

The Aminoacyl-tRNA Synthetase Reaction

19.2 Protein Synthesis

Prokaryotic Protein Synthesis

Eukaryotic Protein Synthesis

19.3 The Proteostasis Network

The Heat Shock Response

The Proteostasis Network and Human Disease

- **Biochemistry IN PERSPECTIVE**
  Trapped Ribosomes: RNA to the Rescue!
- **Biochemistry IN PERSPECTIVE**
  Context-Dependent Coding Reassignment
- **Biochemistry IN THE LAB**
  Proteomics

Chapter Summary

Suggested Readings

Key Words

Review Questions

Available Online

Biochemistry IN PERSPECTIVE

EF-Tu: A Motor Protein

Appendix: Solutions

Glossary

Credits

Index

Chapter-Opening Vignettes

CHAPTER 1

Why Study Biochemistry?

CHAPTER 2

Our Bodies, Our Selves

CHAPTER 3

Water, Water, Everywhere

CHAPTER 4

Energy and Life’s Deep, Dark Secrets

CHAPTER 5

Spider Silk: A Biosteel Protein

CHAPTER 6

Pain, Aspirin, and the Cox Enzymes

CHAPTER 7

Sweet and Bitter Taste: The Roles of Sugar Molecules

CHAPTER 8

Metabolism and Jet Engines

CHAPTER 9

Oxygen and Evolution: Chance and Necessity

CHAPTER 10

Oxygen: A Molecular Paradox

CHAPTER 11

The Low-Fat Diet

CHAPTER 12

Abetalipoproteinemia

CHAPTER 13

Climate Change, Renewable Energy, and Photosynthesis

CHAPTER 14

Nitrogen and the Gulf of Mexico Dead Zone

CHAPTER 15

Degradative Pathways and Human Disorders

CHAPTER 16

Hypertension and Uric Acid: A Diet Connection?

CHAPTER 17

What Makes Us Human?

CHAPTER 18

DNA and Chimeras: A Biological and Legal Mystery

CHAPTER 19

MRSA: The Superbug

Biochemistry in Perspective Essays

CHAPTER 2

Primary Cilia and Human Disease

What effects do nonmotile cilia have on human health?

CHAPTER 3

Cell Volume Regulation and Metabolism

Is there a relationship between metabolism and cell volume?

CHAPTER 4

Nonequilibrium Thermodynamics and the Evolution of Life

How does thermodynamic theory relate to energy flow in living organisms?

CHAPTER 5

Spider Silk and Biomimetics

What properties of spider silk have made it the subject of research worth hundreds of millions of dollars?

CHAPTER 6

Alcohol Dehydrogenase: A Tale of Two Species

Why can humans make and consume alcoholic beverages?

CHAPTER 8

Saccharomyces cerevisiae and the Crabtree Effect

What unique properties of *Saccharomyces cerevisiae* make it so useful in the production of wine, beer, and bread?

Turbo Design Can Be Dangerous

Why must turbo design biochemical pathways be rigorously controlled?

CHAPTER 9

Carcinogenesis: The Warburg Effect and Metabolic Reprogramming

What is aerobic glycolysis, and how does it impact cancerous cells?

CHAPTER 10

Myocardial Infarct: Ischemia and Reperfusion

How are heart cells damaged by the inadequate nutrient and oxygen flow caused by blood clots, and why does the reintroduction of O_2 cause further damage?

CHAPTER 11

NEW! Alzheimer's Disease and Apolipoprotein E4

What structural property of apolipoprotein E4 is responsible for making it a high-risk factor for Alzheimer's disease?

CHAPTER 12

Atherosclerosis

What is the biochemical basis of arterial damage in the disease process called atherosclerosis?

Biotransformation

How does the body metabolize potentially toxic hydrophobic molecules?

CHAPTER 13

The Artificial Leaf: Biomimetic Photosynthesis

Can humans successfully produce solar fuels as a substitute for environmentally damaging fossil fuels?

CHAPTER 14

Gasotransmitters

How do gas molecules, previously thought to be toxic at any concentration, act as signal molecules in the mammalian body?

CHAPTER 15

Disorders of Amino Acid Catabolism

What are the effects on human health of deficiency of a single enzyme in amino acid metabolism?

CHAPTER 16

Diabetes Mellitus

Why does diabetes mellitus, a disease in which glucose transport is compromised, damage the entire body?

Obesity and the Metabolic Syndrome

Why are so many humans predisposed to obesity in the modern world, especially in the past 30 years?

CHAPTER 17

Epigenetics and the Epigenome: Genetic Inheritance beyond DNA Base Sequences

How do covalent modifications of DNA and histones affect the functions of multicellular organisms? How do the more than 200 cell types in humans arise from a fertilized egg?

Forensic Investigations

How is DNA analysis used in the investigation of violent crime?

NEW! Ebola Virus

What is Ebola, and why is it so dangerous?

CHAPTER 18

Carcinogenesis

What is cancer, and what are the biochemical processes that facilitate the transformation of normal cells into those with cancerous properties?

CHAPTER 19

Trapped Ribosomes: RNA to the Rescue!

How are ribosomes bound to damaged mRNAs retrieved so that they can be recycled? What happens when a ribosome attempts to translate an mRNA that is either truncated or missing a stop codon?

Context-Dependent Coding Reassignment

How is selenocysteine, a nonstandard amino acid, incorporated into polypeptides during protein synthesis?

Biochemistry in the Lab Boxes

CHAPTER 1

An Introduction

CHAPTER 2

Cell Technology

CHAPTER 5

Protein Technology

CHAPTER 17

Nucleic Acid Methods

CHAPTER 18

Genomics

NEW! CRISPR

CHAPTER 19

Proteomics

Don t forget your favorite Biochemistry in Perspective essays from past editions, now available on the companion website at www.oup.com/us/mckee:

Organelles and Human Disease (Chapter 2)
Caveolar Endocytosis (Chapter 2)
Water, Abiotic Stress, and Compatible Solutes (Chapter 3)
The Extremophiles: Organisms That Make a Living in Hostile Environments (Chapter 4)
Protein Poisons (Chapter 5)
Protein Folding and Human Disease (Chapter 5)
Molecular Machines and Motor Proteins (Chapter 5)
Myosin: A Molecular Machine (Chapter 5)
Enzymes and Clinical Medicine (Chapter 6)
Quantum Tunneling and Catalysis (Chapter 6)
Humans and Enzymes: A Brief History (Chapter 6)

Scurvy and Ascorbic Acid (Chapter 7)
Sweet Medicine (Chapter 7)
Conversion of Fischer Structures into Haworth Structures (Chapter 7)
Fermentation: An Ancient Heritage (Chapter 8)
Hans Krebs and the Citric Acid Cycle (Chapter 9)
The Evolutionary History of the Citric Acid Cycle (Chapter 9)
Glucose-6-Phosphate Dehydrogenase Deficiency (Chapter 10)
Botulism and Membrane Fusion (Chapter 11)
Sphingolipid Metabolism (Chapter 12)
Starch and Sucrose Metabolism (Chapter 13)
Photosynthesis in the Deep (Chapter 13)
Crassulacean Acid Metabolism (Chapter 13)
The Amine Neurotransmitters (Chapter 14)
The Catecholamines (Chapter 14)
Parkinson s Disease and Dopamine (Chapter 14)
Heme and Chlorophyll Biosynthesis (Chapter 14)
The Essential Amino Acids (Chapter 14)
The Nucleotides: IMP Biosynthesis (Chapter 14)
Lead Poisoning (Chapter 14)
Hyperammonemia (Chapter 15)
Gout (Chapter 15)
Catecholamine Inactivation (Chapter 15)
Heme Biotransformation (Chapter 15)
Mammalian Hormones and the Hormone Cascade System (Chapter 16)
Metabolism in the Mammalian Body: Division of Labor (Chapter 16)
A Short History of DNA Research: The Early Years (Chapter 17)
HIV Infection (Chapter 17)
The Meselson-Stahl Experiment (Chapter 18)
EF-Tu: A Motor Protein (Chapter 19)

Your favorite Biochemistry in the Lab boxes from past editions are also available on the companion website at www.oup.com/us/mckee:

Dialysis (Chapter 3)
Protein Sequence Analysis: The Edman Degradation (Chapter 5)
Glycomics (Chapter 7)
Photosynthetic Studies (Chapter 13)

Preface

Welcome to the seventh edition of *Biochemistry: The Molecular Basis of Life*. Although this textbook has been revised and updated to reflect the latest research in biochemistry, our original mission remains unchanged. We continue to believe that the cornerstone of an education in the life sciences is a coherent understanding of the basic principles of biochemistry. Once they have mastered these foundational biochemical concepts, students are prepared to tackle the complexities of their chosen science fields. To that end, we have sought comprehensive coverage of biochemical systems, structures, and reactions, but within the context of the organism. We have thus sought a unique balance between chemistry and biology, while also considering their applications to medicine and human health.

ORGANIZATION AND APPROACH

CHEMICAL AND BIOLOGICAL PRINCIPLES IN BALANCE As with previous editions, the seventh edition is designed for both life science students and chemistry majors. We provide thorough coverage of biochemical principles, structures, and reactions, but within a biological context that emphasizes their relevance.

A REVIEW OF BASIC PRINCIPLES Few assumptions have been made about a student's chemistry and biology background. To ensure that all students are sufficiently prepared for acquiring a meaningful understanding of biochemistry, the first four chapters review the principles of such topics as organic functional groups, non-covalent bonding, thermodynamics, and cell structure. These chapters can either be covered in class or assigned for self-study.

Several topics are introduced in these early chapters and then revisited throughout the book. Examples include cell volume changes triggered by metabolic processes that alter osmotic balance across membranes, the self-assembly of biopolymers such as proteins into supermolecular structures, and proteostasis, the mechanisms whereby cells protect proteins and when necessary degrade them. Other important concepts that are emphasized include the relationship between biomolecular structure and function and the dynamic, unceasing, and self-regulating nature of living processes. Students are also provided with overviews of the major physical and chemical techniques that biochemists have used to explore life at the molecular level.

REAL-WORLD RELEVANCE Because students who take the one-semester biochemistry course come from a range of backgrounds and have diverse career goals, the seventh edition consistently demonstrates the fascinating connections between biochemical principles and the worlds of medicine, nutrition, agriculture, bioengineering, and forensics. Biochemistry in Perspective essays and chapter-opening vignettes, as well as dozens of examples integrated into the body of the text, help students see the relevance of biochemistry to their chosen fields of study.

SUPERIOR PROBLEM-SOLVING PROGRAM Analytical thinking is at the core of the scientific enterprise, and mastery of biochemical principles requires consistent and sustained engagement with a wide range of problems. The seventh edition continues to present students with a complete problem-solving system. This includes in-chapter worked problems, which illustrate how quantitative problems are solved, as well as dozens of practice questions presented throughout the text when new concepts and high-interest topics are introduced. There is also a large set of end-of-chapter questions, including hundreds of new problems, all reorganized for this edition by section number and question type. These include new study questions designed to help students in pre-health programs prepare for the MCAT test.

SIMPLE, CLEAR ILLUSTRATIONS Biochemical concepts often require a high degree of visualization, and we have crafted an art program that brings complex processes to life. More than 720 full-color figures fill the pages of the seventh edition, many newly enhanced with consistent scale and color for a more vivid presentation.

CURRENCY The seventh edition has been updated to present recent developments in the field, while remaining focused on the "big-picture" principles that are the cornerstone of the one-term biochemistry course. These changes again reflect the goal of balanced and thorough coverage of chemistry within a biological context. A detailed list of updated material follows in the next section.

NEW IN THIS EDITION

As a result of the rapid pace of discovery in the life sciences and our commitment to provide students with the highest-quality learning system available in any biochemistry textbook, we have revised the seventh edition in the following ways:

- **New and Updated Applications.** The seventh edition includes two new Biochemistry in Perspective essays, on Alzheimer's disease and the Ebola virus. In addition, a new Biochemistry in the Lab essay introduces the genome-editing tool CRISPR. Together with revised features on carcionogenesis, diabetes, the atherosclerotic process, proteomics, protein analysis, and bioinformatics, these essays will inspire student interest into the latest developments in medical and biochemistry research.
- **Expanded Problem-Solving Program.** The seventh edition includes hundreds of new end-of-chapter questions organized by section number and problem type—multiple-choice, fill-in-the-blank, short-answer, and critical thinking. New to this edition is a set of questions designed to help students preparing for the MCAT exam.
- **Brand-New Illustrations.** With 34 new figures, the seventh edition incorporates a superior and expanded art program designed to help students develop a strong visual grasp of biochemical processes. Many figures have been enhanced for vivid, clear, and consistent presentation.
- **New, Expanded, and Updated Coverage of Current Topics.** The following is a list of some, but not all, of the updated content that has been introduced in the seventh edition:
  - The **General and Organic Chemistry Review Primer** gets students off to the right start with a review of foundational topics in general and organic chemistry. Revised for this edition, the primer includes two new worked problems and a new figure showing trends in the periodic table.

- **Chapter 2** contains more detailed coverage of the function and structure of the Golgi apparatus, including an updated explanation of the role of anterograde and retrograde transport in the processing of new lipid and protein molecules.
- **Chapter 4** offers an expanded discussion of energy currency molecules that introduces NADH, NADPH, and $FADH_2$ as electron carriers with considerable reducing power.
- **Chapter 5** provides new information on the roles of the peptides vasopressin and oxytocin along with an updated discussion of multifunctional proteins, including moonlighting proteins. There is also new information on intrinsically unstructured proteins and intrinsically disordered regions, and a new figure highlights the unstructured protein domains of p53, an important tumor suppressor protein. The Biochemistry in the Lab essay on protein technology has been revised to include a new brief discussion of high-performance liquid chromatography as a protein analysis technique.
- **Chapter 6** has been dramatically reorganized to present catalysis before the discussion of enzyme kinetics. The chapter opens with a new feature on pain, aspirin, and the COX enzymes. New illustrations enhance the descriptions of the structure and functional properties of α-carbonic anhydrase and alcohol dehydrogenase.
- **Chapter 7** now includes an explanation of the importance of glycosylation reactions in animal biology. The discussion of glycoconjugates has been thoroughly revised, with an expanded explanation of the structure and functions of proteoglycans complete with a new illustration of proteoglycan aggregate structure.
- **Chapter 8** begins with a new overview of metabolism that introduces the concepts of signal transduction and genetic regulatory pathways in relation to core metabolic pathways. The discussion of glycolysis regulation includes a revised presentation of hexokinases.
- **Chapter 9** features a new illustration of the pyruvate dehydrogenase complex, a revised presentation of the eight reactions of the citric acid cycle, and a major updating of the discussion of carcinogenesis and the Warburg effect.
- **Chapter 10** offers a set of new figures to present a clearer illustration of electron transport and its components. There is also a new discussion of creatine kinase and creatine, featuring an illustration of the creatine kinase–phosphocreatine shuttle system, which facilitates intracellular transport of high-energy phosphate. An extensive new description of the redox proteome and the four principles of the redox code explain how sensitive redox signaling is a major feature of numerous organismal processes.
- **Chapter 11** presents a new Biochemistry in Perspective essay on Alzheimer's disease and apolipoprotein E. A new figure illustrating the structures of ApoE3 and ApoE4 helps to show why individuals with the ApoE4 variant have an increased risk for Alzheimer's disease. There is also a new in-chapter question on ATP hydrolysis and the unique digestive issues in alligators.
- **Chapter 12** features extensive revisions to the Biochemistry in Perspective essay on the atherosclerotic process. Included in this discussion is the distinction between the different atherosclerosis risks of small buoyant and small dense LDLs.
- Revised figures in **Chapter 13** better explain the noncyclic and cyclic electron transport pathways in chloroplasts.
- **Chapter 15** contains updates to the explanation of ubiquitination and proteasomes. There is also a new table on the carbon skeletons of glucogenic and ketogenic amino acids.
- In **Chapter 16**, the discussion of hormones and intercellular communication has been

thoroughly revised, with extensive new coverage of G-protein–coupled receptors, including new illustrations of G protein and GPCR and an extended discussion of GPCR regulation. Greater attention is given to the functions of insulin and receptor tyrosine kinase signaling. The Biochemistry in Perspective feature on diabetes has been expanded to cover type 3 diabetes and the link between insulin resistance and obesity, atherosclerosis, and fatty liver disease.

- **Chapter 17** now includes an extended discussion of chromatin, complete with illustrations of chromosome territories and TAD chromosomal boundary structure. There is also a new description of DNA and RNA transposons, as well as new discoveries of transfer RNA and noncoding RNA functions. A new Biochemistry in Perspective essay examines the Ebola virus.
- **Chapter 18** includes updated coverage of DNA synthesis and repair, including a description of the SOS response triggered when bacterial cells are exposed to high levels of UV light or mutagenic chemicals. The Biochemistry in the Lab feature on genomics includes updated information on genomic libraries, DNA microarrays, and bioinformatics. A new Biochemistry in the Lab feature examines CRISPR, the genome-editing technology now widely used in biomedical research. The Biochemistry in Perspective essay on carcinogenesis features an extended discussion of the link between cancer and lncRNAs.
- **Chapter 19** includes new material on the initiation phase of eukaryotic protein synthesis, including an explanation of the 43S preinitiation complex and of transcript localization as a mechanism for directing mRNAs to discrete cellular locations. The Biochemistry in the Lab essay includes new methods in proteomics that analyze protein complex mixtures by combining mass spectrometry and high-performance liquid chromatography.

LEARNING PACKAGE

We have created a comprehensive set of additional resources to accompany the seventh edition. These are designed to enrich the educational experience for students and instructors alike.

For Students

STUDENT STUDY GUIDE AND SOLUTIONS MANUAL Written by the textbook authors, this manual provides the solutions to all of the exercises from the text that are not included in the book itself. Each solution has been independently checked for accuracy by a panel of expert reviewers.

ANIMATION AND VIDEO GUIDE Revised for the seventh edition, the student companion website includes a curated guide to dozens of relevant, high-quality videos and animations to help students visualize complex biochemical processes.

WEB QUIZZES At www.oup.com/us/mckee7e, students seeking to test their knowledge of biochemistry can gain access to more than 600 questions written by Dan Sullivan (University of Nebraska at Omaha). Students receive a feedback summary with each graded quiz.

INTERACTIVE 3D MOLECULES At www.oup.com/us/mckee7e, students will find an updated selection of hundreds of interactive 3D molecules in JMOL format. Students can manipulate and study individual molecules and their structures, take self-guided concept tutorials, and test their understanding through interactive self-quizzes.

For Instructors

ANCILLARY RESOURCE CENTER The Ancillary Resource Center (ARC) at **www.oup-arc.com** is a convenient, instructor-focused single destination for resources that accompany the text. Accessed online through individual user accounts, the ARC provides instructors with up-to-date ancillaries at any time while guaranteeing the security of grade-significant resources. The following instructor's resources are available on the McKee ARC:

- **All text images in electronic format.** Instructors who adopt the seventh edition gain access to every numbered illustration, photo, figure caption, and table from the text in high-resolution electronic format. Labels have been enlarged, and multipart figures have been broken down into separate components for clearer projection in large lecture halls. Images are available on both the Instructor's Resource CD and the seventh edition website, www.oup.com/us/mckee7e.
- **Computerized test bank.** Written by the authors, the test bank includes more than 1,400 questions that can be edited and combined with an instructor's own questions to create customized quizzes and exams that can then be exported for use in learning management systems or printed for paper-based assessments

LECTURE NOTES SLIDES This set of more than 1100 editable lecture notes slides makes preparing lectures faster and easier than ever. Available in PowerPoint format, the lecture notes for the seventh edition now include embedded links from the new curated animations and video guide.

COURSE MANAGEMENT SYSTEMS All instructor and student resources, including the text images, the test bank files, and the online self-quiz questions, are compatible with a variety of management systems.

ACKNOWLEDGMENTS

We wish to express our appreciation for the efforts of the dedicated individuals who provided detailed content and accuracy reviews of the text and the supplemental materials of the seventh edition:

Shohreh Amini – Temple University
Samuel Besong – Delaware State University
Vijay Antharam – Methodist University
Curtis L Ashendel – Purdue University
Jennifer Chaytor – Saginaw Valley State University
Christopher Cottingham – University of North Alabama
Joe Deweese – Lipscomb University
Stephanie Dingwall – University of California, Riverside
Michael Gunther – West Virginia University
David Hawkinson – University of South Dakota
Angela Hoffman – University of Portland
Susanne Hoffmann-Benning – Michigan State University
Joseph Johnson – University of Minnesota Duluth

Marjorie A. Jones – Illinois State University
Eun Hoo Kim – Methodist University
Youngjoo Kim – SUNY College at Old Westbury
Bryan Knuckley – University of North Florida
Jason Kowalski – Marian University of Wisconsin
Colleen Krause – University of Hartford
George B. Kudolo – University of Texas Health Science Center at San Antonio
Chandrika Kulatilleke – The City University of New York – Baruch College and the Graduate Center
Karen Lewis – Texas State University
Arthur Lesk – Penn State University
Dennis Matlock – Harding University
Cornelia Mihai – Northwestern Oklahoma State University
Eugene Mueller – University of Louisville
Kyle Murphy – Rutgers University
Michelle Marie Osovitz – St. Petersburg College
Scott Pinkerton – Bellevue University
Gerry Prody - Western Washington University
Niina Ronkainen – Benedictine University
Michael Sehorn – Clemson University
Ivan Shibley – Penn State
Lisa Stephens – Marist College
Kari Stone – Benedictine University
Toni Trumbo Bell – Bloomsburg University of Pennsylvania
Craig Vander Kooi – University of Kentucky
Eric Wagner – Western Washington University
Thao Yang – University of Wisconsin-Eau Claire

We would also like to thank the individuals who reviewed the first six editions of this text:

Erika L. Abel – Baylor University
Gul Afshan – Milwaukee School of Engineering
Kevin Ahern – Oregon State University
Mark Annstron – Blackburn College
Josephine Arogyasami – Southern Virginia University
Curt Ashendel – Purdue University
Donald R. Babin – Creighton University
Stephanie Baker – Erksine College
Bruce Banks – University of North Carolina at Greensboro
Thurston Banks – Tennessee Technological University
Sandra L. Barnes – Alcorn State University
Ronald Bartzatt – University of Nebraska, Omaha
Deborah Bebout – The College of William and Mary

Werner Bergen – Auburn University
Steven Berry – University of Minnesota, Duluth
Allan Bieber – Arizona State University
Ruth E. Birch – Saint Louis University
Karl Bishop – Husson University
Albert Bobst – University of Cincinnati
Michael G. Borland – University of Cincinnati
Brenda Braaten – Framingham State College
John Brewer – University of Georgia
Martin Brock – Eastern Kentucky University
David W. Brown – Florida Gulf Coast University
Weiguo Cao – Clemson University
Edward J. Carroll Jr. – California State University, Northridge
Srikripa Chandrasekaran – Clemson University
Jiann-Shin Chen – Virginia Tech
Alice Cheung – University of Massachusetts, Amherst
Oscar P. Chilson – Washington University
Randolph A. Coleman – The College of William and Mary
Sean Coleman – University of the Ozarks
Kim K. Colvert – Ferris State University
Sulekha R. Coticone – Florida Gulf Coast University
Elizabeth Critser – Columbia College
Michael Cusanovich – University of Arizona
Anjuli Datta – Pennsylvania State University
Bansidhar Datta – Kent State University
Danny J. Davis – University of Arkansas
Patricia DePra – Carlow University
Siegfried Detke – University of North Dakota
William Deutschman – State University of New York, Plattsburgh
Matthew R. Dintzner – Western New England University
Robert P. Dixon – Southern Illinois University–Edwardsville
Patricia Draves – University of Central Arkansas
Lawrence K. Duffy – University of Alaska, Fairbanks
Charles Englund – Bethany College
Paula L. Fischhaber – California State University, Northridge
Nick Flynn – Angelo State University
Clarence Fouche – Virginia Intermont College
Thomas Frielle – Shippensberg University
Matthew Gage – Northern Arizona University
Paul J. Gasser – Marquette University
Eric R. Gauthier – Laurentian College
Frederick S. Gimble – Purdue University

Mark Gomelsky – University of Wyoming
George R. Green – Mercer University
Gregory Grove – Pennsylvania State University
Joseph Hajdu – California State University, Northridge
Marlin Halim – California State University, East Bay
James Hawker – Florida State University
Terry Helser – State University of New York, Oneonta
Kristin Hendrickson – Arizona State University
Tamara Hendrickson – Wayne State University
Pui Shing Ho – Oregon State University
Angela Hoffman – University of Portland
Charles Hosler – University of Wisconsin
Andrew J. Howard – Illinois Institute of Technology
Jeffrey Hoyt – Paradise Valley Community College
Christine A. Hrycyna – Purdue University
Holly Huffman – Arizona State University
Larry L. Jackson – Montana State University
Sajith Jayasinghe – California State University, San Marcos
John R. Jefferson – Luther College
Craig R. Johnson – Carlow University
Gail Jones – Texas Christian University
Ivan Kaiser – University of Wyoming
Michael Kalafatis – Cleveland State University
Peter Kennelly – Virginia Tech University
Barry Kitto – University of Texas, Austin
Paul Kline – Middle Tennessee State University
James Knopp – North Carolina State University
Vijaya L. Korlipara – Saint John's University
Christa R. Koval – Colorado Christian University
Allison Lamanna – Boston University
Hugh Lawford – University of Toronto
C. Martin Lawrence – Montana State University
Carol Leslie – Union University
Duane LeTourneau – University of Idaho
Robley J. Light – Florida State University
Rich Lomneth – University of Nebraska at Omaha
Maria O. Longas – Purdue University, Calumet
Cran Lucas – Louisiana State University–Shreveport
Jerome Maas – Oakton Community College
Arnulfo Mar – University of Texas–Brownsville
Larry D. Martin – Morningside College
Carrie May – University of New Hampshire

Dougals D. McAbee – California State University, Long Beach
Martha McBride – Norwich University
Kristi L. McQuade – Bradley University
Gary Means – Ohio State University
Alexander Melkozernov – Arizona State University
Joyce Miller – University of Wisconsin–Platteville
Robin Miskimins – University of South Dakota
David Moffet – Loyola Marymount University
Rakesh Mogul – California Polytechnic State University
Joyce Mohberg – Governors State University
Jamil Momand – California State University, Los Angeles
Bruce Morimoto – Purdue University
Kyle Murphy – Rutgers University
Alan Myers – Iowa State University
George Nemecz – Campbell University
Harvey Nikkei – Grand Valley State University
Michael Nosek – Fitchburg State University
Peter Oelkers – University of Michigan–Dearborn
Peter M. Palenchar – Villanova University
Treva Palmer – Jersey City State College
Ann Paterson – Williams Baptist College
Scott Pattison – Ball State University
Allen T. Phillips – Pennsylvania State University
Jerry L. Phillips – University of Colorado at Colorado Springs
Jennifer Powers – Kennesaw State University
Dominic F. Qualley – Berry College
Ramin Radfar – Wofford College
Rachel Roberts – Texas State University–San Marcos
Niina J. Ronkainen – Benedictine University
Abbey Rosen – University of Minnesota, Morris
Gordon Rule – Carnegie Mellon University
Tom Rutledge – Ursinus College
Ben Sandler – Ashford University
Richard Saylor – Shelton State Community College
Michael G. Sehorn – Clemson University
Steve Seibold – Michigan State University
Edward Senkbeil – Salisbury State University
Ralph Shaw – Southeastern Louisiana University
Andrew Shiemke – West Virginia University
Aaron Sholders – Colorado State University
Kevin R. Siebenlist – Marquette University
Vijay Singh – University of North Texas

Ram P. Singhal – Wichita State University
Deana J. Small – University of New England
Maxim Sokolov – West Virginia University
Madhavan Soundararajan – University of Nebraska
Salvatore Sparace – Clemson University
David Speckhard – Loras College
Narasimha Sreerama – Colorado State University
Ralph Stephani – St. John's University
Dan M. Sullivan – University of Nebraska, Omaha
William Sweeney – Hunter College
Blair R. Szymczyna – Western Michigan University
Christine Tachibana – Pennsylvania State University
John M. Tomich – Kansas State University
Anthony Toste – Southwest Missouri State University
Toni Trumbo-Bell – Bloomsburg University of Pennsylvania
Craig Tuerk – Morehead State University
Sandra L. Turchi-Dooley – Millersville University
Shashi Unnithan – Front Range Community College
Timothy Vail – Northern Arizona University
Harry van Keulan – Cleveland State University
Ales Vancura – Saint John's University
William Voige – James Madison University
Alexandre G. Volkov – Oakwood College
Justine Walhout – Rockford College
Linette M. Watkins – Southwest Texas State University
Terry Watt – Xavier University of Louisiana
Athena Webster – California State University, East Bay
Lisa Wen – Western Illinois University
Rosemary Whelan – University of New Haven
Kenneth O. Willeford – Mississippi State University
Alfred Winer – University of Kentucky
Beulah Woodfin – University of New Mexico
Kenneth Wunch – Tulane University
Ryan D. Wynne – St. Thomas Aquinas College
Les Wynston – California State University, Long Beach
Wu Xu – University of Louisiana at Lafayette
Laura S. Zapanta – University of Pittsburgh

We also wish to express our appreciation to Jason Noe, senior editor; Eric Sinkins, development editor; Katie Tunkavige, assistant editor; Tina Chapman, marketing manager; Petra Recter, director of content and digital strategy; and John Challice, vice president and publisher. The excellent efforts of the Oxford University Press production team are gratefully acknowledged. We are especially appreciative of the efforts of William Murray, senior production editor; Lisa Grzan,

content operations manager; and Michele Laseau, art director. We give a very special thank you to Susan Brown, whose consistent diligence on this project has ensured the accuracy of the text.

We wish to extend our deep appreciation to those individuals who have encouraged us and made this project possible: Ira and Jean Cantor and Josephine Rabinowitz. Finally we thank our son, James Adrian McKee, for his patience and encouragement.

Trudy McKee
James R. McKee

ABOUT THE AUTHORS

Trudy McKee is a biochemist who has taught Biochemistryat Thomas Jefferson University, Rosemont College, ImmaculataCollege, and the University of the Sciences.

James R. McKee is Professor of Chemistry at theUniversity of the Sciences.

Biochemistry

The Molecular Basis of Life

Common Abbreviations in Biochemistry

A	adenine
ACTH	adrenocorticotropic hormone
ACP	acyl carrier protein
ADP	adenosine-5′-diphosphate
AIDS	acquired immune deficiency syndrome
ALA	δ-aminolevulinate
AMP	adenosine-5′-monophosphate
ATP	adenosine-5′-triphosphate
BCAA	branched chain amino acids
BH_2	dihydrobiopterin (oxidized form)
BH_4	tetrahydrobiopterin (reduced form)
bp	base pair
BPG	2,3-bisphosphoglycerate
C	cytosine
CAP	catabolite gene activator protein
CDP	cytidine-5′-diphosphate
CMP	cytidine-5′-monophosphate
CTP	cytidine-5′-triphosphate
CoA or CoASH	coenzyme A
cAMP	adenosine-3′,5′-cyclic monophosphate
cGMP	guanosine-3′,5′-cyclic monophosphate
cyt	cytochrome
DAG	diacylglycerol
DHAP	dihydroxyacetone phosphate
DNA	deoxyribonucleic acid
ssDNA	single-stranded DNA
dsDNA	double-stranded DNA
DNase	deoxyribonuclease
DNP	2,4-dinitrophenol
EAA	essential amino acids
EF	elongation factor
EGF	epidermal growth factor
ER	endoplasmic reticulum
ESR	electron spin resonance
FAD	flavin adenine dinucleotide (oxidized form)
$FADH_2$	flavin adenine dinucleotide (reduced form)
fMet	N-formylmethionine
FMN	flavin mononucleotide (oxidized form)
G	guanine or *Gibbs free energy*
G protein	guanine-nucleotide binding protein
GH	growth hormone

GDP	guanosine-5′-diphosphate
GMP	guanosine-5′-monophosphate
GSH	glutathione
GSSG	glutathione (oxidized form)
GTP	guanosine-5′-triphosphate
Hb	hemoglobin
HDL	high-density lipoprotein
HETPP	hydroxyethyl-thiamine pyrophosphate
HGPRT	hypoxanthine-guaninephosphoribosyltransferase
HIV	human immunodeficiency virus
HMG-CoA	*β*-hydroxy-*β*-methylglutaryl-CoA
HPLC	high pressure liquid chromatography
HRE	hormone response element
hsp	heat shock protein
IF	initiation factor
IGF	insulinlike growth factor
IgG	immunoglobulin G
IL	interleukin
IMP	inosine-5′-monophosphate
IP_3	inositol-1,4,5-triphosphate
K_m	Michaelis constant
kb	kilobases
kDa	kilodalton
LDL	low-density lipoprotein
LHC	light harvesting complex
Man	mannose
NAA	nonessential amino acids
NAD^+	nicotinamide adenine dinucleotide (oxidized form)
NADH	nicotinamide adenine dinucleotide (reduced form)
$NADP^+$	nicotinamide adenine dinucleotide phosphate (oxidized form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
ncRNA	noncoding RNA
NDP	nucleoside-5′-diphosphate
NMR	nuclear magnetic resonance
NO	nitric oxide
nt	nucleotides
NTP	nucleoside-5′-triphosphate
P_i	orthophosphate (inorganic phosphate)
PC	plastocyanin
PDGF	platelet-derived growth factor
PEP	phosphoenolpyruvate
PFK	phosphofructokinase
PIP_2	phosphatidylinositol-4,5-bisphosphate
PP_i	pyrophosphate

PRPP	phosphoribosylpyrophosphate
PS	photosystem
PQ(Q)	plastoquinone (oxidized)
PQH_2 (QH_2)	plastoquinone (reduced)
RER	rough endoplasmic reticulum
RF	releasing factor
RFLP	restriction-frament length polymorphism
RNA	ribonucleic acid
dsRNA	double-stranded RNA
mRNA	messenger RNA
rRNA	ribosomal RNA
siRNA	small interfering RNA
snRNA	small nuclear RNA
ssRNA	single-stranded RNA
tRNA	transfer RNA
snRNP	small ribonucleoprotein particles
RNase	ribonuclease
S	Svedberg unit
SAH	S-adenosylhomocysteine
SAM	S-adenosylmethionine
SDS	sodium dodecyl sulfate
SER	smooth endoplasmic reticulum
SRP	signal recognition particle
T	thymine
THF	tetrahydrofolate
TPP	thiamine pyrophosphate
U	uracil
UDP	uridine-5′-diphosphate
UMP	uridine-5′-monophosphate
UTP	uridine-5′-triphosphate
UQ	ubiquinone (coenzyme Q)(oxidized form)
UQH_2	ubiquinone (reduced form)
VLDL	very low density lipoprotein
XMP	xanthosine-5′ monophosphate

Names and Abbreviations of the Standard Amino Acids

Amino Acid	Three-Letter Abbreviations	One-Letter Abbreviations
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G

Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

The Genetic Code

First position (5′ end)	Second Position: U	Second Position: C	Second Position: A	Second Position: G	Third position (3′ end)
U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA STOP UAG STOP	UGU Cys UGC Cys UGA STOP UGG Trp	U C A G
C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	U C A G
A	AUU Ile AUC Ile AUA Ile AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U C A G
G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U C A G

General and Organic Chemistry Review Primer

OUTLINE

GENERAL CHEMISTRY

Atomic Structure: The Basics
- Atomic Number and Mass Number
- Radioactivity
- Atomic Theory
- Electron Configuration in Atoms
- The Periodic Table

Chemical Bonding
- Lewis Dot Notation
- Molecular Structure
- Valence Bond Theory and Orbital Hybridization

Chemical Reactions
- Reaction Kinetics
- Chemical Reactions and Equilibrium Constants
- Acid–Base Equilibria and pH
- Reaction Types
- Measuring Chemical Reactions

ORGANIC CHEMISTRY

Hydrocarbons
- Cyclic Hydrocarbons
- Aromatic Hydrocarbons

Substituted Hydrocarbons
- Alcohols
- Aldehydes
- Ketones
- Carboxylic Acids
- Esters
- Ethers

Amines
Amides
Thiols
Organic Reactions: Substitutions and Eliminations
Substitution Reactions
Elimination Reactions

BIOCHEMISTRY COURSES ARE ALWAYS FAST-PACED AND CHALLENGING.

It is for this reason that success is highly dependent on a student's background in general and organic chemistry. Although courses in these subjects are prerequisites, students often have trouble recalling the detailed chemical information that will help them understand the chemical processes in living organisms. This review is divided into two sections: general chemistry and organic chemistry. General chemistry topics include atomic structure, chemical bonding, acids and bases, and the chemical properties of the principal elements found in living organisms. Topics in the organic chemistry section include the structure and chemical properties of carbon-containing compounds, nucleophiles and electrophiles, functional group structure and chemical behavior, and organic reaction classes. Topics that are directly relevant to biochemistry (e.g., biomolecule classes, pH, buffers, kinetics, and thermodynamics) are described within the textbook.

GENERAL CHEMISTRY

Chemistry is the investigation of matter and the changes it can undergo. Matter, which can be described as physical substances that occupy space and have mass, is composed of various combinations of the chemical elements. Each chemical element is a pure substance that is composed of one type of atom. About 98 of the 118 known elements occur on Earth, and an even smaller number occurs naturally in living organisms. These elements fall into three categories: *metals* (substances such as sodium and magnesium with high electrical and heat conductivity, metallic luster, and malleability), *nonmetals* (elements such as nitrogen, oxygen, and sulfur, which are defined as a group because of their lack of metallic properties), and *metalloids* (elements such as silicon and boron, which have properties intermediate between metals and nonmetals).

The review of general chemistry includes an overview of atomic structure, atomic electron configurations, the periodic table, chemical bonds, valence bond theory, chemical reaction types, reaction kinetics, and equilibrium constants.

Atomic Structure: The Basics

Atoms are the smallest units of an element that retain the property of that element. Atomic structure consists of a positively charged central nucleus surrounded by one or more negatively charged electrons. With the exception of the element hydrogen (H), the dense, positively charged nucleus contains positively charged protons and neutrons, which have no charge. (The hydrogen nucleus consists of a single proton.) Atoms are electrically neutral, so the number of protons is

equal to the number of electrons. When atoms gain or lose one or more electrons, they become charged particles called ions. Ions formed when atoms lose electrons, called *cations*, are positively charged because they have fewer electrons than protons. For example, when a sodium atom (Na) loses an electron, it becomes the positively charged ion Na^+. Ions formed by the gain of electrons, called *anions*, are negatively charged. Chlorine (Cl) gains an electron to form the chlorine ion Cl^-.

ATOMIC NUMBER AND MASS NUMBER Elements are identified by their atomic number and mass number. The *atomic number* of an element is the number of protons in its nucleus. The atomic number uniquely identifies an element. Carbon (C) has 6 protons in its nucleus, so its atomic number is 6. Any atom with 16 protons in its nucleus is an atom of sulfur (S).

The *mass number* of an element, measured in atomic mass units, is equal to the number of protons and neutrons. Calculating an element's mass number is complicated by the existence of *isotopes*, atoms of an element with the same number of protons but different numbers of neutrons.

Many naturally occurring elements exist as a mixture of isotopes. For example, carbon has three naturally occurring isotopes containing six, seven, and eight neutrons, called carbon-12, carbon-13, and carbon-14, respectively. Carbon-12, the most abundant carbon isotope, is used as a reference standard in the measurement of atomic mass. An atomic mass unit (μ) or dalton (Da), named after the chemist John Dalton (1766–1844), is defined as one twelfth of the mass of an atom of carbon-12. Because the isotopes of an element do not occur with equal frequency, the *average atomic mass unit* (the weighted average of the atomic masses of the naturally occurring isotopes) is used. For example, hydrogen has three isotopes: hydrogen-1, hydrogen-2 (deuterium), and hydrogen-3 (tritium), which contain zero, one, and two neutrons, respectively. The average atomic mass for hydrogen is 1.0078 μ. This number is very close to 1.0 because hydrogen-1 has an abundance of more than 99.98%.

RADIOACTIVITY Some isotopes are radioactive (i.e., they undergo *radioactive decay*, a spontaneous process in which an atomic nucleus undergoes a change that is accompanied by an energy emission). For example, relatively unstable carbon-14 undergoes a form of radioactive decay, referred to as β-decay. In β-decay, one neutron in the atom's nucleus is converted into a proton and an electron. The new proton converts the carbon-14 atom to a stable nitrogen-14 atom. The newly created electron is emitted as a β-particle. With its consistent decay rate and a half-life (p. 220) of 5700 years, carbon-14 is used in the radiocarbon dating of organic materials up to 60,000 years ago. Carbon-14 is also used as a radioactive tracer in biochemical and medical research.

ATOMIC THEORY According to the Bohr model of atoms, electrons are in circular orbits with fixed energy levels that occur at specific distances from the nucleus. When an atom absorbs energy, an electron moves from its "ground state" to a higher-energy level. The electron returns to its ground state when the atom releases the absorbed energy. As quantum theory revolutionized physics in the early twentieth century, it became apparent that the theory explained many properties of atoms that the Bohr model did not.

Quantum theory is based on the principle that both matter and energy have the properties of particles and waves. Using quantum theory, physicists and chemists eventually described an atomic model in which electrons are predicted to occur in complex orbitals that are essentially probability clouds. An orbital is a probability distribution (i.e., variations in an orbital's cloud density correlate with the probability of finding an electron). The different shapes and sizes of orbital clouds depend on the energy level of the electrons within them. Together, four quantum numbers describe the configuration of the electrons and the orbitals in an atom.

The *principal quantum number n* defines the average distance of an orbital from the nucleus where $n = 1, 2, 3$, and so on. In other words, the quantum number n designates the principal energy shell. The higher its n value, the farther an electron is from the nucleus.

The *angular momentum quantum number l* (lower-case L) determines the shape of an orbital. The l values of 0, 1, 2, 3, and 4 correspond to the s, p, d, and f subshells. Note that the value of n indicates the total number of subshells within the principal energy shell. So if $n = 3$, the atom's principal shell has three subshells with l values of 0, 1, and 2. In such an atom the principal energy shell would contain s, p, and d orbitals. Each subshell also has a specific shape. The s orbital is spherical with the nucleus at its center. Each p orbital is dumbbell-shaped, and each d orbital is double dumbbell-shaped. The shapes of f orbitals are increasingly more complex and are rarely relevant in biochemical reactions.

The *magnetic quantum number m* describes an orbital's orientation in space. Values of m range from −l to +l. With an s orbital, $l = 0$, so the value of m is 0. For p orbitals, the value of l is 1, so m is equal to −1, 0, or +1; that is, there are three orbitals designated p_x, p_y, and p_z (**Figure 1**). For d orbitals $l = 2$, so there are five possible orientations: –2, –1, 0, +1, or +2.

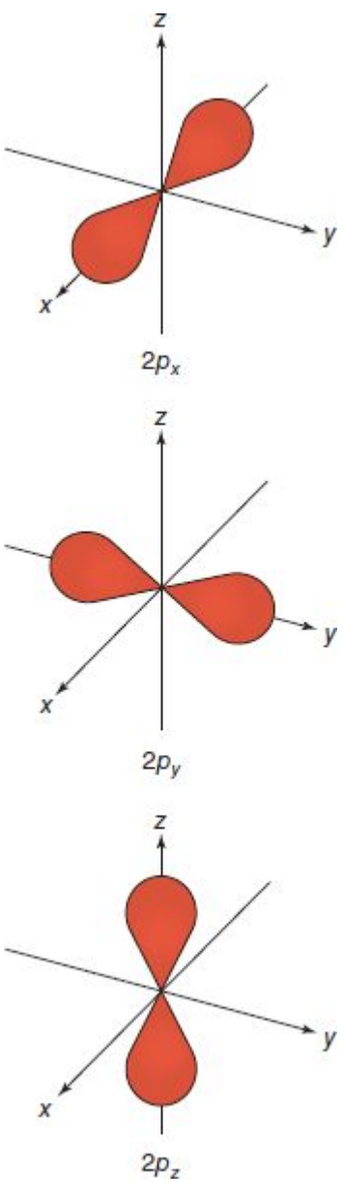


FIGURE 1

The 2p Orbitals

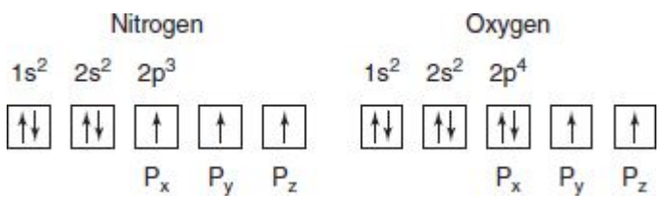
Three 2p orbitals are oriented at right angles to each other.

The fourth quantum number is the *spin quantum number* m_s, which describes the direction in which an electron is rotating: clockwise or counterclockwise. The values for m_s can be either +1/2 or −1/2. Because the *Pauli exclusion principle* states that each electron in an atom has a unique set of the four quantum numbers, it follows that when two electrons are in the same orbital, they must have opposite spins. Such spins are described as "paired." The spinning of an electron creates a magnetic field. *Diamagnetic* atoms such as nitrogen are not attracted to magnets because they have paired electrons (i.e., the magnetic fields of the paired electrons cancel out). Atoms that contain unpaired electrons (e.g., oxygen) are referred to as *paramagneti*c because they are attracted to magnets.

ELECTRON CONFIGURATION IN ATOMS Knowing how electrons are distributed in atoms is essential to any understanding of how chemical bonds are formed. There are several rules concerning electron distribution. The most basic rule is the *Aufbau principle*, which stipulates that electrons are put into orbitals, two at a time, in the order of increasing orbital energy (i.e., the inner orbitals are filled before the outer, higher-energy orbitals). Chemists use a shorthand method to illustrate how electrons are arranged around the nucleus of a ground-state atom (an atom at its lowest possible energy state). The electron configuration pattern useful for the elements relevant to living organisms is as follows: $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 5s^2 4d^{10} 5p^6 6s^2$.

The superscripts in the electron configuration pattern indicate the maximum number of electrons in each subshell. Note that because of orbital overlaps, the order of the orbitals being filled becomes more complicated as the filling pattern progresses. **Figure 2** is a diagram that will aid in recalling the order in which the subshells are filled.

Determining an element's electron configuration requires knowing its atomic number (the number of protons), which is also equal to the number of electrons. Using the electron configuration pattern, the element's electrons are then used to fill in the orbitals beginning with the lowest energy level. For example, the electron configurations of hydrogen (1 electron) and helium (2 electrons) are $1s^1$ and $1s^2$, respectively. Similarly, the electron configurations for carbon (6 electrons) and chlorine (17 electrons) are $1s^2 2s^2 2p^2$ and $1s^2 2s^2 2p^6 3s^2 3p^5$, respectively. According to *Hund's rule*, when an energy subshell has more than one orbital (e.g., p and d orbitals) there is only one electron allowed in each orbital until all the orbitals have one electron. Such electrons have parallel spins. As additional electrons enter the orbitals, they will spin pair with the previously unpaired electrons. The orbital diagrams of nitrogen and oxygen illustrate this rule.



For many elements, an electron configuration also reveals how many valence electrons there are. *Valence electrons*, the electrons in the s and p orbitals of the outermost energy level, determine the element's chemistry (i.e., how it will react with other elements). For example, oxygen atoms with an electron configuration of $1s^2 2s^2 2p^4$ have six valence electrons (i.e., there are a total of six electrons in its 2s and 2p orbitals). Chlorine has seven valence electrons because there are seven electrons in its 3s and 3p orbitals. For many elements, atoms will react so that their outermost

energy level or *valence shell* is filled, which is the most stable configuration they can have. The term *octet rule* is used to describe this phenomenon because the atoms of most elements react so that their valence shells contain eight electrons. Hydrogen and lithium are two obvious exceptions. Because a hydrogen atom only has one electron in its 1s orbital, it can gain one electron when it reacts to form a hydride ion (H^-) with a $1s^2$ orbital, or it can give up an electron to form a proton (H^+). With three electrons, lithium (Li) has a $1s^2 2s^1$ configuration. By losing its one valence electron, lithium atoms gain stability by having a filled 1s shell (two electrons). In the reaction of lithium with chlorine forming lithium chloride (LiCl), lithium gives up one electron to become a lithium ion (Li^+). The lithium valence electron is donated to chlorine to form the chloride ion (Cl^-). Chlorine has thereby increased its valence shell from seven to eight electrons. Understanding the significance of electron configurations, valence, and other properties of the elements is enhanced by familiarity with the periodic table of the elements, which is discussed next. It should be noted that the term *oxidation state* is often used in reference to atoms that have gained or lost electrons. The lithium ion, for example, has a +1 oxidation state, and the chloride ion has a −1 oxidation state.

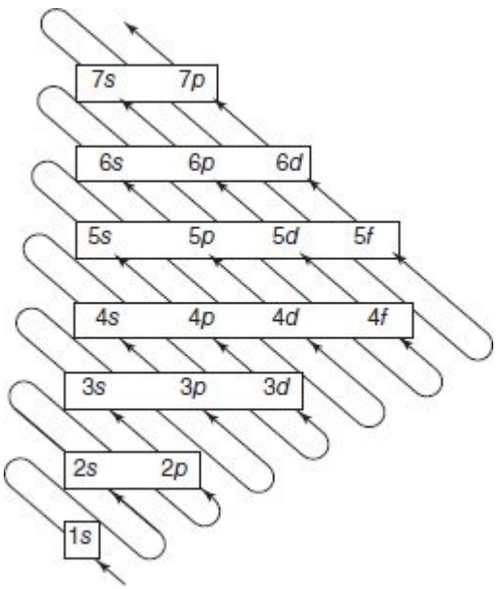


FIGURE 2

Subshell-Filling Sequence

All of the subshells of a given value of *n* are on the same horizontal line. The filling sequence is determined by following the arrows starting at the lower left.

THE PERIODIC TABLE The modern periodic table (**Figure 3**) is a chart based on the *periodic law*, which states that the electron configurations of the elements vary periodically with their atomic number. The properties of the elements that depend on their electronic configuration, therefore, also change with increasing atomic number in a periodic pattern. The periodic table is arranged in vertical rows called *groups* or *families* and horizontal rows called *periods*. Certain characteristics of the elements increase or decrease along the vertical or horizontal rows. These characteristics, which affect chemical reactivity, include atomic radius, ionization energy, electron affinity, and electronegativity. The *atomic radius* of a neutral atom is the distance from the nucleus to the outermost electron orbital. *Ionization energy* is defined as the amount of energy required to remove the highest energy electron from each atom in 1 mol of the atoms in the gaseous state (i.e., how strongly an atom holds on to its electrons). *Electron affinity* is the energy that is released

when an electron is added to an atom. *Electronegativity* is the tendency of an atom to attract electrons to itself. The trends of these properties are illustrated in **Figure 4**.

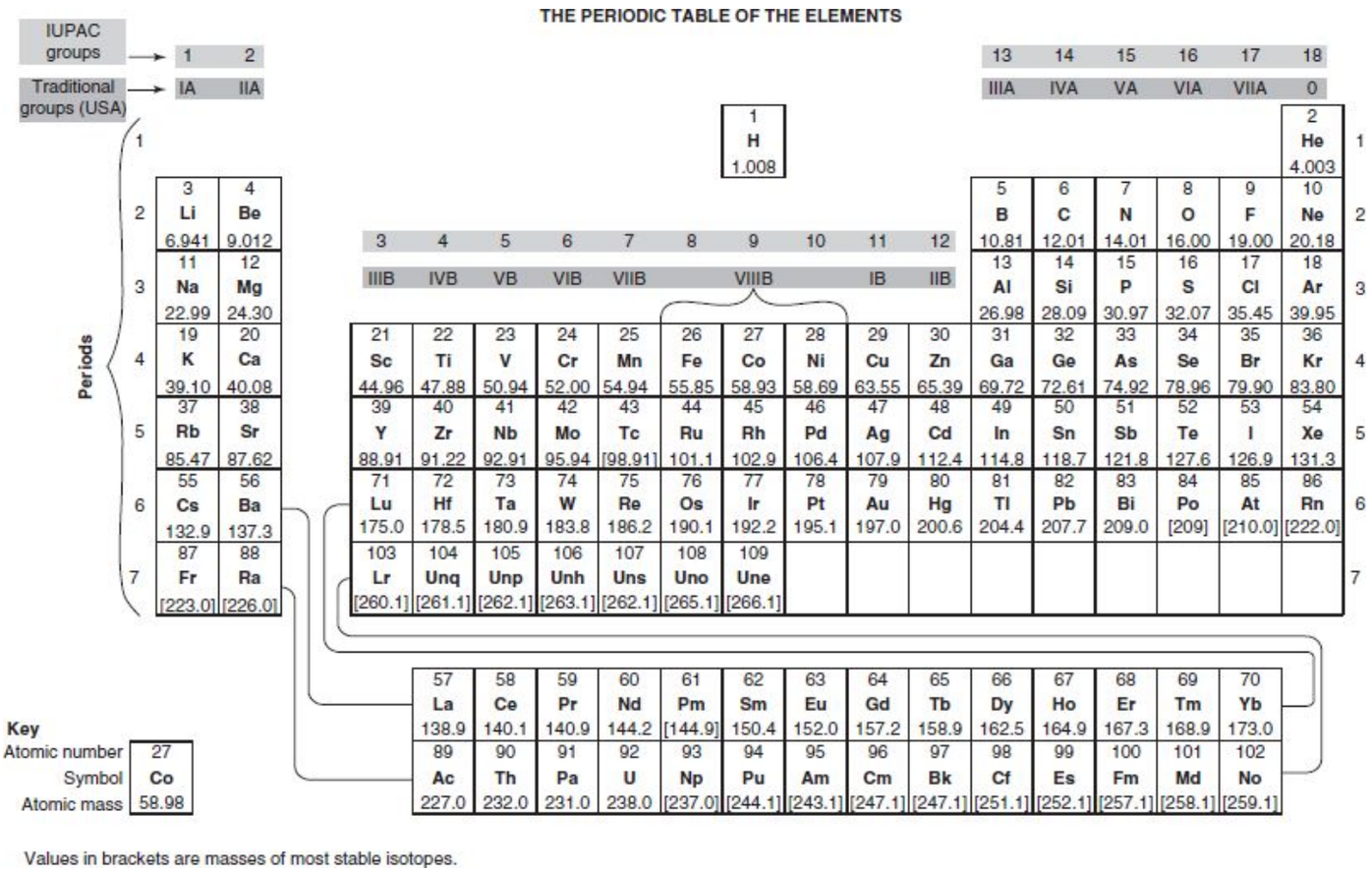


FIGURE 3

The Periodic Table

In the modern periodic table, elements are organized on the basis of their atomic numbers, electron configurations, and recurring chemical properties. Note that the Lanthanides (elements 57–70) and the Actinides (elements 89–102) are not relevant to biochemistry and are not discussed.

Each of the seven horizontal rows of the periodic table, called *periods*, begins with an element with a new shell with its first electron. For example, there is one electron in the 2s, 3s, and 4s subshells of lithium, sodium, and potassium (K), respectively. The atomic radii of the elements in groups 1, 2, and 13 to 18 decrease from left to right. As the number of positively charged protons at the center of the atom increases, the negatively charged electrons are attracted more strongly (i.e., the electrons are drawn closer to the nucleus). The same trend is not seen in the atomic radii of the elements in groups 3 to 12. There is little reduction in atomic size in these elements because of the repulsion between the 4s and 3d electrons.

The ionization energies of the elements in a period typically increase with increasing atomic number. As the atomic radii decrease across a period (i.e., the distance between the outer electron and the nucleus of the atoms decreases), more energy is required to remove the outer electron. For example, it is easier to remove an electron from lithium (atomic number = 3) than from nitrogen (N) (atomic number = 7).

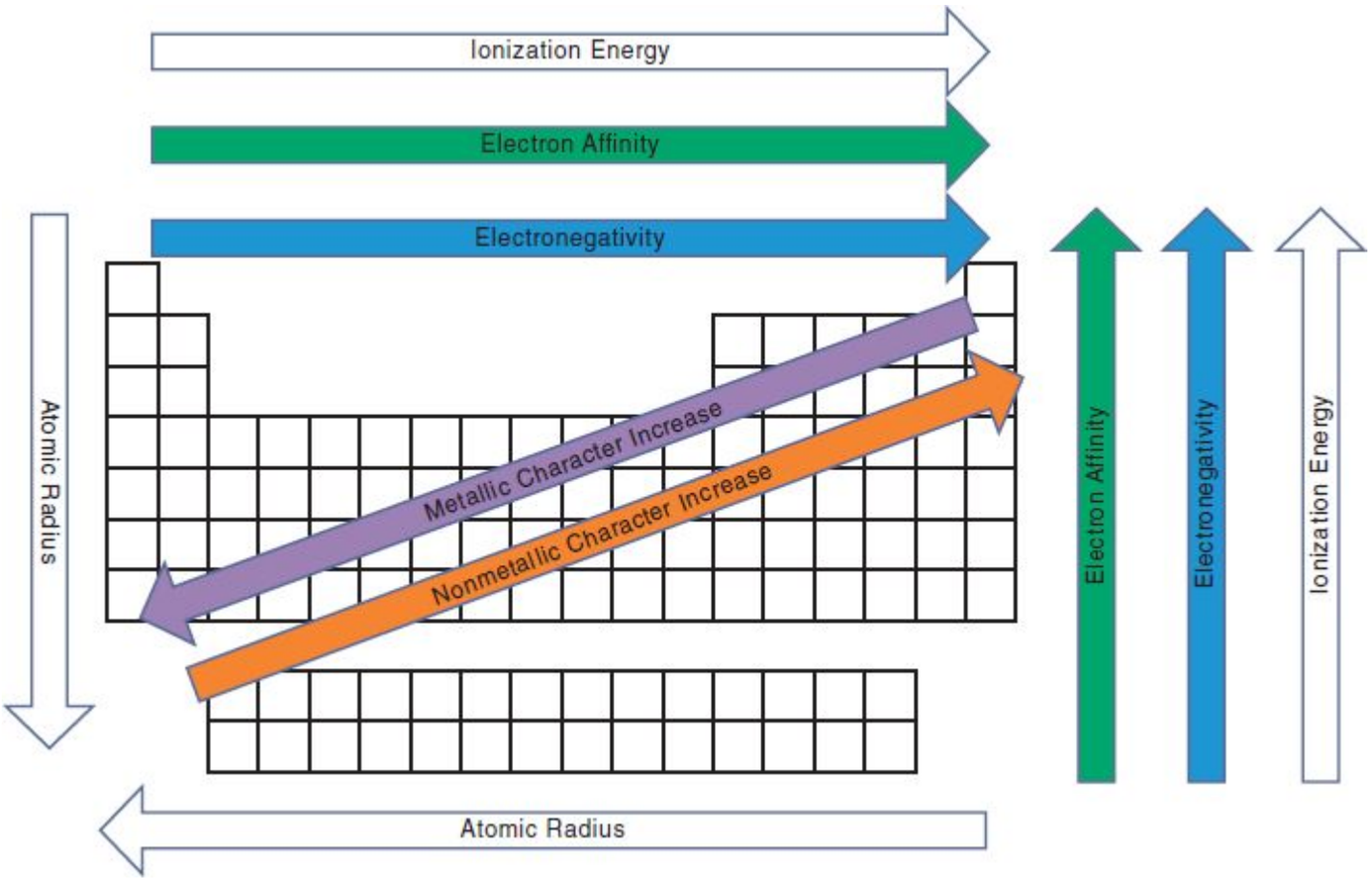


FIGURE 4

Periodic Table Trends

Investigations of the elements in the periodic table have revealed trends in certain properties. Among these are atomic radius (decreases across a period from left to right and increases down a group), ionization energy (increases from left to right across a period and increases up a group), electron affinity (increases from left to right across a period and increases up a group), and electronegativity (increases from left to right across a period and increases up a group). The metallic character of an element, defined as the readiness of an atom to lose an electron, also shows periodic trends. Metallic properties increase from right to left diagonally across periods. Nonmetallic properties (e.g., the ability to gain electrons) increase diagonally from left to right across periods.

Electron affinity, which can be thought of as the likelihood that a neutral atom will gain an electron, increases from left to right across periods. Metals such as sodium ($1s^2 2s^2 2p^6 3s^1$) have low electron affinities because they become more stable when they lose valence electrons. Elements on the right side of the periodic table have high electron affinities because of vacancies in their valence shells. Chlorine ($1s^2 2s^2 2p^6 3s^2 3p^5$) has a very high electron affinity because it releases a large amount of energy to become more stable as the chloride ion when it fills its valence shell by gaining one electron. The noble gases [e.g., helium (He), neon (Ne), and argon (Ar)] in group 18 do not conform to this trend because their valence shells are filled; hence, they are chemically unreactive.

Electronegativity, a measure of an atom's affinity for electrons in a chemical bond, increases across a period as atomic radii are decreasing. In the water molecule (H_2O), for example, oxygen is more electronegative than the hydrogen atoms because oxygen's larger nucleus strongly attracts its electrons. Hence, in water molecules the electrons in the bonds between each of the two hydrogen atoms and the oxygen atom are shared unequally.

The 18 vertical rows of the periodic table consist of elements with similar chemical and physical

properties. The group 1 elements all have one electron in their outermost shell. With the exception of hydrogen, all of the group 1 elements [lithium (Li), sodium (Na), potassium (K), rubidium (Rb), cesium (Cs), and the rare radioactive francium (Fr)] are referred to as the *alkali metals* because they react vigorously with water to form hydroxides (e.g., NaOH). The alkali metals do so because they readily lose their single valence electron to form cations with a +1 charge. For example, sodium ($1s^2 2s^2 2p^6 3s^1$) reacts with water to form sodium hydroxide (NaOH) and hydrogen gas (H_2). NaOH then dissociates to form Na^+ and OH^-. Because the alkali metals donate their valence electron so readily, they are considered especially strong reducing agents. (*Reducing agents* are elements or compounds that donate electrons in chemical reactions.) Of all the alkali metals, only sodium and potassium have normal functions in living organisms. For example, the balance of sodium ions and potassium ions across the plasma membrane of neurons is critical to the transmission of nerve impulses.

The group 2 alkaline earth metals [beryllium (Be), magnesium (Mg), calcium (Ca), strontium (Sr), barium (Ba), and radium (Ra)] have two electrons in their outermost shell. The electron configurations for the biologically important group 2 elements magnesium (DNA structure and enzyme function) and calcium (bone structure and muscle contraction) are $1s^2 2s^2 2p^6 3s^2$ and $1s^2 2s^2 2p^6 3s^2\ 3p^6 4s^2$, respectively. With the exception of beryllium, the alkaline earth metals lose their two valence electrons to form cations with a +2 charge [e.g., they react with water to form metal hydroxides such as $Ca(OH)_2$]. Like the group 1 metals, the alkaline earth metals are strong reducing agents, although each element is somewhat less reactive than the alkali metal that precedes it.

Groups 3 to 12 are referred to as the d-block elements because electrons progressively fill the d orbitals. The majority of the d-block elements are the *transition elements*, which have incompletely filled d orbitals. Zinc (Zn, atomic number = 30) is not considered a transition metal because its 3d subshell has 10 electrons. Because electron configurations of elements with high atomic numbers are unwieldy, chemists use a simplification for an element's electron configuration that is an abbreviation for the electron configuration of the noble gas immediately preceding the element. For example, zinc's electron configuration can be described as $[Ar]3d^{10}4s^2$.

The transition elements are metals with special properties. Among these are the capacities to have more than one oxidation state and to form colored compounds. Iron is a transition metal found in a large number of proteins in all living organisms. It most notably occurs in hemoglobin, the oxygen-transport protein that gives blood its red color. Iron atoms ($[Ar]3d^6 4s^2$) can form a wide range of oxidation states from −2 to +6, but its most common oxidation states are +2 and +3. Neutral iron atoms can form the +2 ion because the 4s orbital and the 3d orbitals have very similar energies, so that the removal of two electrons to form Fe^{+2} ($[Ar]3d^6$) requires little energy. The loss of an additional electron to form Fe^{+3} {$[Ar]3d^5$} requires more energy.

In addition to iron, several other d-block elements are important in living organisms. Manganese (Mn) is found in numerous enzymes in all living organisms. Cobalt (Co) is an important component of vitamin B_{12} structure (p. 545). Nickel (Ni) occurs in several enzymes in microorganisms and plants. Copper (Cu) is found in several energy generation proteins. Zinc (Zn) occurs in over 100 enzymes and has structural roles in numerous proteins. Molybdenum (Mo) has a vital role in nitrogen fixation.

The remaining elements in the standard periodic table, groups 13 to 18, are in the *p block*, so named because electrons progressively fill p orbitals. The p-block elements found in living organisms (carbon, nitrogen, phosphorus, sulfur, chlorine, and iodine) are nonmetals. Carbon

($1s^2 2s^2 2p^2$) in group 14 has four electrons available to form stable bonds both with other carbon atoms and with a variety of other elements (most notably hydrogen, oxygen, nitrogen, and sulfur). As a result, carbon can form an almost infinite number of compounds. Carbon is the crucial element in most biomolecules, with the exceptions of molecules such as water and ammonia and the electrolytes (e.g., Na^+, K^+, and Mg^{+2}). (An *electrolyte* is an ionic species that influences the distribution of electric charge and the flow of water across membranes.) Because nitrogen ($1s^2 2s^2 2p^3$) has five electrons in its outer shell, its valence is −3. In biomolecules, nitrogen is found in amines (R-NH_2, where R is a carbon-containing group) and amides (e.g., bonds between amino acids in proteins). Phosphorus ($1s^2 2s^2 2p^6 3s^2 3p^3$), in the nitrogen family, most commonly occurs in living organisms as phosphate (PO_4^{-3}) (e.g., in the nucleic acids DNA and RNA and as a structural component of bones and teeth). In living organisms, oxygen ($1s^2 2s^2 2p^4$) is found most abundantly in water molecules, where it has an oxidation state of −2. Oxygen atoms are also found in all the major classes of biomolecules (e.g., proteins, carbohydrates, fats, and nucleic acids). Sulfur ($[Ne]3s^2 3p^6$), the second member of the oxygen family, is found in proteins and small molecules such as the vitamin thiamine. It often occurs in biomolecules in the form of thiols (R-SH) and disulfides (R-S-S-R). Of all the members of the halogen family (group 17), only chlorine ($[Ne]3s^2 3p^5$) and iodine ($[Kr]5s^2\ 4d^{10} 5p^5$) routinely occur in living organisms. The functions of chlorine in the form of the chloride ion include the digestion of protein in the animal stomach [hydrochloric acid (HCl)] and its function as an electrolyte. Iodine is a component of the thyroid hormones, which regulate diverse metabolic processes in the animal body. It should be noted that phosphorus, sulfur, and chlorine are usually assigned valences of −3, −2, and −1, respectively. However, because they have vacant d orbitals, they can expand their valence shell and form different oxidation states. For example, phosphorus has a +5 valence in phosphoric acid (H_3PO_4), sulfur has a valence of +6 in the sulfate ion (SO_4^{2-}), and chlorine has a +1 valence in the hypochlorite ion (ClO^-).

WORKED PROBLEM 1

Consider the element potassium (atomic number = 19). What is its electronic configuration? How many electrons are in its valence shell? Is elemental potassium paramagnetic or diamagnetic?

SOLUTION

The number of electrons in the potassium atom is equal to its atomic number, 19. Using the Aufbau and Pauli exclusion principles and Hund's rule, the electronic configuration is $1s^2 2s^2 2p^6 3s^2 3p^6 4s^1$. Potassium has one valence electron because its outermost energy level (4s) has one electron. Elemental potassium is paramagnetic because it has one unpaired electron in its valence shell. ■

Chemical Bonding

A chemical bond is a strong, attractive force between the atoms in a chemical compound. Chemical

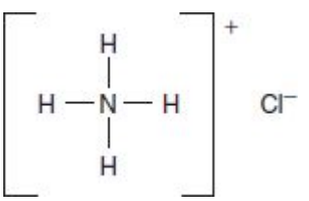
bonds form because of interactions between atoms that involve the rearrangement of their outer shell electrons. According to the octet rule, atoms react so as to achieve the outer electron configuration of the noble gases. Their complete outer valence shells are stable because of a reduction in stored potential energy.

Two major types of chemical bonds are ionic and covalent. The bonds differ in how valence electrons are shared among the bonded atoms. *Ionic bonds* form when electrons are transferred from atoms with a tendency to release electrons (e.g., alkali or alkaline earth metals) to electronegative atoms that tend to gain electrons. The transfer process results in the formation of oppositely charged atoms called *ions*. The positively charged ion product is called a *cation*, and the negatively charged ion product is called an *anion*. For example, the transfer of sodium's single electron to chlorine (seven valence electrons) yields the cation Na^+ and the anion Cl^-. The ionic bond in NaCl is the electrostatic attraction between the positive and negative ions.

In *covalent bonds*, electrons are shared between atoms with similar electronegativity values. A single covalent bond consists of two shared electrons. For example, there is one covalent bond in molecular hydrogen (H_2). The two hydrogen atoms, with one electron each, complete their valence shell by sharing their electrons. Elements such as carbon, nitrogen, and oxygen can form multiple covalent bonds. Carbon, for example, can form double and triple bonds. In the molecule ethylene, a carbon–carbon double bond involves the sharing of two sets of valence electrons. The triple covalent bond in molecular nitrogen (N_2) is an example of the sharing of three sets of valence electrons.

Covalent bonds between atoms with moderate differences in electronegativity are referred to as *polar covalent bonds*. In such bonds, the electrons are shared unequally, with the electron density shifted toward the atom with the greater electronegativity. The electrical asymmetry in such bonds causes one end of the molecule to possess a slightly negative charge and the other end a slightly positive charge. These partial charges are indicated by the lowercase Greek letter δ: δ^+ and δ^-. For example, in the water molecule H_2O the oxygen atom has a significantly larger electronegativity value than the hydrogen atoms. As a result, the electron pairs between oxygen and each of the hydrogen atoms are drawn closer to the oxygen atom. Each hydrogen atom has a partial positive charge (δ^+), and the oxygen has a partial negative charge (δ^-).

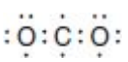
In a *coordinate covalent bond*, a shared pair of electrons in the bond comes from one atom. The reaction between ammonia (NH_3) and HCl provides a simple example. The product ammonium chloride (NH_4Cl) results when a covalent bond is formed between the nitrogen with its lone pair and the proton that has dissociated from HCl.



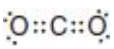
LEWIS DOT NOTATION Chemists often describe chemical bonds using Lewis dot structures. Devised by the chemist G. N. Lewis (1875–1946), Lewis dot structures are a shorthand notation for explaining how the valence electrons of the atoms in various compounds combine to form covalent bonds. Molecular hydrogen (H_2) is a simple example. Because each hydrogen atom has one electron, the Lewis dot structure for the hydrogen atom is H·. The energy shell of H_2 has a maximum of two electrons, and the formation of H_2 is depicted as

The following rules facilitate the drawing of Lewis structures for more complicated molecules:

1. **Determine the number of valence electrons for each atom in the molecule.** For example, carbon dioxide (CO_2) has one carbon atom with four valence electrons and two oxygen atoms, each with six valence electrons.
2. **Determine the identity of the central atom in the Lewis structure.** This atom will often be the one with the lowest electronegativity. Recall that electronegativity decreases from right to left across and from top to bottom of the periodic table. In the case of CO_2, carbon is less electronegative than oxygen, so carbon is the central atom.
3. **Arrange the electrons so that each atom donates one electron to a single bond between it and another atom and then count the electrons around each atom.** Are the octets complete? For CO_2, a first try would yield

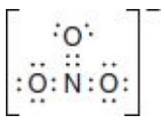


Note, however, that in this structure each carbon only has six electrons and each oxygen atom has seven electrons. Because the octets are incomplete, more electrons must be shared, an indication that there are double or triple bonds in the molecule. In the case of CO_2, rearranging the electrons results in the following Lewis dot structure in which there are two double bonds and all three atoms have a full octet of electrons.

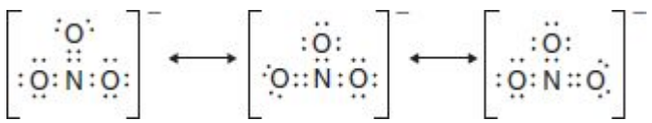


In molecules such as ammonia (NH_3), it is obvious that although the nitrogen atom is more electronegative than the hydrogens, it is the only atom in the molecule that can form multiple bonds. Hence, the nitrogen atom is the central atom in ammonia molecules. Its Lewis dot structure is

For a large number of molecules, there is more than one valid Lewis dot structure. The nitrate ion (NO^-_3) is a typical example. Inasmuch as nitrogen has five valence electrons and each oxygen atom has six valence electrons, the following Lewis dot structure of the nitrate ion satisfies the octet rule:



However, there is no reason why the double bond should appear where it does in this formula. It could easily appear in either of the two other locations around the nitrogen atom. Therefore, there are three valid Lewis dot structures for the nitrate ion.



When this situation occurs, the ion or molecule is said to be a *resonance hybrid*. (The double-headed arrows are used in the representation of resonance structures.) In the case of the nitrate ion, it is considered to have a structure that is the average of these three states.

MOLECULAR STRUCTURE Molecules are three-dimensional arrangements of atoms. Understanding molecular structure, also referred to as molecular shape, is important because structure provides insight into the physical and chemical properties of molecules. Physical properties that are affected by molecular shape include boiling point, melting point, and water solubility. The shape of molecules also powerfully affects chemical reactivity.

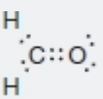
According to the *valence shell electron pair repulsion* (VSEPR) *theory*, repulsive forces between valence shell bonding and nonbonding electrons (lone pairs) determine molecular geometry (molecular shape). In other words, the valence electron pairs on the central atom in a molecule orient themselves in space so that repulsion is minimized (i.e., their total energy is minimized). Lone pairs of electrons have a greater repulsive effect than bonding pairs. (A *lone pair* is a valence electron pair on a central atom that is not involved in bonding.) The term *electron group* is used in discussions of VSEPR theory. An *electron group* is defined as a set of valence electrons in a region around a central atom that exerts repulsion on other valence electrons. Electron groups include bonding and nonbonding electron pairs or the pairs of electrons in double or triple bonds.

WORKED PROBLEM 2

What is the Lewis electron dot formula for formaldehyde ($H_2C{=}O$)?

SOLUTION

1. Calculate the number of valence electrons for each atom in the molecule. The valence electrons for hydrogen, carbon, and oxygen are 2 (one for each atom), 4, and 6, respectively, for a total of 12 electrons. Single bonds between the elements account for 6 electrons, leaving 6 electrons unaccounted for.
2. Group the remaining 6 electrons around the most electronegative atom (oxygen) until a total of 8 electrons (bonding and nonbonding) is reached. Using one pair of these electrons to form a double bond between carbon and oxygen completes the carbon octet. The final Lewis structure is given below.



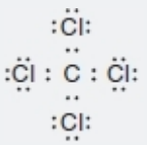
■

WORKED PROBLEM 3

Draw the Lewis dot structure for the organic solvent molecule carbon tetrachloride (CCl_4).

SOLUTION

1. Calculate the number of valence electrons for each atom in the molecule. The carbon atom has 4 valence electrons and each of the 4 chloride atoms have 7 valence electrons for a total of 32 valence electrons.
2. Determine the number of bonding electrons. Since there are four chlorines bonded to the central carbon atom, there are eight bonding electrons. (Carbon is the central atom because it has the lowest electronegativity.)
3. Calculate the number of remaining electrons and distribute them about each of the chloride atoms so that each atom in the molecule has an octet. Subtracting 8 bonding electrons from 32 valence electrons yields 24 electrons. The remaining 24 electrons are distributed around the 4 chlorine atoms as lone pairs. The final Lewis structure is given as follows.



■

Ascertaining a molecule's three-dimensional shape begins with a correct Lewis dot structure. The molecule's geometry is then determined based on the number of bonding and nonbonding electrons on the central atom (**Figure 5**). If there are two electron pairs, the molecule has a linear shape. Carbon dioxide (CO_2), for example, is a *linear* molecule with two electron groups. Its bond angle is 180^o. Formaldehyde ($H_2C{=}O$), with three electron groups, has *trigonal planar* geometry with bond angles of 120°. Molecules with a central atom with four pairs of electrons have a tetrahedral shape. Methane (CH_4), with its four carbon–hydrogen bonds, has bond angles of 109.5°. If one of the four electron groups in a tetrahedron is a lone pair, the molecular shape is *trigonal pyramidal*. Because of the strong repulsion of the lone pair, bond angles are less than 109.5°. For example, the lone pair in NH_3 forces the NH bonding electron pairs closer together with bond angles of 107.3°.

Three-dimensional shape also affects molecular polarity. In polar covalent bonds, there is an unequal sharing of electrons because the atoms have different electronegativities. This separation of charge is called a *dipole*. Although a polar molecule always contains polar bonds, some molecules with polar bonds are nonpolar. Molecular polarity requires an asymmetric distribution of polar bonds. For example, CO_2 contains two C—O dipole bonds. Carbon dioxide is a nonpolar molecule because of its linear shape (i.e., its bond dipoles are symmetrical and cancel each other out). Water, which also has two polar bonds (two O—H bonds), is a polar molecule because of its geometry. Water contains four electron groups: two bonding pairs and two lone pairs. As a result of the greater repulsions from the lone pairs on the oxygen, however, the bond angle of a water molecule is 104.5^o. Water's "bent" geometry (refer to **Figure 5**) makes it an asymmetric molecule and therefore polar.

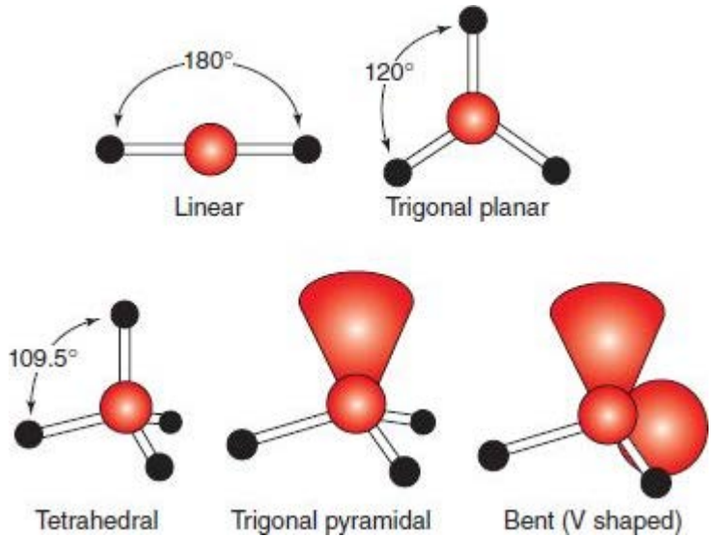


FIGURE 5

Common Molecular Geometrics

These structures illustrate the spatial orientations of electron groups. Note that lone pair electrons are indicated by an enlarged representation of an orbital.

WORKED PROBLEM 4

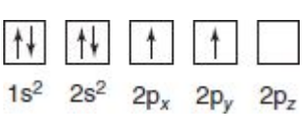
Dimethyl ether has the following formula: $CH_3—O—CH_3$. What is the shape of dimethyl ether? Does this molecule have a dipole moment?

SOLUTION

The oxygen in dimethyl ether has four electron pairs, including two lone electron pairs. As a result, the oxygen has a tetrahedral shape with an overall bent shape for the dimethyl ether molecule. Because the molecule has a bent shape and the electron distribution is uneven, dimethyl ether has a dipole moment. ■

VALENCE BOND THEORY AND ORBITAL HYBRIDIZATION Although the VSEPR theory accounts for molecular shape, it does not explain how the orbitals of the individual atoms interact to form the covalent bonds in molecules. The concept of *orbital hybridization*, the result of quantum-mechanical calculations, explains how the mixing of atomic orbitals results in the formation of the more stable hybrid orbitals found in molecules. Each type of hybrid orbital corresponds to a type of electron group arrangement predicted by VSEPR theory. The three most common hybrid orbitals observed in biomolecules are sp^3, sp^2, and sp.

Carbon has an electron configuration of $1s^2 2s^2 2p^2$, which can also be represented as



It appears from this diagram that carbon only has two bonding electrons. The carbon atoms in molecules such as methane, however, are bonded to four hydrogen atoms in a tetrahedral

arrangement. During methane formation, as a result of the attraction of each of the hydrogen nuclei (i.e., protons) for carbon's lower-energy valence electrons, the two 2s electrons move into 2p orbitals.

As they do so, they mix, forming four identical sp^3 orbitals (**Figure 6**).

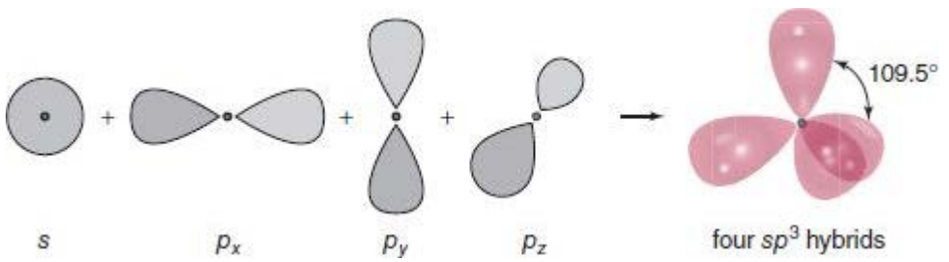


FIGURE 6

sp^3 Orbitals

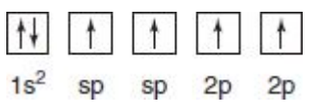
Hybridization of an s orbital and all three p orbitals gives four identical sp^3 orbitals.

In the methane molecule (**Figure 7**), each of the four sp^3 hybrid orbitals overlaps with the 1s orbital of hydrogen to form a sigma bond. A *sigma bond* (σ), which is formed by the overlapping by the outermost orbitals of two atoms, is the strongest type of covalent bond.

Each of the two carbon atoms in the molecule ethene ($H_2C{=}CH_2$) is bonded to three atoms in trigonal planar geometry. Carbon's 2s orbital mixes with two of the three available 2p orbitals to form three sp^2 orbitals.

Two of the three sp^2 orbitals of each carbon atom overlap the orbital of a hydrogen atom, forming a total of four σ bonds. The third sp^2 orbital of the two carbon atoms overlaps to form a carbon–carbon σ bond. The p orbitals, one on each carbon, overlap to form a pi (π) bond (**Figure 8**). A *double bond* in molecules, such as ethene, consists of a σ bond and a π bond.

Acetylene (C_2H_2) is a molecule with a triple bond, with each carbon bonded to two other atoms in a linear geometry. Carbon's 2s orbital



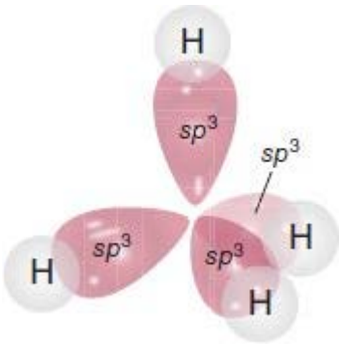


FIGURE 7

Structure of Methane

Methane (CH_4) has a tetrahedral geometry with four σ bonds formed by the overlap of four sp^3 orbitals of carbon with four 1s orbitals of hydrogen atoms.

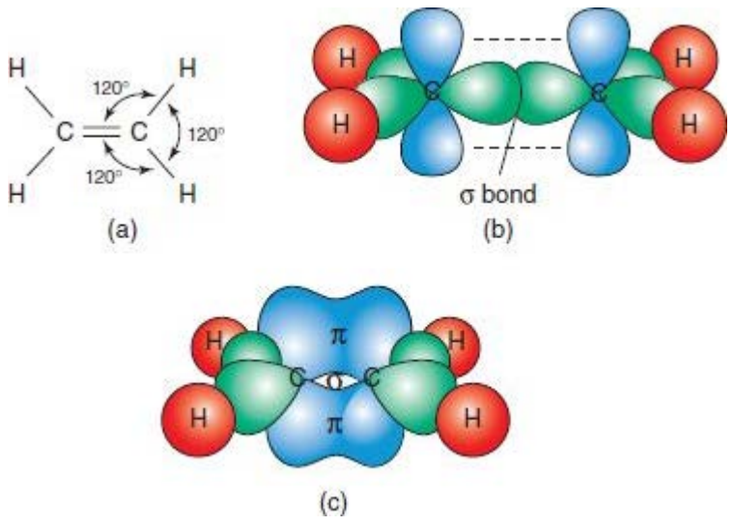


FIGURE 8

Ethene Structure

Each carbon atom in ethene (also known as ethylene) has three sp^2 orbitals with bond angles of 120°, which have a trigonal planar geometry. (b) Two of the sp^2 orbitals of each carbon (green) overlap with an s orbital of hydrogen (red), forming a total of 4 σ bonds. The remaining two sp^2 orbitals, one from each carbon, overlap to form a carbon–carbon σ bond. (c) Two p orbitals (blue), one from each carbon atom, overlap to form a π bond.

mixes with one 2p orbital to form 2 sp hybrid orbitals. Each carbon also possesses two unhybridized 2p orbitals. Acetylene has a triple bond consisting of one σ bond and two π bonds. The carbon–carbon σ bond is formed by the overlap of an sp hybrid orbital from each carbon atom. Each π bond is formed by the overlap of two of carbon's 2p orbitals. Each of the two carbon–hydrogen σ bonds is formed by the overlap of carbon's other sp orbital with the 1s orbital of hydrogen.

Chemical Reactions

In chemical reactions, the atoms in chemical substances are rearranged to form new substances as chemical bonds are broken and formed. According to the *collision theory*, the reaction rate in bimolecular reactions depends on the frequency of successful collisions between the chemical

species. Successful collisions occur when there is sufficient energy at the moment of impact (called the *activation energy*) and the colliding species are oriented during the collision in a manner that favors the rearrangement of the atoms and electrons. *Catalysts* are substances that increase reaction rates without being affected by the reaction. They do so by lowering the activation energy of the reaction by providing an alternative pathway for the reaction. For example, the metal iron is used as a catalyst in the Haber process, the industrial method of converting nitrogen and hydrogen gases into ammonia (NH_3). As the N_2 and H_2 molecules become adsorbed onto the surface of the metal, where they are more likely to be in a favorable orientation for successful collisions, the bonds in both molecules are weakened. Once formed, the ammonia molecules desorb from the catalyst.

Chemical reactions are described with chemical equations. The substances undergoing the reaction, called *reactants*, appear on the left-hand side of the equation, and the *products* of the reaction are on the right-hand side. An arrow between the reactants and products symbolizes the chemical change that occurs as a result of the reaction. For example, the chemical equation for the reaction in which molecule A reacts with molecule B to form molecules C and D is

$$A + B \rightarrow C + D$$

Other symbols that may appear in a chemical equation provide information about the physical state of the reactants and products or a required energy source. The equation for the decomposition of calcium carbonate, for example, is

$$CaCO_3\,(s) \xrightarrow{\Delta} CaO\,(s) + CO_2\,(g)$$

In this equation, the letter s indicates that the reactant $CaCO_3$ and the product CaO (calcium oxide) are solids. The letter g indicates that CO_2 is a gas. The uppercase Greek letter delta (Δ) above the arrow indicates that the reaction requires an input of energy in the form of heat. Reactions that require an input of energy in the form of heat are described as *endothermic*. If light energy is involved in a reaction, hν (ν is the lowercase Greek letter nu) is placed above the arrow.

All chemical equations must obey the *law of conservation of matter*, which states that during chemical reactions mass is neither created nor destroyed. In other words, the mass of the reactants must be equal to the mass of the products. For example, in the equation for the reaction in which methane (CH_4) reacts with molecular oxygen (O_2) to form carbon dioxide (CO_2) and water, the number of each type of atom on both sides of the arrow must be equal.

$$CH_4 + 2O_2 \rightarrow CO_2 + 2\,H_2O$$

In this balanced equation, the same number of carbon, hydrogen, and oxygen atoms is on both sides of the arrow because the number 2 has been placed before the formulas for molecular oxygen and water.

WORKED PROBLEM 5

Consider the following reaction equation:

$$KClO + H_2S \rightarrow KCl + H_2SO_4$$

Balance the equation and identify the elements that are oxidized, reduced, or unchanged by the reaction. Use the generalizations that the oxidation state of hydrogen and group 1 metals is +1 and that of oxygen is −2.

SOLUTION

Balancing a chemical reaction equation requires that the number and types of atoms be the same. To satisfy this requirement, the number 4 is placed before the reactant KClO and the product KCl. The reaction becomes

$$4KClO + H_2S \rightarrow 4KCl + H_2SO_4$$

Using the oxidation state information given above, oxidation numbers are assigned to each element:

$$K(+1)Cl(+1)O(-2) + H_2(+1)S(-2) \rightarrow K(+1)Cl(-1) + H2(+1)S(+6)O_4(-2)$$

Sulfur is the element that is oxidized (i.e., its oxidation number increases from −2 to +6). Chlorine is reduced because its oxidation number is lowered from +1 to −1. Elements whose oxidation numbers remain unchanged are hydrogen, potassium, and oxygen. ■

REACTION KINETICS As informative as a chemical reaction's equation is, it reveals nothing about several important properties of the reaction: (1) How fast does the reaction occur? (2) When the reaction ends, what will be the ratio of product to reactant molecules? (3) Does the reaction require or release energy? The science of chemical kinetics seeks to answer these and other questions about chemical reaction rates (i.e., the change in the number of product and reactant molecules as the reaction progresses).

Reaction rate is defined as the change in the concentration of reactant or product per unit time. For the general reaction

$$aA + bB \rightarrow cC + dD$$

the rate is equal to $k[A]^m[B]^n$, where k is the rate constant and [A] and [B] are the concentrations of the reactants A and B, respectively. The exponents *m* and *n* are used to determine the reaction's order, a number that relates the rate at which a chemical reaction occurs to the concentrations of the reactants. For example, if *m* is equal to 1, the reaction rate doubles when the concentration of reactant A doubles. If *m* is equal to 2, the rate quadruples when the concentration of reactant A doubles. (Refer to pp. 219–20 for a more detailed description of reaction order.) The rate constant and order of a reaction can only be determined by experiment. Experiments performed over the course of the past century have revealed that the following factors influence reaction rate:

1. **Reactant structure.** The nature and the strength of chemical bonds affect reaction rates. For example, salt formation, the exchange of ions, is a fast process compared with the breaking and forming of covalent bonds.
2. **Reactant concentration.** The number of molecules of a substance per unit volume affects the likelihood of collisions. Reaction rates increase as the reactant molecules become more crowded.
3. **Physical state.** Whether the reactants are in the same phase (solid, liquid, or gas) affects

reaction rates because reactants must come into contact with each other. When reactants are in the aqueous phase, for example, thermal motion brings them into contact. When reactants are in different phases, contact only occurs at the interface between the phases. In such circumstances, increasing interface surface area raises a reaction's rate. For example, when reactants are in the solid and liquid phases, grinding the solid into small pieces increases its surface area that is in contact with the liquid phase.

4. **Temperature.** At higher temperatures, molecules have more thermal energy and are, therefore, more likely to collide with each other.
5. **Catalysts.** Catalysts are substances that accelerate a reaction's rate but remain unchanged afterward. A catalyst provides a different pathway for the reaction, thereby lowering the activation energy.

CHEMICAL REACTIONS AND EQUILIBRIUM CONSTANTS Many chemical reactions are reversible (i.e., they occur in both forward and reverse directions). Reversible reactions are indicated in reaction equations with double arrows. When a reversible reaction begins (i.e., when the reactants are mixed together), reactants begin to be converted into product. At some point in time, which differs for each reaction, some product molecules are reconverted back into reactant molecules. Eventually, the reaction reaches a dynamic equilibrium state in which both the forward and reverse reactions occur, but there is no net change in the ratio of reactant and product molecules. The extent to which the reaction proceeds to product is measured by an equilibrium constant (K_{eq}), which reflects the concentrations of reactants and products under specific conditions of temperature and pressure. For a reaction with the equation

$$aA + bB \rightleftharpoons cC + dD$$

K_{eq} is calculated as the ratio of the molar concentrations of product and reactant, each of which is raised to the power of its coefficient.

$$K_{eq} = \frac{[C]^c\,[D]^d}{[A]^a\,[B]^b}$$

Note that K_{eq} is also equal to k_f/k_r, the ratio of the forward and reverse rates of the reaction. A high K_{eq} value (significantly greater than 1) indicates that when a reaction reaches equilibrium, the concentration of the reactant is low (i.e., the reaction favors the production of product). If the K_{eq} value is lower than 1, then product concentration is lower than reactant concentration when equilibrium has been reached. When K_{eq} is greater than 1000, the reaction has gone to completion (i.e., almost all reactants have been converted to product).

In 1885, the French chemist Henri Louis Le Chatelier reported his discovery of a remarkable feature of systems at equilibrium. For a chemical reaction at equilibrium, a change in the conditions of the reaction (e.g., temperature, pressure, or the concentrations of its components) triggers a shift in the equilibrium to counteract the change. Chemists and chemical engineers use *Le Chatelier's principle* to manipulate chemical reactions to maximize product synthesis. The Haber–Bosch process for making ammonia (NH_3) from N_2 and H_2 is a prominent example.

All living organisms require a source of usable nitrogen-containing molecules. As a result of the extraordinary difficulty in breaking the stable triple bond of N_2, *nitrogen fixation* (the conversion of N_2 to NH_3, a molecule that can be assimilated into organic molecules such as amino acids) is largely limited to a select group of microorganisms. Note that the synthesis of ammonia is an

exothermic reaction (i.e., it releases heat energy):

$$N_2\,(g) + 3H_2\,(g) \rightarrow 2NH_3\,(g) + 98\ kJ$$

where a joule (J) is a unit of energy and a kilojoule (kJ) is 1000 joules.

The Haber–Bosch industrial process for synthesizing ammonia maximizes the reaction's yield in several ways:

1. An iron-based catalyst (iron oxide with small amounts of other metal oxides), which increases the rate at which equilibrium is attained, converts a slow reaction to one that is fast enough to be commercially feasible.
2. Ammonia, the product of the reaction, is removed from the reaction vessel. As a result, the system produces more NH_3 to reestablish equilibrium.
3. An increase in the pressure within the reaction vessel (to 200 atm), obtained by decreasing volume, causes an increase in ammonia synthesis. Note that in this reaction 4 mol of reactant molecules are converted to 2 mol of product. The equilibrium shifts toward ammonia synthesis because there are fewer molecules of this gas.
4. By lowering the temperature of the reaction (i.e., by removing heat from an exothermic reaction), the equilibrium is shifted toward more ammonia synthesis. There is a limit to how much the temperature can be lowered, however, because the catalyst requires heat to be efficient. As a result, the reaction vessel operates at 400°C, a temperature that is hot enough for the catalyst, yet relatively cool for an industrial process.

ACID–BASE EQUILIBRIA AND pH When acids and bases dissolve in water, they dissociate, forming ions. Hydrochloric acid and acetic acid (CH_3COOH) are two well-known acids. HCl dissociates in water to yield chloride and hydrogen ions, and acetic acid dissociates to yield acetate (CH_3COO^-) and hydrogen ions. Sodium hydroxide (NaOH) and methylamine (CH_3NH_2) are examples of bases. In water, NaOH dissociates to yield sodium and hydroxide ions, and methylamine forms methylammonium ($CH_3NH_3^+$) and hydroxide (−OH) ions. The strength of an acid or a base is determined by the degree to which it dissociates. HCl is a strong acid because its dissociation in water is complete (i.e., 100% of HCl molecules dissociate into chloride and hydrogen ions). *Weak acids* and weak bases are so named because they dissociate only to a limited extent and establish a dynamic equilibrium with their ions. The general equation for the dissociation of a weak acid is

$$HA \rightleftharpoons A^- + H^+$$

where HA is the undissociated acid and A^- is the conjugate base of the acid.

The degree to which a weak acid dissociates is expressed as an acid dissociation constant K_a, the quotient of the equilibrium concentrations of the ions A^- and H^+ and the undissociated acid (HA).

$$K_a = \frac{[A^-][H^+]}{[HA]}$$

The dissociation constants of weak acids and bases are usually expressed as the negative log of the equilibrium constant (–log K_a or –log K_b), where the term –log is replaced by the letter p. The extent to which a weak acid dissociates is referred to as its pK_a value. The dissociation constant

and pK_a for acetic acid at 25°C, for example, are 1.8×10^{-5} and 4.76, respectively. The behavior of weak acids and bases is especially important in biochemistry because many biomolecules possess carboxylate, amino, and other functional groups that can accept or donate hydrogen ions. For example, refer to pp. 142–43 for a description of the effect of hydrogen ion concentration on amino acids, the molecules used to construct proteins.

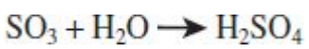
Water also has a slight capacity to dissociate into ions.

The hydrogen ion concentration of pure water at 25°C is 1.0×10^{-7} M. Because one hydroxide ion is produced for each hydrogen ion, the hydroxide ion concentration is also 1.0×10^{-7} M. The product of these two values (i.e., $[H^+][OH^-]$), referred to as the ion product for water, is 1×10^{-14}. The concentrations of hydrogen and hydroxide ions change depending on the substances that are dissolved in water, but their product is always 1×10^{-14}.

For weak acids and bases, hydrogen ion concentrations in aqueous solution can vary from 1 M to 1×10^{-14} M. For the sake of convenience, hydrogen ion concentrations are usually converted to pH values. The term *pH* simply means that the concentration of hydrogen ions in a solution has been converted to its negative log value (i.e., pH = −log $[H^+]$. Refer to pp. 93–94 for a more detailed description of pH and the pH scale, a convenient means of expressing the acidity or alkalinity of substances.

REACTION TYPES There are several basic types of chemical reaction: synthesis reactions, decomposition reactions, displacement reactions, double displacement reactions, acid–base reactions, and redox reactions. Each is discussed briefly.

Synthetic reactions (also referred to as combination reactions) involve two or more substances that combine to form a single new substance. The reaction of sulfur trioxide (SO_3) with water, for example, yields sulfuric acid (H_2SO_4).



WORKED PROBLEM 6

The K_a for acetic acid is 1.8×10^{-5}. Determine the hydrogen ion concentration of a 0.1 M solution of acetic acid in water. What is the pH of this solution?

SOLUTION

The equation for the dissociation of acetic acid is

$$K_a = [\text{acetate}][H^+]/[\text{acetic acid}]$$

Because acetic acid is a weak acid, it is assumed that the dissociation of acetic acid has no substantive effect on acetic acid concentration. The values of the concentrations of acetate and hydrogen ions are equal to each other and are set at *x*. The equation for determining the hydrogen ion concentration in a 0.1 M solution is

$$1.8 \times 10^{-5} = x^2/0.1, \text{ which becomes}$$
$$1.8 \times 10^{-6} = x^2$$

Solving for x yields 1.35×10^{-3}, the hydrogen ion concentration in the acetic acid solution.

The pH of the 0.1 M solution of acetic acid is calculated as follows:

$$\begin{aligned} \text{pH} &= -\log[\text{H}^+] \\ &= -\log(1.35 \times 10^{-3}) \\ &= 3 - 0.13 = 2.87 \end{aligned}$$

■

In *decomposition reactions*, a compound breaks down to form simpler products when the reactant absorbs enough energy so that one or more of its bonds breaks. For example, ammonium sulfate, $(NH_4)_2SO_4$, decomposes upon heating to yield ammonia (NH_3) and H_2SO_4.

$$(NH_4)_2SO_4 \xrightarrow{\Delta} 2\ NH_3 + H_2SO_4$$

In *displacement* or substitution reactions a more reactive element replaces a less active element. For example, if an iron nail is placed in an aqueous solution of copper (II) sulfate (i.e., copper with a +2 oxidation state), the color of the solution turns from blue to green because the iron displaces the copper from copper sulfate to yield iron sulfate.

$$Fe + CuSO_4 \xrightarrow{\Delta} FeSO_4 + Cu$$

The surface of the iron nail turns to reddish brown because of the deposition of metallic copper. Predicting whether a specific metal will displace another is accomplished by referring to the *activity series of metals*, a list of metals (found in general chemistry textbooks) that is arranged in order of strength of metal reactivity from highest to lowest.

In *double displacement reactions*, two compounds exchange their ions to form two new compounds. For example, silver nitrate reacts with potassium bromide in aqueous solution to yield silver bromide and potassium nitrate.

$$AgNO_3 + KBr \rightarrow AgBr + KNO_3$$

The silver bromide product is insoluble in water and precipitates out of solution.

Acid–base reactions are a type of double displacement reaction. The *Bronsted–Lowry theory* defines acids and bases as proton donors and proton acceptors, respectively. For example, hydrogen chloride reacts with water to yield the hydronium ion H_3O^+ and chloride ion:

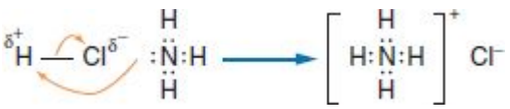
$$HCl + H_2O \rightarrow H_3O^+ + Cl^-$$

In this reaction, the hydrogen chloride donates a proton (H^+) to H_2O (acting as a base because it accepts the proton) to form the H_3O^+ and chloride ion. In this reaction, Cl^- is the *conjugate base* of the acid HCl. Together these two species constitute a *conjugate acid–base pair*. Similarly, H_3O^+ is the conjugate acid of H_2O. They also form a conjugate acid–base pair.

In another way of explaining acid–base reactions, referred to as the *Lewis acid and base theory*,

acids and bases are defined in terms of atomic structure and bonding. A *Lewis acid* is a chemical species that accepts an electron pair and has a vacant low-energy orbital. Examples of Lewis acids include cations such as Cu^{+2} and Fe^{+2} and molecules such as carbon monoxide (CO) with multiple bonds and atoms with different elctronegativities. A Lewis base is defined as a chemical species that donates an electron pair and possesses lone pair electrons. Examples include NH_3, OH^-, and cyanide ions (CN^-). The product of a Lewis acid–base reaction contains a new covalent bond.

In the reaction of HCl with ammonia, HCl is polarized, with the slightly positive hydrogen and the chloride slightly negative.



Ammonia ($:NH_3$) acting as a Lewis base is attracted to the hydrogen atom. As the lone pair on the nitrogen approaches the HCl, the latter becomes more polarized (i.e., the hydrogen becomes more positive), eventually causing the formation of a coordinate covalent bond between the nitrogen and the hydrogen as the hydrogen–chloride bond breaks.

Oxidation-reduction reactions, also referred to as *redox reactions*, involve the exchange of electrons between chemical species. Oxidation occurs when an ion, atom, or molecule loses electrons. In a reduction, there is a gain of electrons. In the reaction of metallic zinc with molecular oxygen to form zinc oxide, for example, the zinc atoms are oxidized (i.e., they lose electrons) and oxygen atoms are reduced (i.e., they gain electrons).

$$2\ Zn\ (s) + O_2\ (g) \rightarrow 2\ ZnO\ (s)$$

Although oxidation and reduction occur simultaneously, for convenience they may be considered two separate half-reactions, one involving oxidation and the other reduction. The oxidation half-reaction is

$$2\ Zn \rightarrow 2\ Zn^{+2} + 4e^-$$

where the two zinc atoms lose two electrons each.
In the reduction half-reaction

$$O_2 + 4e^- \rightarrow 2O^{2-}$$

the two atoms of oxygen gain a total of four electrons. In redox reactions, the species that gives up or "donates" electrons is referred to as the *reducing agent*. The species that accepts the electrons is referred to as the *oxidizing agent*. In the reaction of zinc with molecular oxygen, zinc is the reducing agent and molecular oxygen serves as the oxidizing agent. It should be noted that any type of reaction in which the oxidation state of the reactants changes could also be classified as a redox reaction. For example, in the Haber reaction

$$N_2\ (g) + H_2\ (g) \rightarrow 2\ NH_3\ (g)$$

in which molecular nitrogen reacts with molecular hydrogen to form ammonia, the oxidation number of nitrogen atoms changes from 0 to −3, and hydrogen atoms change from 0 to +1. The displacement reaction described on P-19, in which iron displaces the copper ion in copper (II) sulfate, is also a redox reaction because the oxidation state of iron changes from 0 to +2 and that of copper changes from +2 to 0.

Combustion reactions are a type of redox reaction in which fuel molecules react with an oxidizing agent to release large amounts of energy, usually in the form of heat and light. Reactions that release energy are described as *exothermic*. The burning of the hydrocarbon methane (natural gas) is a typical combustion reaction.

$$CH_4\,(g) + O_2\,(g) \rightarrow CO_2\,(g) + 2\,H_2O\,(g)$$

The oxidation half-reaction is

$$CH_4 + O_2 \rightarrow CO_2 + 8e^- + 4\,H^+$$

The reduction half-reaction is

$$O_2 + 4\,H^+ + 8e^- \rightarrow 2\,H_2O$$

Molecular oxygen is the oxidizing agent in the combustion of methane. The eight electrons removed from methane, the reducing agent, are used in combination with four protons to reduce the oxygen atoms to form two water molecules. It should be noted that cellular respiration, the biochemical mechanism whereby aerobic (oxygen-utilizing) living cells extract energy from fuel molecules such as the sugar glucose, is a slower, controlled combustion reaction process.

WORKED PROBLEM 7

Consider the following combustion reaction equation:

$$C_6H_{12}\,(l) + 9\,O_2\,(g) \rightarrow 6\,CO_2\,(g) + 6\,H_2O\,(g)$$

Identify the elements that are being reduced, oxidized, or unchanged in this reaction, and the reducing or oxidizing agents. The oxidation state of hydrogen is +1 throughout the reaction. The oxidation state of an atom in its elemental state is 0. The oxidation state of oxygen atoms in compounds is usually −2.

SOLUTION

The oxidation number of oxygen atoms in diatomic oxygen gas changes from 0 to −2 in both water and carbon dioxide molecules. In the following half-reaction, a decrease in the oxidation number for oxygen indicates that each oxygen atom has gained two electrons; that is, it has been reduced.

$$9\,O_2\,(g) + 36\,e^- \rightarrow 6\,CO_2\,(g) + 6\,H_2O\,(g)$$

The other half-reaction illustrates that the oxidation number of carbon in C_6H_{12} (cyclohexane) has changed from −2 to +4 in carbon dioxide.

$$C_6H_{12}\,(l) \rightarrow 6\,CO_2\,(g) + 36\,e^-$$

The increase in oxidation number indicates that each carbon atom has lost 6 electrons for a total of 36 e^-; that is, the carbon atoms have been oxidized.

In a redox reaction, the oxidizing agent is reduced and the reducing agent is oxidized. Therefore, in the cyclohexane combustion reaction, cyclohexane is the reducing agent and diatomic oxygen is the oxidizing agent. ■

MEASURING CHEMICAL REACTIONS Chemists use the mole concept as a means of determining the amounts of the reactants and products in chemical reactions. A *mole* is defined as the amount of a substance that contains as many particles (e.g., atoms, molecules, or ions) as there are atoms in 12 g of carbon-12. This number, which is 6.022×10^{23} particles, is referred to as *Avogadro's number*. So there are 6.022×10^{23} molecules in 1 mol of H_2O and 6.022×10^{23} sodium ions in 1 mol of NaCl.

The molar mass of substances (mass per mole of particles) is used to determine the amounts of reactants and products in a reaction. For example, in the reaction of methane (CH_4) with O_2 to yield carbon dioxide and water, how much water is produced from the combustion of 8 g of methane? Solving this problem begins with a balanced equation:

$$CH_4 + 2\,O_2 \rightarrow CO_2 + 2\,H_2O$$

According to this reaction equation, the combustion of every mole of methane yields 2 mol of water. The number of moles of methane is calculated by dividing the mass of methane (8 g) by the molecular mass of methane, which is 16 g (the carbon atom has a mass of 12 g, and each of the four hydrogens is 1 g). (Refer to the periodic table for atomic mass numbers.) By this calculation, there are 0.5 mol of methane in the reaction. Because the ratio of methane to water is 1 to 2, the 0.5 mol of methane are multiplied by 2 to yield 1 mol of water. Because the molecular mass of water is 18 g, the combustion of 8 g of CH_4 produces 18 g of H_2O.

Moles are also used to express concentrations of substances in solution. *Molarity* is defined as the number of moles in 1 liter (l) of solution. For example, what is the molarity of a solution of 5 g of NaCl in 2 l of water? First, the number of moles of NaCl must be determined by dividing the mass of NaCl (5 g) by the formula weight of NaCl (58.5 g; i.e., 23 g for sodium and 35.5 g for chlorine). By this calculation (i.e., 5/58.5), there are 0.085 mol of NaCl in the 2 l of solution. Molarity of the solution is determined by dividing the number of moles by the number of liters. The molarity of the solution in this problem is 0.085 mol/2 l, which is equal to 0.0425 M (moles per liter, or molar). This number is rounded off to 0.043 M because of the rule of significant figures. Refer to a general chemistry textbook for a discussion of significant figures.

WORKED PROBLEM 8

The empirical formula of the sugar glucose is $C_6H_{12}O_6$. (a) How many moles are there in 270 g of glucose? (b) Calculate the molarity of a solution of 324 g of glucose dissolved in 2.0 l of water.

SOLUTION

(a) The number of moles of glucose is calculated by dividing the molecular mass of glucose by its mass. First, the molecular mass of glucose must be determined by adding the sums of the masses of each atom in glucose.

$$\text{Carbon: } 12\text{ g} \times 6\text{ atoms} = 72\text{ g}$$
$$\text{Hydrogen: } 1\text{ g} \times 12\text{ atoms} = 12\text{ g}$$
$$\text{Oxygen: } 16\text{ g} \times 6\text{ atoms} = 96\text{ g}$$

Adding these numbers, the molecular mass (m) of glucose is determined to be 180 g. The number of moles of glucose in 270 g is calculated by dividing the mass of glucose (270 g) by the molecular mass (180 g).

$$\text{Moles} = 270\text{ g} / 180\text{ g} = 1.5\text{ mol}$$

There are 1.5 mol of glucose in 270 g of the substance.

(b) The molarity of the glucose solution is calculated by first determining the number of moles of glucose in 324 g.

$$\text{Moles} = 324\text{ g} / 180\text{ g} = 1.8\text{ mol}$$

The molarity of the glucose solution is then calculated by dividing the number of moles by the number of liters.

$$\text{Mol/l} = 1.8\text{ mol} / 2.0\text{ l} = 0.9\text{ M}.$$

The molarity of the glucose solution is 0.9 M. ■

ORGANIC CHEMISTRY

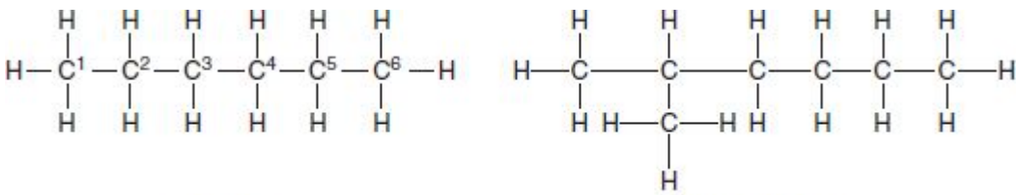
Organic chemistry is the investigation of carbon-containing compounds. An entire field is devoted to molecules composed of carbon because of its astonishing versatility. In addition to its capacity to form stable covalent bonds with other carbon atoms to form long chains, branch chains, and rings, carbon also forms stable covalent bonds with a variety of other elements (e.g., hydrogen, oxygen, nitrogen, and sulfur). Carbon can also form carbon–carbon double and triple bonds. As a result of these properties, the possibilities for molecules with different arrangements of carbon and the other elements are virtually limitless. For students embarking on the study of biochemistry, a thorough understanding of the principles of organic chemistry is essential because, as stated previously, with the exceptions of inorganic molecules such as H_2O, O_2, NH_3, and CO_2 and several minerals (e.g., Na^+, Ca^{2+}, and Fe^{2+}), biomolecules are organic molecules. The structural and functional properties of proteins, nucleic acids (DNA and RNA), fats, and sugars can only be appreciated when students know how carbon-based molecules behave. This review will focus on the structures and the chemical properties of the major classes of organic molecules: the *hydrocarbons* (molecules containing only carbon and hydrogen) and *substituted hydrocarbons* (hydrocarbon molecules in which one or more hydrogens has been replaced with another atom or group of atoms).

Hydrocarbons

Because hydrocarbon molecules contain only carbon and hydrogen, they are nonpolar. They dissolve in nonpolar solvents such as hexane and chloroform but not in water. Such molecules are described as *hydrophobic* ("water-hating"). The hydrocarbons are classified into four groups: (1) saturated hydrocarbons (molecules containing only single bonds), (2) unsaturated hydrocarbons (molecules with one or more carbon–carbon double or triple bonds), (3) cyclic hydrocarbons (molecules containing one or more carbon rings), and (4) aromatic hydrocarbons (molecules that contain one or more aromatic rings, which can be described as cyclic molecules with alternating double and single bonds).

The *saturated hydrocarbons*, referred to as the *alkanes*, are either normal (straight chains) or branched chains. These molecules are "saturated" because they will not react with hydrogen. The straight-chain alkanes belong to a homologous series of compounds that differ in the number of carbon atoms they contain. Their formula is C_nH_{2n+2}. The first six members of this series are methane (CH_4), ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}), pentane (C_5H_{12}), and hexane (C_6H_{14}). Note that the prefix in each of these names indicates the number of carbon atoms (e.g., meth- = 1 carbon atom) and the suffix -ane indicates a saturated molecule. Hydrocarbon groups that are derived from alkanes are called *alkyl groups*. For example, a methyl group is a methane molecule with one hydrogen atom removed.

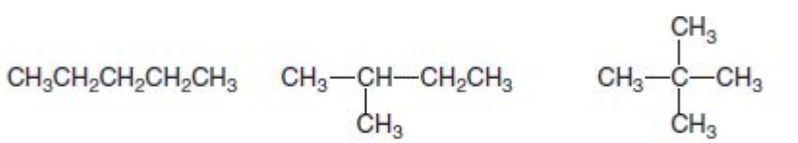
As their name suggests, the branched-chain hydrocarbons are carbon chains with branched structures. If one hydrogen atom is removed from carbon-2 of hexane, for example, and a methyl group is attached, the branched product is 2-methylhexane.



Hexane 2-Methylhexane

Note that branched-chain molecules are named by first identifying the longest chain and that the number of the carbon that is bonded to the sidechain group is the lowest one possible.

One of the most remarkable features of the hydrocarbons is the capacity to form *isomers*, molecules with the same type and number of atoms that are arranged differently. For example, there are three molecules, each with its own set of properties, with the molecular formula (C_5H_{12}): pentane, 2-methylbutane, and 2,2-dimethylpropane.

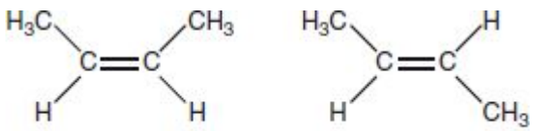


Pentane 2-Methylbutane 2,2-Dimethylpropane

The alkanes are unreactive except for combustion (p. P-20) and halogenation reactions. In halogenation reactions, alkane molecules react at elevated temperatures or in the presence of light, forming *free radicals* (atoms or molecules with an unpaired electron). For example, when methane reacts with chlorine gas (Cl_2), the molecule breaks down to form two chlorine radicals, which then initiate a chain reaction with methane molecules that yields several chlorinated products: CH_3Cl (methyl chloride), CH_2Cl_2 (methylene chloride), $CHCl_3$ (chloroform), and CCl_4 (carbon

tetrachloride).

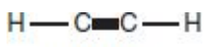
There are two types of unsaturated hydrocarbons: the *alkenes*, which contain one or more double bonds, and the *alkynes*, which contain one or more triple bonds. The double bond in alkenes is formed from the overlap of two carbon sp^2 orbitals (a σ bond) and the overlap of two unhybridized p orbitals (one from each carbon) to form a π bond. The homologous family of alkenes (formula = C_nH_{2n}) are named by taking the names of the alkane with the same number of carbons and substituting the suffix -ene for -ane. Ethene ($H_2C{=}CH_2$), also known by the older name ethylene, is the first member of the series. For alkenes with more than three carbons, the carbons are numbered in reference to the double bond so that the numbers are the lowest possible. For example, $CH_2{=}CH{-}CH_2{-}CH_2{-}CH_2{-}CH_3$ is named 1-hexene, not 5-hexene. Alkenes with four or more carbons have structural isomers in which the position of the carbon–carbon double bond is different. For example, 1-butene and 2-butene are referred to as positional isomers. The rigidity of the carbon–carbon double bond prevents rotation, thereby producing another class of isomers: geometric isomers. *Geometric isomers* occur when each of the carbons in the double bond has two different groups on it. For example, there are two geometric isomers of 2-butene: *cis*- 2-butene, in which the methyl groups are on the same side of the double bond, and *trans*-2-butene, in which the methyl groups are on opposite sides of the double bond.



Cis-2-Butene *Trans*-2-Butene

Note that 1-butene does not form geometric isomers because one of the double-bonded carbons does not have two different groups.

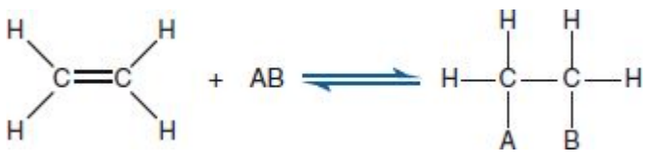
The alkynes such as ethyne (or acetylene)



Acetylene

contain triple bonds composed of 1 σ bond and 2 π bonds. The carbon–carbon triple bond is rare in biomolecules and is not discussed further.

The principal reaction of alkenes is the *electrophilic addition reaction* in which an *electrophile* (an electron-deficient species) forms a bond by accepting an electron pair from a *nucleophile* (an electron-rich species).



Electrophiles have positive charges or they may have an incomplete octet. Examples include H^+, CH_3^+, and polarized neutral molecules such as HCl. Nucleophiles have negative charges (e.g., OH^-), contain atoms with lone pairs (e.g., H_2O and NH_3), or have π bonds.

Hydrogenation and hydration are two addition reactions that occur frequently in living organisms. In a laboratory or industrial hydrogenation reaction, a metal catalyst (e.g., nickel or platinum) is required to promote the addition of H_2 to an alkene to yield an alkane.

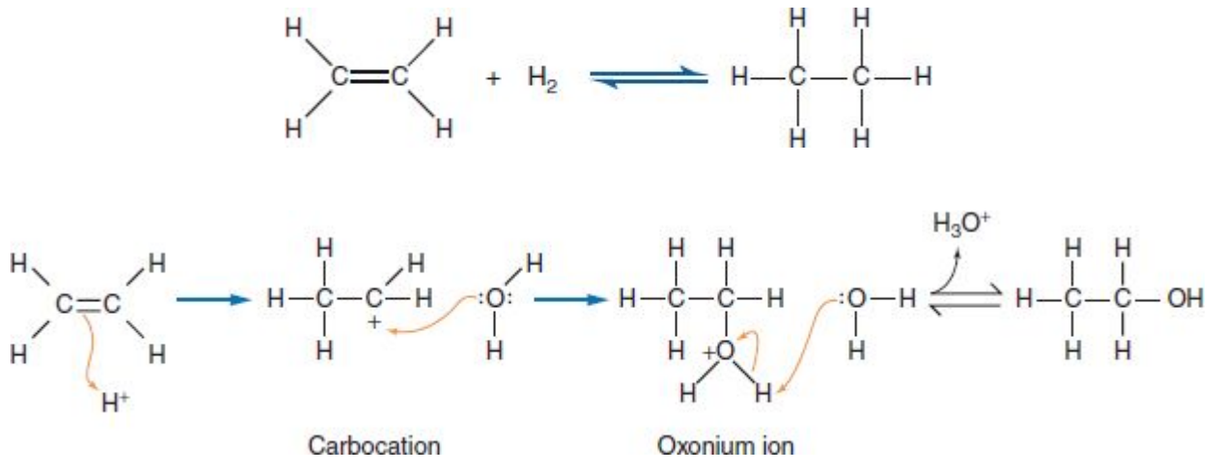


FIGURE 9

Acid-Catalyzed Hydration of an Alkene

In the first step, protonation of the double bond forms a carbocation. Nucleophilic attack by water gives a protonated alcohol. Deprotonation by a water molecule yields the alcohol. Note that the narrow, colored arrows in this and other reaction mechanisms indicate the movement of electrons.

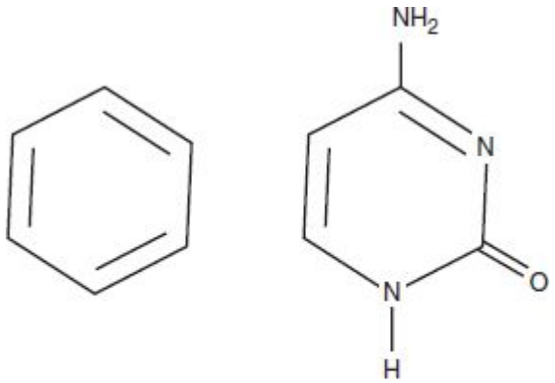
Hydration reactions of alkenes are electrophilic addition reactions that yield alcohols. The reaction (**Figure 9**) requires a small amount of a strong acid catalyst such as sulfuric acid (H_2SO_4) because water is too weak an acid to initiate protonation of the alkene. In the presence of the sulfuric acid, the electron pair in the π cloud polarizes toward the hydronium ion (H_3O^+) and forms a new carbon–hydrogen ion. The newly formed *carbocation* (a molecule containing a positively charged carbon atom) is then attacked by the nucleophilic water molecule to yield an *oxonium ion* (a molecule containing an oxygen cation with three bonds). The alcohol product is formed as the oxonium ion transfers a proton to a water molecule. It is important to note that in hydration reactions of propene and larger alkenes, the most highly substituted carbocation will form (Markovnikov's rule). For example, the principal product of propene hydration is 2-propanol and not 1-propanol.

CYCLIC HYDROCARBONS As their name suggests, the cyclic hydrocarbons are the cyclic counterparts of the alkanes and alkenes. The general formula of the cyclic alkanes, C_nH_{2n}, has two fewer hydrogens than the alkane formula C_nH_{2n+2}. As with the alkanes, the cycloalkanes undergo combustion and halogenation reactions.

Ring strain, observed in cycloalkane rings with three or four carbons, is caused by unfavorable bond angles that are the result of distortion of tetrahedral carbons. As a result, the carbon–carbon bonds in these molecules are weak and reactive. There is minimal or no ring strain in cycloalkane rings with five to seven carbons. Cyclohexane has no ring strain because it is puckered so that its bond angles are near the tetrahedral angles. The most stable puckered conformation is the chair form.

Because of greater ring strain, cyclopropene and cyclobutene are even less stable than cyclopropane and cyclobutane.

AROMATIC HYDROCARBONS Aromatic hydrocarbons are planar (flat) hydrocarbon rings with alternating single and double bonds. Heterocyclic aromatic compounds have two or more different elements in their rings. The simplest aromatic hydrocarbon is benzene. Cytosine, a pyrimidine base found in DNA and RNA, is an example of a heterocyclic aromatic molecule.



Benzene Cytosine

Despite the presence of double bonds, benzene and the other aromatic molecules do not undergo reactions typical of the alkenes. In fact, aromatic compounds are remarkably stable. This stability is the result of the unique bonding arrangement of aromatic rings. Each carbon has three sp^2 orbitals that form three σ bonds with two other carbon atoms and with one hydrogen atom. The 2p orbital of each of the six carbon atoms overlaps side to side above and below the plane of the ring to form a continuous circular π bonding system. Instead of two alternate structures of benzene, benzene is a resonance hybrid, which is indicated by the fact that all of the carbon–carbon bonds in benzene are the same length. (In alkenes, carbon–carbon double bonds are shorter than carbon–carbon single bonds.) Each carbon atom is joined to its neighbors by the equivalent of one and a half bonds. Resonance explains why aromatics do not undergo the addition reactions observed with alkenes: the delocalizing of the π electrons around an aromatic ring confers considerable stability to the molecule.

Alternate Structures of Benzene

Not all cyclic compounds that contain double bonds are aromatic molecules. According to *Huckel's rule*, to be aromatic a ring molecule must be planar and have $4n + 2$ electrons in the π cloud, where n is a positive integer. For benzene, which has six π electrons, n is equal to 1 [4(1) + 2 = 6]. In addition, every atom in an aromatic ring has either a p orbital or an unshared pair of electrons. For example, cytosine is aromatic because the NH adjacent to the carbonyl group donates its lone pair of electrons to the ring's π electron cloud.

Despite their resistance to addition reactions, aromatic compounds are not inert. They can undergo substitution reactions. In *electrophilic aromatic substitution reactions*, an electrophile reacts with an aromatic ring and substitutes for one of the hydrogens. For example, benzene reacts with HNO_3 in the presence of H_2SO_4 to yield nitrobenzene and water (**Figure 10**). In the first step in the reaction, a strong electrophile is generated. In this case, the nitronium ion ($^+NO_2$) is created when the sulfuric acid protonates the nitric acid on the OH group and the resulting water molecule leaves. In the second step, a pair of π electrons in the benzene attack the electrophile, resulting in

the formation of a resonance-stabilized carbocation intermediate. In the final step, the aromatic ring is regenerated when a water molecule abstracts a proton from the carbon atom bonded to the electrophile.

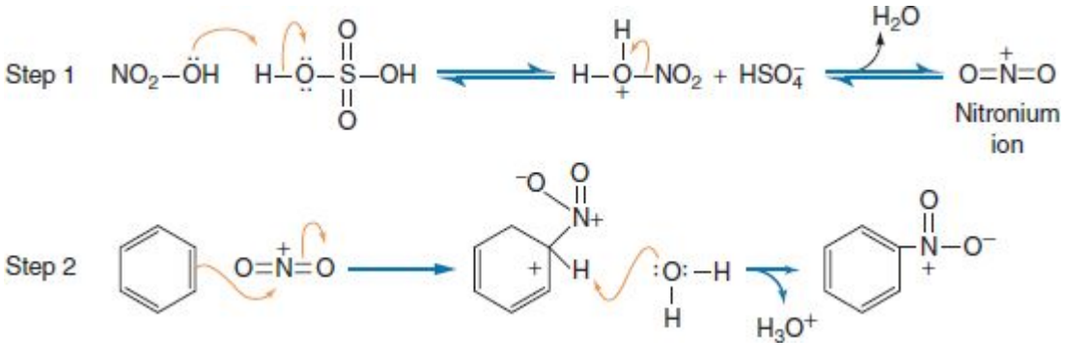


FIGURE 10

Nitration of Benzene

The nitronium ion, a powerful electrophile, is created by the protonation of HNO_3 by H_2SO_4. The product loses a water molecule to leave O=N+=O In the second step, the nitronium ion reacts with the nucleophilic benzene. The aromaticity of nitrobenzene is restored with the loss of a proton to a water molecule.

Substituted Hydrocarbons

Substituted hydrocarbons are produced by replacing one or more hydrogens on hydrocarbon molecules with functional groups. A *functional group* is a specific group of atoms within a molecule that is responsible for the molecule's chemical reactivity. Functional groups also separate the substituted hydrocarbons into families. For example, methanol (CH_3—OH), a member of the alcohol family of organic molecules, is the product when the functional group –OH is substituted for a hydrogen atom on methane (CH_4). Three general classes of functional groups are important in biomolecules: oxygen-containing, nitrogen- containing, and sulfur-containing molecules. The structural and chemical properties of each class are briefly discussed. Also refer to **Table 1.1** on p. 6 of the textbook for a brief overview of the functional groups. There are six major families of organic molecules that contain oxygen: alcohols, aldehydes, ketones, carboxylic acids, esters, and ethers. Amines and amides possess nitrogen-containing functional groups. The major type of sulfur- containing functional group is the sulfhydryl group, which occurs in thiols.

ALCOHOLS In alcohols, the hydroxyl group (–OH) is bonded to an sp^3 hybridized carbon. The presence of the polar –OH group makes alcohol molecules polar, allowing them to form hydrogen bonds with each other and with other polar molecules. A *hydrogen bond* is an attractive force between a hydrogen atom attached to an electronegative atom (e.g., oxygen or nitrogen) of one molecule and an electronegative atom of a different molecule. For alcohols with up to four carbons (methanol, ethanol, propanol, and butanol), the polar OH group allows them to dissolve in water because hydrogen bonds form between the hydrogen of the OH group of the alcohol and the oxygen of a water molecule. Such molecules are described as *hydrophilic* ("water-loving"). Alcohols with five or more carbons are not water-soluble because the hydrophobic properties of hydrocarbon components of these molecules are dominant. Alcohols can be classified by the number of alkyl groups (designated as R groups) attached to the carbon adjacent to the –OH group. Ethanol (CH_3CH_2—OH) is a primary alcohol. In a secondary alcohol (RR′CH—OH) such as 2-propanol, the carbon atom bonded to the OH group is also attached to two alkyl groups.

Tertiary alcohols (RR′R″C—OH) such as 2-methyl-2-propanol have three alkyl groups bonded to the carbon bearing the OH group. Alcohols are weak acids (i.e., a strong base can remove the proton from an alcohol's hydroxyl group to form the alkoxide ion R—O^-). Tertiary alcohols are less acidic than primary alcohols because the alkyl groups inhibit the solvation of the alkoxide ion. The increased electron density on the oxygen atom in these molecules also decreases proton removal.

Alcohols react with carboxylic acids to form esters. They can also be oxidized to give the carbonyl group-containing aldehydes, ketones, or carboxylic acids. The *carbonyl group* (C═O) in which a carbon is double-bonded to an oxygen atom is a structural feature of the aldehydes, ketones, carboxylic acids, and esters. (Amides, which contain both nitrogen atoms and carbonyl groups, are described on p. P-30.) The carbonyl group is polar because of the difference in electronegativity between oxygen and carbon. The slightly positive carbon is electrophilic and therefore able to react with nucleophiles.

ALDEHYDES The functional group of the aldehydes is a carbonyl group bonded to a hydrogen atom [–(C═O)—H]. The simplest aldehyde is formaldehyde (also referred to as methanal) in which the aldehyde group is bonded to a hydrogen atom. In all other aldehydes the aldehyde group is bonded to an alkyl group. The general formula for aldehydes is abbreviated as R-CHO. Acetaldehyde (CH_3CHO) is the oxidation product of ethanol. The reaction of an aldehyde with an alcohol yields a hemiacetal (**Figure 11**). Note that hemiacetals are unstable and their formation is readily reversible. The reaction of the aldehyde group of aldose sugars with an intramolecular OH group to form the more stable cyclic hemiacetals (see p. 250 in the textbook) is an important feature of the chemistry of carbohydrates (Chapter 7).

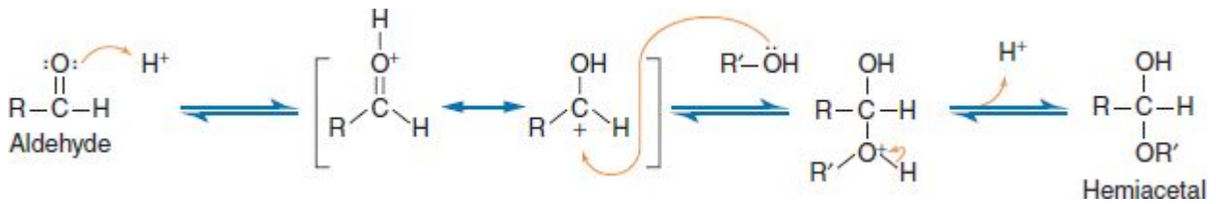


FIGURE 11

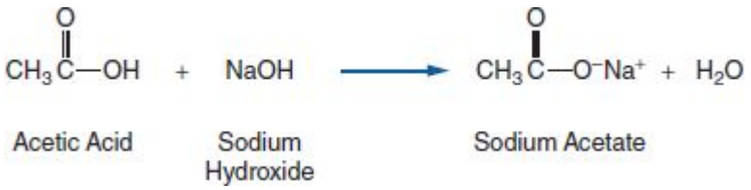
Hemiacetal Formation

The reaction begins with the acid catalyst protonating the carbonyl group. The alcohol, acting as a nucleophile, attacks the resonance stabilized carbocation. The hemiacetal product forms with the release of a proton from the positively charged intermediate.

KETONES Ketones are molecules in which a carbonyl group is flanked by two R groups [R—(C═O)—R′]. The names of members of the ketone family end in -one. For example, dimethyl ketone is usually referred to by its original name, acetone. Ethylmethyl ketone is also referred to as 2-butanone.

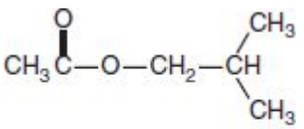
In ketoses, sugars with a ketone group (most notably fructose), the carbonyl group reacts with an OH group on the sugar molecule to form a cyclic hemiketal.

CARBOXYLIC ACIDS Carboxylic acids (RCOOH) contain the carboxyl group, which is a carbonyl linked to an OH group. These molecules function as weak acids (i.e., they are proton donors) because the carboxylate group (COO^-), the conjugate base of a carboxylic acid, is resonance stabilized. Carboxylic acids react with bases to form carboxylate salts. For example, acetic acid reacts with sodium hydroxide to yield sodium acetate and water.



The simplest carboxylic acid is formic acid (HCOOH), which is found in ant and bee stings. Carboxylic acids with more than two carbon atoms are often named using the hydrocarbon precursor name followed by the ending -oic acid. For example, the carboxylic acid derived from the four-carbon molecule butane is butanoic acid. In living organisms, the longer-chained carboxylic acids, called *fatty acids*, are important components of biological membranes and the triacylglycerols, a major energy-storage molecule.

ESTERS Found widely in nature, esters [R(C═O)—OR′] are responsible for the aromas of numerous fruits. An ester is the product of a *nucleophilic acyl substitution reaction* in which a carboxylic acid reacts with an alcohol. For example, isobutanol reacts with acetic acid to form isobutyl acetate, an ester found in cherries, raspberries, and strawberries.



Isobutyl acetate

The formation of the ester methyl acetate from acetic acid and methanol is illustrated in **Figure 12**.

Fats and vegetable oils, also called triacylglycerols (see p. 248 in the textbook), are triesters of the trialcohol molecule glycerol and three fatty acids.

ETHERS Ethers have the general formula R—O—R′. Diethyl ether (CH_3CH_2—O— CH_2CH_3), the best-known ether, was the first anesthetic used in surgery (late nineteenth century) and is still used as a solvent. Ethers are relatively inert chemically, but they do convert over time into explosive peroxides (e.g., diethyl ether hydroperoxide) when exposed to air.

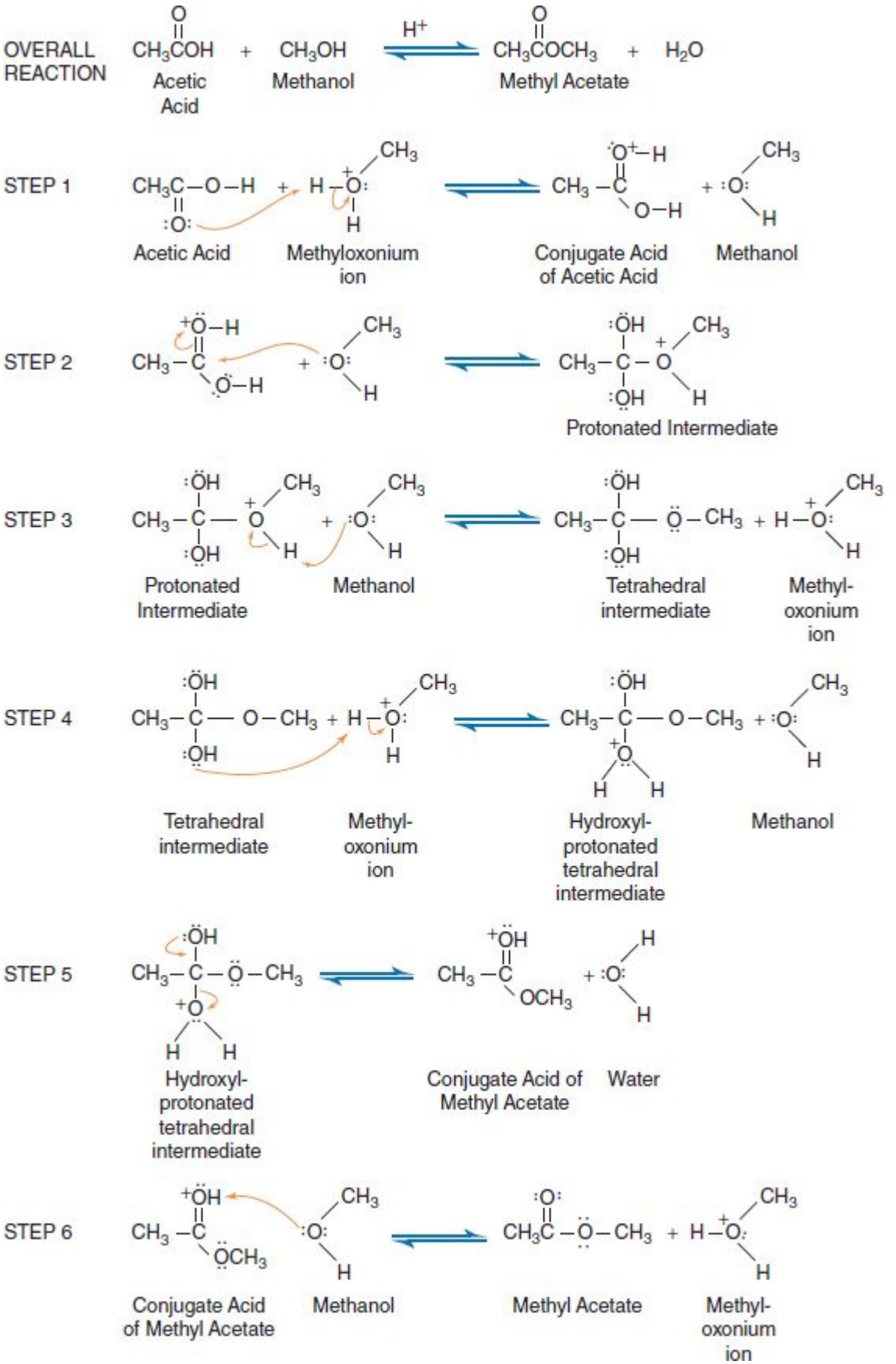
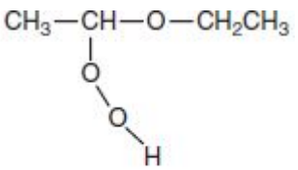


FIGURE 12

Formation of Methyl Acetate

Step 1 Acetic acid is protonated on its carbonyl oxygen to form the conjugate acid of acetic acid. **Step 2** A molecule of methanol, acting as a nucleophile, attacks the electrophilic carbon of the protonated acetic acid. **Step 3** The oxonium ion (the protonated intermediate formed in step 2) loses a proton to form a neutral tetrahedral intermediate. **Step 4** The tetrahedral intermediate is protonated on one of its hydroxyl oxygens.

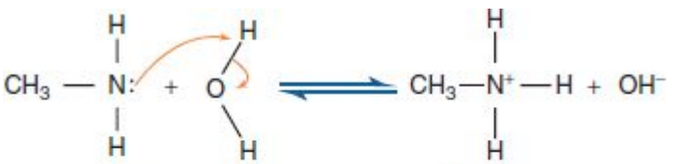
Step 5 The hydroxyl protonated intermediate formed in step 4 loses a molecule of water to yield the protonated form of the ester. **Step 6** The loss of a proton from the protonated product of step 5 (the conjugate acid of methyl acetate) yields methyl acetate.



Diethylether hydroperoxide

In living organisms, the ether linkage occurs in biomolecules such as the carbohydrates.

AMINES Amines are organic molecules that can be considered derivatives of ammonia (NH_3). Primary amines ($R—NH_2$) are molecules in which only one of the hydrogen atoms of ammonia has been replaced by an organic group (e.g., alkyl or aromatic groups). Methylamine (CH_3NH_2) is an example of a primary amine. In secondary amines, such as dimethylamine ($CH_3—NH—CH_3$), two hydrogens have been replaced by organic groups. Tertiary amines such as triethylamine [$(CH_3CH_2)_3N$] are molecules in which all three hydrogens have been replaced with organic groups. Amines with small organic groups are water-soluble, although the solubility of tertiary amines is limited because they do not have any hydrogen atoms bonded to the electronegative nitrogen atom. Like ammonia, amines are weak bases because of the lone pair of electrons on the nitrogen atom, which can accept a proton. Protonation of the nitrogen converts the amine into a cation.



Methyl amine

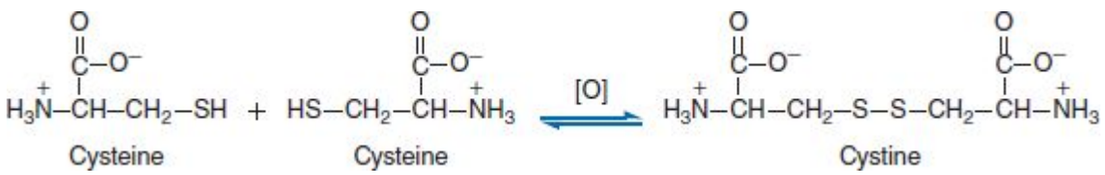
Methyl-ammonium ion

There are an enormous number of biomolecules that contain amine nitrogens. Examples include the amino acids (components of proteins), the nitrogenous bases of the nucleic acids, and the alkaloids (complex molecules produced by plants such as caffeine, morphine, and nicotine that have significant physiological effects on humans).

AMIDES Amides are amine derivatives of carboxylic acids with the general formula [$R(C{=}O)—NR_2$] where the R groups bonded to the nitrogen can be hydrogens or hydrocarbon groups. In contrast to the amines, amides are neutral molecules. The C—N bond is a resonance hybrid because of the attraction of the carbonyl group for the nitrogen's lone pair. As a result, amides are not weak bases (i.e., they have little capacity to accept protons).

Resonance hybridization explains why the amide functional group is planar, with the nitrogen's sp^2 orbital forming a π bond with the carbonyl carbon atom. In living organisms, the amide functional group is the linkage, referred to as the peptide bond, that connects amino acids in polypeptides. Amides are classified according to how many carbon atoms are bonded to the nitrogen atom. Molecules with the molecular formula R(C═O)—NH_2 are primary amides. The substitution of one of the nitrogen's hydrogens with an alkyl group yields a secondary amide [R(C═O)—NHR′]. Amides with two alkyl groups attached to the nitrogen are tertiary amides [R(C═O)—NR_2].

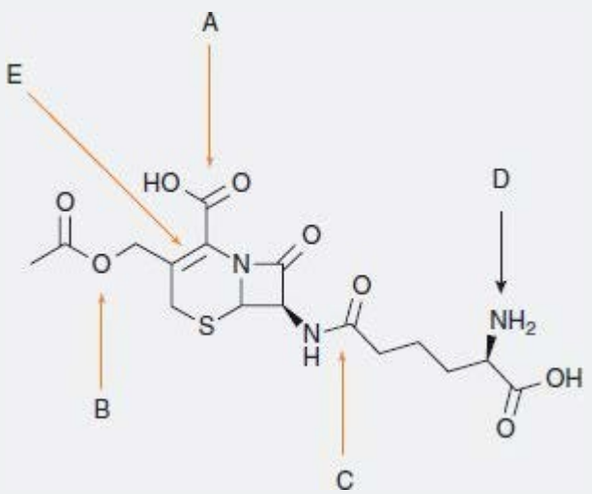
THIOLS A thiol is a molecule in which an sp^3 carbon is bonded to a sulfhydryl group (–SH). Although thiols are considered the sulfur analogues of alcohols, the low polarity of the SH bonds limits their capacity to form hydrogen bonds. As a result, thiols are not as soluble in water as their alcohol counterparts. Thiols are stronger acids than their alcohol equivalents, however, in part because of the weakness of the S—H bond. For the same reason, thiolates ($R—S^-$), the conjugate bases of the thiols, are weaker bases than the alkoxides ($R—O^-$). Thiolates are excellent nucleophiles because sulfur's 3p electrons are easily polarized. The sulfhydryl group of thiols is easily oxidized to form disulfides (RS—SR). For example, two molecules of the amino acid cysteine react to form cystine, which contains a disulfide bond.



This reaction is especially important in numerous proteins that contain cysteine. The disulfide bond that forms when cysteine residues are linked is an important stabilizing feature of protein structure.

WORKED PROBLEM 9

Cephalosporin c is one of a class of antibiotics, known as the cephalosporins, that kill bacteria by preventing the cross-linking of peptidoglycan, a key structural element in the cell wall. Its structure is shown as follows. Identify the indicated functional groups.



SOLUTION

A = carboxylic acid B = ester C = amide D = amine E = alkene ■

Organic Reactions: Substitutions and Eliminations

There are a very large number of organic reaction types. Electrophilic addition reactions (p. P-24), electrophilic aromatic substitution reactions (p. P-26), and nucleophilic acyl substitution (p. P-28) have already been described. There are two additional reaction classes, aliphatic substitution and elimination reactions, which students should be familiar with. (The term *aliphatic* refers to nonaromatic hydrocarbon compounds.)

SUBSTITUTION REACTIO NS Aliphatic substitution reactions, which involve tetrahedral carbons, are designated as either S_N1 or S_N2. *S_N1* (substitution nucleophilic unimolecular) reactions, which often involve secondary or tertiary alkyl halides (molecules in which a halogen atom such as chloride has been substituted for a hydrogen in an alkane) or alcohols, proceed in two steps. In the first step of an S_N1 reaction, a planar carbocation forms as the leaving group (a stable ion or a neutral molecule) is displaced. S_N1 reactions are considered unimolecular because the rate of the reaction depends only on the rate of carbocation formation. In the second step, the nucleophile attacks the electrophilic carbocation to form the product. Good reactants for S_N1 reactions are molecules with tertiary carbons that can form a stable carbocation when the leaving group is released. Examples of nucleophiles in S_N1 reactions include alcohols and water. Because carbocations are sp^2 hybridized and have an empty p orbital, the nucleophile can attack on either side of the ion. As a result, two isomeric products may be produced. A typical example of an S_N1 reaction is the reaction of *t*-butyl bromide with methanol (**Figure 13**).

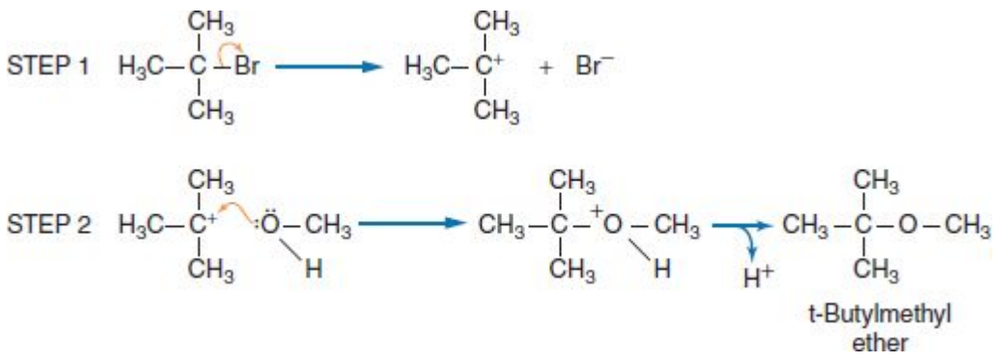


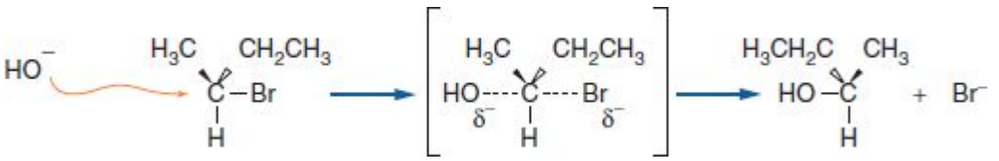
FIGURE 13

Example of S_N1 Reaction

In the first and slowest step, the alkyl halide *t*-butyl bromide forms the *t*-butyl carbocation as the leaving group bromide is released. In step 2, the nucleophilic oxygen of methanol attacks the carbocation. The product methyl *t*-butyl ether is formed as a proton is released from the oxonium ion into the solvent.

S_N2 (substitution nucleophilic bimolecular) reactions differ from S_N1 reactions in that there are no carbocation-like intermediates and reaction rates are determined by the concentrations of both the nucleophilic and the electrophilic reactants. S_N2 reactions proceed in one step: as the

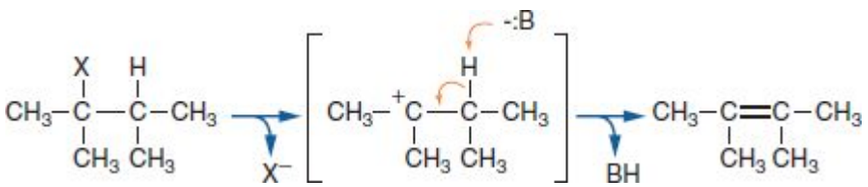
nucleophile, functioning as a Lewis base (an electron pair donor), donates its electron pair to an electrophilic carbon that has been polarized by an electronegative atom. As a result, the leaving group leaves. If the electrophilic carbon is asymmetric (i.e., four different groups are attached), an inverted configuration around this carbon will be observed in the product. Because the nucleophile attacks the back of the electrophilic reactant, S_N2 reactions occur most rapidly with primary carbons, followed by secondary carbons. Molecules with tertiary carbons do not undergo S_N2 reactions because of *steric hindrance* (the reactive site on a molecule is blocked by adjacent groups).



A class of enzymes referred to as the *S*-adenosylmethionine-dependent methyltransferases catalyzes some of the best-known examples of S_N2 reactions in biochemistry. *S*-Adenosylmethionine (SAM) is widely used in methylation reactions as a methyl group donor. SAM's methyl group is readily donated because it is attached to an electron-withdrawing sulfur atom. The inactivation of the neurotransmitter epinephrine (adrenalin) catalyzed by catechol-*O*-methyltransferase (COMT) is illustrated in **Figure 14**.

ELIMINATION REACTIONS As their name suggests, elimination reactions involve the loss of two atoms or groups from a molecule. This loss is usually accompanied by the formation of a π bond. There are three types of elimination reactions: E1, E2, and E1cb.

In E1 (elimination unimolecular) reactions, the first and slowest step is carbocation formation. For this reason E1 reactions are similar to S_N1 reactions because the reaction rates of both depend on one molecule, the precursor of the carbocation. In the second and faster step, a weak base (often a solvent molecule) abstracts a proton from a carbon atom next to the carbocation to yield a carbon–carbon double bond.



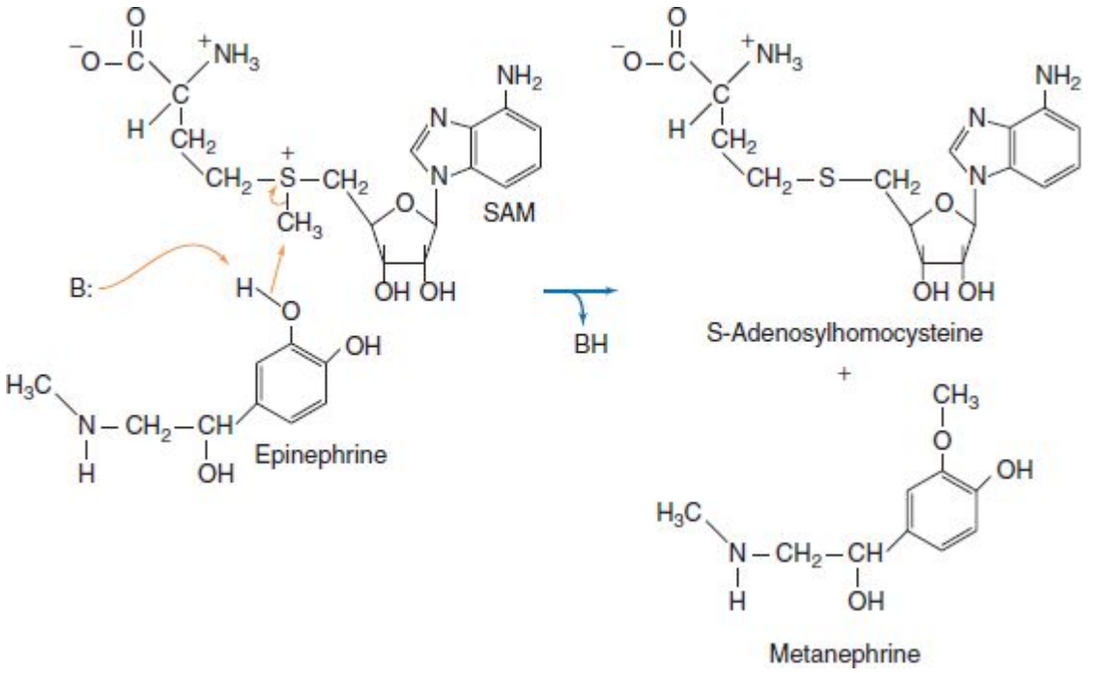


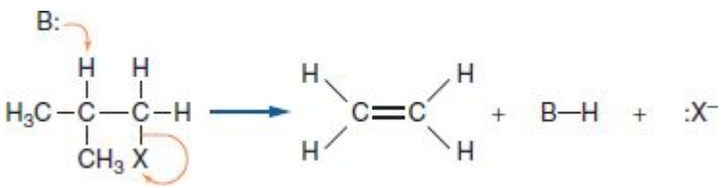
FIGURE 14

COMT-Catalyzed Methylation of Epinephrine: An S_N1 Mechanism

When epinephrine enters the active site of the enzyme, a basic amino acid R group deprotonates one of the hydroxyl groups of epinephrine, forming a nucleophilic alkoxide group. The alkoxide then attacks the methyl carbon atom of SAM. This carbon is electrophilic because it is bonded to an electron-withdrawing positively charged sulfur atom. As the alkoxide attacks, the sulfur–carbon bond begins to break to form the more stable sulfide leaving group. The products of the reactions are *S*- adenosylhomocysteine and epinephrine's inactive methylated derivative metanephrine.

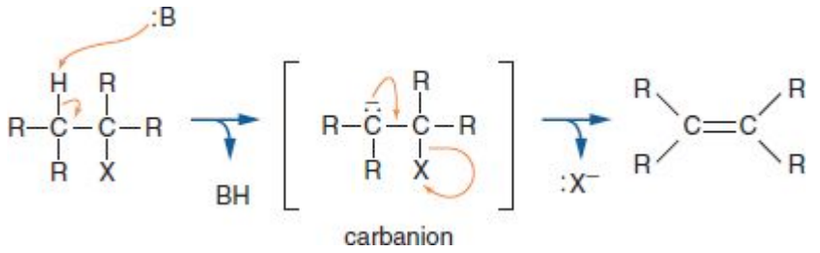
Note that the carbon bonded to "X" in the reactant molecule is sp^3 hybridized, whereas the same carbon in the product is sp^2 hybridized. Molecules with tertiary carbons groups are good substrates for E1 reactions because reaction rates depend on the stability of the carbocation.

E2 (elimination bimolecular) reactions involve the simultaneous removal of a β- proton by a strong base and release of the leaving group. (A β-proton is bonded to the carbon adjacent to the carbon bearing the leaving group.) Molecules with tertiary carbons typically undergo E2 reactions.



For an E2 reaction mechanism to be possible, the β-hydrogen and the leaving group must be anticoplanar (i.e., in a geometric arrangement in which they are 180° from each other on neighboring carbons).

Although all three forms of elimination reactions occur in living organisms, the E1cB is more commonly observed. E1cb (elimination unimolecular conjugate base) reactions involve the formation of a *carbanion* (an organic ion bearing a negatively charged carbon). Carbanions are nucleophiles that are stabilized by adjacent electronegative atoms and resonance effects. The E1cb mechanism involves the removal of a proton from the reactant to form the carbanion, followed by the slower loss of the leaving group.



The second phase, the loss of the leaving group, is the rate-limiting step in E1cb reactions. The presence of an electron-withdrawing atom or group (X in this illustration) makes the C— H acidic. Note that the carbanion form of the reactant is the conjugate base referred to in the term *E1cb*. The conversion of glycerate-2-phosphate, an intermediate in glucose degradation, to phosphoenolpyruvate provides an example of an E1cb mechanism (**Figure 15**). This reaction is catalyzed by the enzyme enolase.

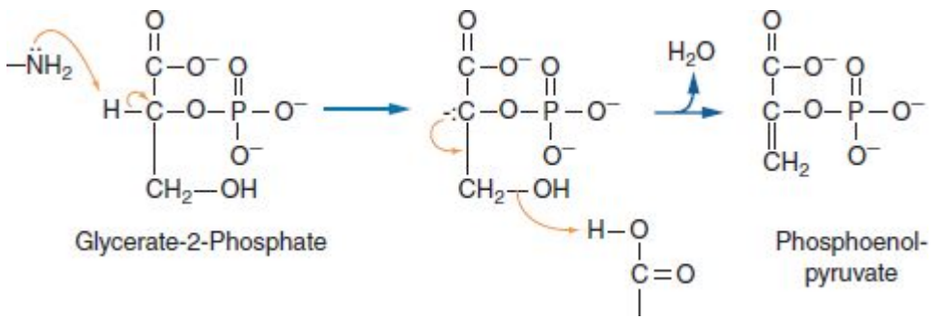


FIGURE 15

Enolase-Catalyzed Dehydration of Glycerate-2-Phosphate: An E1cb Mechanism

Within the active site of the enzyme enolase, the amino nitrogen of a precisely oriented side chain of an amino acid, acting as a base, abstracts the acidic hydrogen at carbon-2 to produce a carbanion. The carbanion electrons displace the hydroxyl (–OH) group from carbon-3, forming a carbon–carbon π bond. The actual leaving group is H_2O, the hydroxyl group having been protonated by a nearby carboxyl (–COOH) group.

CHAPTER 1

Biochemistry: An Introduction

Biochemistry and the Life Sciences Throughout most of its history, biochemical research has provided the life sciences with biochemical knowledge and ever more sophisticated technologies that have revealed living processes at the molecular level. Current examples of the medical relevance of biochemistry include immunotherapy (harnessing a patient's immune system to treat diseases such as cancer) and CRISPR (clustered regularly interspaced short palindromic repeats—a gene-editing technique used to alter DNA in living cells).

OUTLINE

WHY STUDY BIOCHEMISTRY?

1.1 WHAT IS LIFE?

1.2 BIOMOLECULES

Functional Groups of Organic Biomolecules

Major Classes of Small Biomolecules

1.3 IS THE LIVING CELL A CHEMICAL FACTORY?

Biochemical Reactions
Energy
Overview of Metabolism
Biological Order

1.4 SYSTEMS BIOLOGY

Emergence
Robustness
Systems Biology Model Concepts
Living Organisms and Robustness

Biochemistry in the Lab

An Introduction

AVAILABLE ONLINE

Life: It Is a Mystery!

Why Study Biochemistry?

Why study biochemistry? For students embarking on careers in the life sciences, the answer should be obvious: biochemistry, the scientific discipline concerned with chemical processes within living organisms, is the bedrock on which all of the modern life sciences are built. During the past two decades, the influence of biochemistry and the allied field of molecular biology has increased exponentially. Life sciences as diverse as agronomy (the science of soil management and crop production), forensics, marine biology, plant biology, and ecology are now being explored with powerful biotechnological tools. As a result, there is now a vast array of career opportunities in federal or state government agencies and industry (e.g., pharmaceutical, biotechnology, and agribusiness companies) for recent graduates with life science degrees. Examples of such fields include biomedical and clinical research, forensic analysis, plant or animal genetics, environmental protection, and wildlife biology.

Economic conditions often dictate life science career choices. (The Occupational Outlook Handout on the U.S. Bureau of Labor Statistics website offers an unbiased assessment of future employment prospects.) No matter the economic conditions when students graduate, employment opportunities are always better for those who have undergraduate research experience. Developing a network of connections beginning with professors and expanding into the student's field or interests (e.g., by attending science career fairs or professional society conferences) also increases employment opportunities. Furthermore, writing, data analysis, problem solving, and communication are skills that employers always value highly. For students not interested in research careers, there are opportunities in science journalism, education, and software engineering. Examples of alternative careers where a life science degree is an asset include public policy (e.g., public health risk assessment and health product regulation), law (e.g., lawyers for pharmaceutical and biotech companies and environmental organizations), and marketing and sales (e.g., drugs and medical devices).

Overview

FROM MODEST BEGINNINGS IN THE LATE NINETEENTH CENTURY, THE SCIENCE OF BIOCHEMISTRY HAS PROVIDED INCREASINGLY MORE sophisticated intellectual and laboratory tools for the investigation of living processes. Today, in the early years of the twenty-first century, we find ourselves in the midst of a previously unimagined biotechnological revolution. Life sciences as diverse as medicine, agriculture, and forensics have generated immense amounts of information. The capacity to understand and appreciate the significance of this phenomenon begins with a thorough knowledge of biochemical principles. This chapter provides an overview of these principles. The chapters that follow focus on the structure and functions of the most important biomolecules and the major biochemical processes that sustain life.

This textbook is designed to provide an introduction to the basic principles of biochemistry. This opening chapter provides an overview of the major components of living organisms and the processes that sustain the living state. A brief description of the nature of the living state is followed by an introduction to the structures and functions of the major biomolecules and then an overview of the most important biochemical processes. The chapter concludes with a brief discussion of the concepts of modern experimental biochemistry and an introduction to *systems biology*, an investigative strategy used to improve our understanding of living organisms as integrated systems rather than collections of isolated components and chemical reactions.

1.1 WHAT IS LIFE?

What is life? Despite the work of life scientists over several centuries, a definitive answer to this deceptively simple question has been elusive. Much of the difficulty in delineating the precise nature of living organisms lies in the overwhelming diversity of the living world and the apparent overlap in several properties of living and nonliving matter. Consequently, life has been viewed as an intangible property and is usually described in operational terms, such as movement, reproduction, adaptation, and responsiveness to external stimuli. The work of life scientists, made possible by the experimental approaches of biochemistry, has revealed that all organisms obey the same chemical and physical laws that rule the universe. Life's diverse properties include the following:

1. **Life is complex and dynamic**. All organisms are composed of the same set of chemical elements, primarily carbon, nitrogen, oxygen, hydrogen, sulfur, and phosphorus. **Biomolecules**, the molecules synthesized by living organisms, are organic (carbon-based). Living processes, such as growth and development, involve thousands of chemical reactions in which vast quantities and varieties of vibrating and rotating molecules interact, collide, and rearrange into new molecules.
2. **Life is organized and self-sustaining**. Living organisms are hierarchically organized systems: they consist of patterns of organization from smallest (atom) to largest (organism) **(Figure 1.1)**. In biological systems, the functional capacities of each level of organization are derived from the structural and chemical properties of the level below it. Biomolecules are composed of atoms, which in turn are formed from subatomic particles. Certain biomolecules become linked to form polymers called **macromolecules**. Examples include nucleic acids, proteins, and polysaccharides, which are formed from nucleotides, amino

acids, and sugars, respectively. Cells are composed of a diversity of biomolecules and macromolecules, some of which form more complex supermolecular structures. At the molecular level, there are hundreds of biochemical reactions that together sustain the living state. Catalyzed by biomolecular catalysts called **enzymes**, these reactions are organized into pathways. (A *biochemical pathway* is a series of reactions in which a specific molecule is converted through a single or many steps into a terminal product.) The sum of all the reactions in a living organism is referred to as **metabolism**. The capacity of living organisms to regulate metabolic processes despite variability in their internal and external environments is called **homeostasis**. In multicellular organisms, other levels of organization include tissues, organs, and organ systems.

3. **Life is cellular**. Cells, the basic units of living organisms, differ widely in structure and function, but each is surrounded by a membrane that controls the transport of substances into and out of the cell. The membrane also mediates the response of the cell to the extracellular environment. If a cell is divided into its component parts, it will cease to function in a life-sustaining way. Cells arise only from the division of existing cells.
4. **Life is information-based**. Organization requires information. Living organisms can be considered information-processing systems because maintenance of their structural integrity and metabolic processes involves interactions among a vast array of molecules within and between cells. Biological information is expressed in the form of coded messages that are inherent in the unique three-dimensional structure of biomolecules. Genetic information stored in genes, the linear sequences of nucleotides in deoxyribonucleic acid (DNA), in turn specifies the linear sequence of amino acids in proteins and how and when those proteins are synthesized. Proteins perform their function by interacting with other molecules. The unique three-dimensional structure of each protein allows it to bind to, and interact with, a specific molecule that has a precise complementary shape. Information is transferred by the binding process. For example, the binding of the protein insulin to insulin receptor molecules on the surface of certain cells is a signal that initiates the uptake of the nutrient molecule glucose.

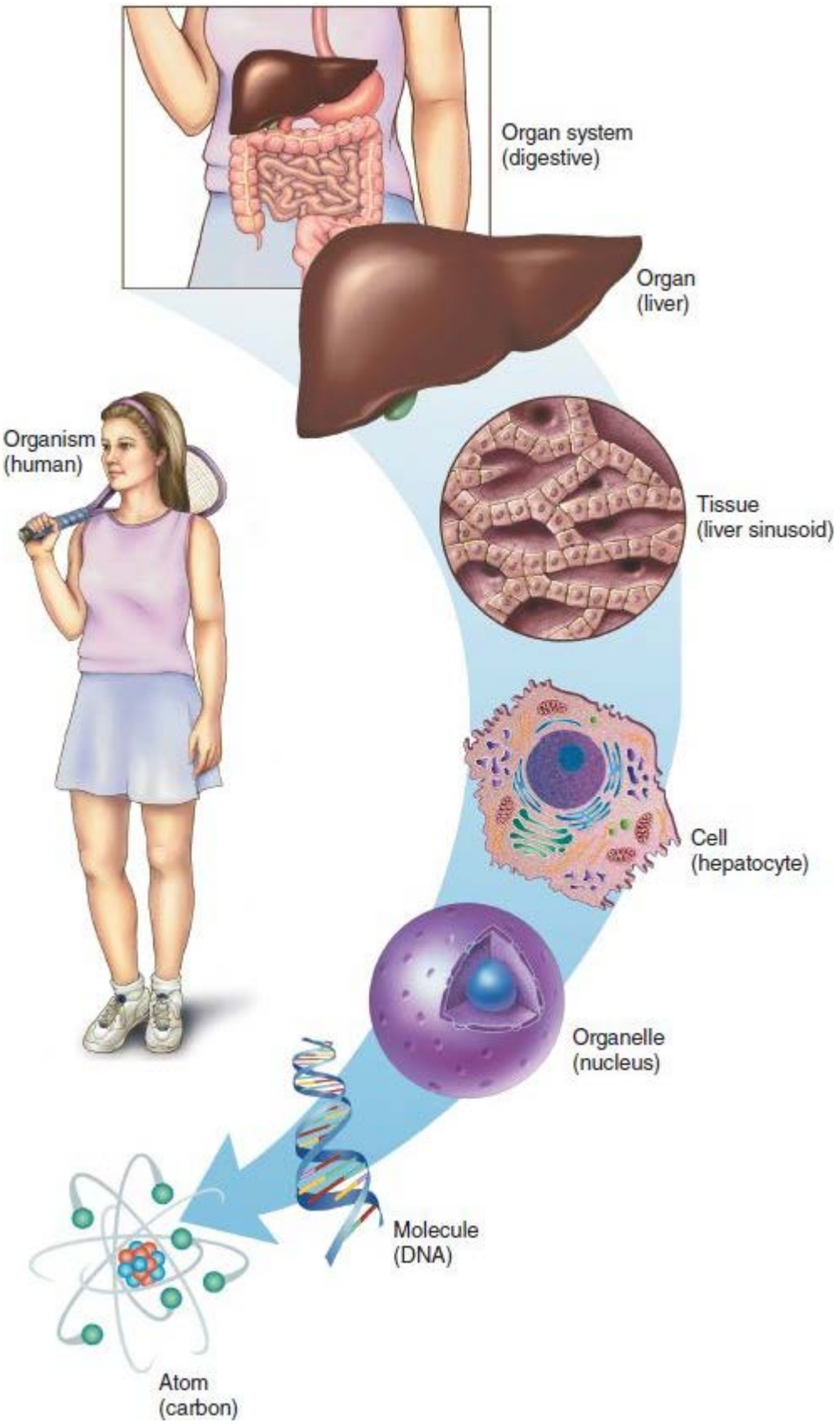


FIGURE 1.1

Hierarchical Organization of a Multicellular Organism: The Human Being

Multicellular organisms have several levels of organization: organ systems, organs, tissues, cells, organelles, molecules, and atoms. The digestive system and one of its component organs (the liver) are shown. The liver is a multifunctional organ that has several digestive functions. For example, it produces bile, which facilitates fat digestion, and it processes and distributes the food molecules absorbed in the small intestine to other parts of the body. DNA contains the genetic information that

controls cell function.

5. **Life adapts and evolves**. All life on Earth has a common origin, with new forms arising from older forms. When an individual organism in a population reproduces itself, **mutations** or sequence changes can arise as a result of stress-induced DNA modifications and errors that occur when DNA molecules are copied. Most mutations are silent: they either are repaired or have no effect on the functioning of the organism. Some, however, are harmful, serving to limit the reproductive success of the offspring. On rare occasions, mutations may contribute to an increased ability of the organism to survive, to adapt to new circumstances, and to reproduce. A principal driving force in this process is the capacity to exploit energy sources. Individuals possessing traits that allow them to better exploit a specific energy source within their habitat may have a competitive advantage when resources are limited. Over many generations, the interplay of environmental change and genetic variation can lead to the accumulation of favorable traits and eventually to increasingly different forms of life.

KEY CONCEPTS

- All living organisms obey the chemical and physical laws.
- Life is complex, dynamic, organized, and self-sustaining.
- Life is cellular and information-based.
- Life adapts and evolves.

1.2 BIOMOLECULES

Living organisms are composed of thousands of different kinds of inorganic and organic molecules. Water, an inorganic molecule, may constitute 50 to 95% of a cell's content by weight, and ions such as sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}), and calcium (Ca^{2+}) may account for another 1%. Almost all the other molecules in living organisms are organic. Life's organic molecules are principally composed of six elements: carbon, hydrogen, oxygen, nitrogen, phosphorus, and sulfur, and they contain trace amounts of certain metallic and other nonmetallic elements. The atoms of each of the most common elements found in living organisms can readily form stable covalent bonds, the kind that allow the formation of such important molecules as proteins.

The remarkable structural complexity and diversity of organic molecules are made possible by the capacity of carbon atoms to form four strong, single covalent bonds either to other carbon atoms or to atoms of other elements. Organic molecules with many carbon atoms can form complicated shapes such as long, straight structures or branched chains and rings.

Functional Groups of Organic Biomolecules

Most biomolecules can be considered derived from the simplest type of organic molecule, called the **hydrocarbons**. Hydrocarbons (**Figure 1.2**) are carbon- and hydrogen-containing molecules that are **hydrophobic**, or insoluble in water. All other organic molecules are formed by attaching other atoms or groups of atoms to the carbon backbone of the hydrocarbon. The chemical properties of these derivative molecules are determined by the specific arrangement of atoms

called **functional groups** (**Table 1.1**). For example, alcohols result when hydrogen atoms are replaced by hydroxyl groups (—OH). Thus methane (CH_4), a component of natural gas, can be converted into methanol (CH_3OH), a toxic liquid that is used as a solvent in many industrial processes.

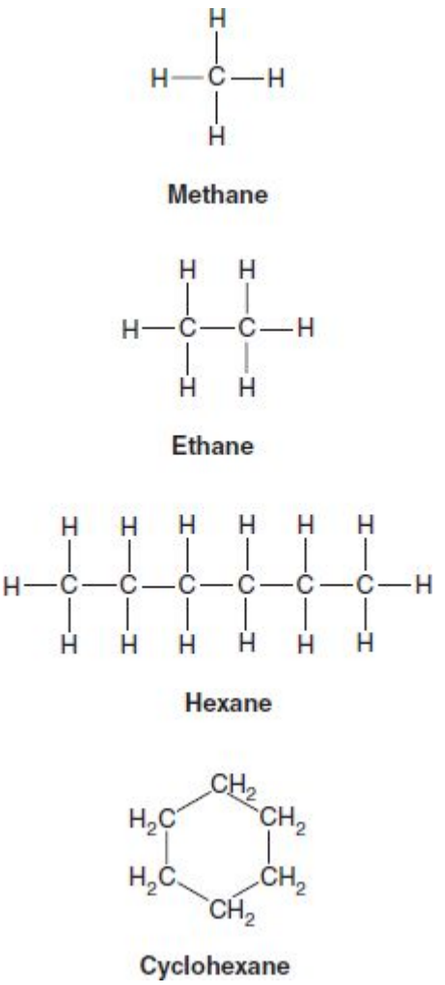
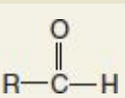
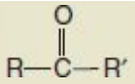
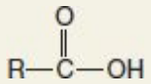
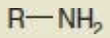
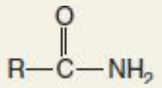
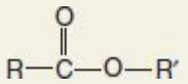
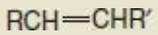


FIGURE 1.2
Structural Formulas of Several Hydrocarbons

Most biomolecules contain more than one functional group. For example, many simple sugar molecules have several hydroxyl groups and an aldehyde group. Amino acids, the building-block molecules of proteins, have both an amino group and a carboxyl group. The distinct chemical properties of each functional group contribute to the behavior of any molecule that contains it.

TABLE 1.1 Important Functional Groups in Biomolecules

Family Name	Group Structure	Group Name	Significance
Alcohol	R—OH	Hydroxyl	Polar (and therefore water-soluble), forms hydrogen bonds
Aldehyde		Carbonyl	Polar, found in some sugars

Ketone		Carbonyl	Polar, found in some sugars
Acids		Carboxyl	Weakly acidic, bears a negative charge when it donates a proton
Amine		Amino	Weakly basic, bears a positive charge when it accepts a proton
Amide		Amido	Polar but does not bear a charge
Thiol		Thiol	Easily oxidized; can form —S—S—(disulfide) bonds readily
Ester		Ester	Found in certain lipid molecules
Alkene		Double bond	Important structural component of many biomolecules (e.g., found in lipid molecules)

Major Classes of Small Biomolecules

Many of the organic compounds found in cells are relatively small, with molecular masses of less than 1000 daltons (Da). (One dalton, one atomic mass unit, is equal to $^{1}/_{12}$ of the mass of one atom of ^{12}C.) Cells contain four families of small molecules: amino acids, sugars, fatty acids, and nucleotides (**Table 1.2**). Members of each group serve several functions. First, they are used in the synthesis of larger molecules, many of which are polymers. For example, proteins, certain carbohydrates, and nucleic acids are polymers composed of amino acids, sugars, and nucleotides, respectively. Fatty acids are components of lipid (water-insoluble) molecules of several types.

TABLE 1.2 Major Classes of Biomolecules

Small Molecule	Polymer	General Functions
Amino acids	Proteins	Catalysts and structural elements
Sugars	Carbohydrates	Energy sources and structural elements
Fatty acids	N.A.	Energy sources and structural elements of complex lipid molecules
Nucleotides	DNA	Genetic information
	RNA	Protein synthesis

Second, some molecules have special biological functions. For example, the nucleotide adenosine triphosphate (ATP) serves as a cellular reservoir of chemical energy. Finally, many small organic molecules are involved in complex reaction pathways. Examples of each class of molecule are described next.

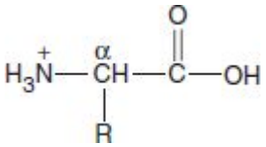


FIGURE 1.3

General Formula for *α*-Amino Acids

For 19 of the 20 standard amino acids found in proteins, the *α*-carbon is bonded to a hydrogen atom, a carboxyl group, an amino group, and an R group.

AMINO ACIDS AND PROTEINS There are hundreds of naturally occurring **amino acids**, each of which contains an amino group and a carboxyl group. Amino acids are classified *α*, *β*, or *γ* according to the location of the amino group in reference to the carboxyl group. In *α*-amino acids, the most common type, the amino group is attached to the carbon atom (the *α*-carbon) immediately adjacent to the carboxyl group (**Figure 1.3**). In *β*- and *γ*-amino acids, the amino group is attached to the second and third carbon, respectively, from the carboxyl group. Also attached to the *α*-carbon is another group, referred to as the side chain or R group. The chemical properties of each amino acid, once incorporated into protein, are determined largely by the properties of its side chain. For example, some side chains are hydrophobic (i.e., they have low solubility in water), whereas others are **hydrophilic** (i.e., they dissolve easily in water). Several examples of *α*-amino acids are presented in **Figure 1.4**.

Twenty standard *α*-amino acids occur in proteins. Some standard amino acids have unique functions in living organisms. For example, glycine and glutamic acid function in animals as **neurotransmitters**, signal molecules released by nerve cells. Proteins also contain nonstandard amino acids that are modified versions of the standard amino acids. The structure and function of protein molecules are often altered by conversion of certain amino acid residues to derivatives via phosphorylation, hydroxylation, and other chemical modifications. (The term *residue* refers to a small biomolecule that is incorporated in a macromolecule, e.g., amino acid residues in a protein.) For example, many of the residues of proline are hydroxylated in collagen, the connective tissue protein. Many naturally occurring amino acids are not *α*-amino acids. Prominent examples include *β*-alanine, a precursor of the vitamin pantothenic acid, and *γ*-aminobutyric acid (GABA), a neurotransmitter found in the brain (**Figure 1.5**).

Amino acid molecules are used primarily in the synthesis of long, complex polymers known as **polypeptides**. Up to a length of about 50 amino acids, these molecules are called **peptides**. **Proteins** consist of one or more polypeptides. Polypeptides play a variety of roles in living organisms, including transport, structure, and catalysis.

The individual amino acids are connected in peptides (**Figure 1.6**) and polypeptides by peptide bonds. **Peptide bonds** are amide linkages that form in a nucleophilic substitution reaction (p. P-30) in which the amino group nitrogen of one amino acid attacks the carbonyl carbon in the activated carboxyl group of another. For many proteins, their three-dimensional structure and biological function result largely from interactions among the R groups (**Figure 1.7**).

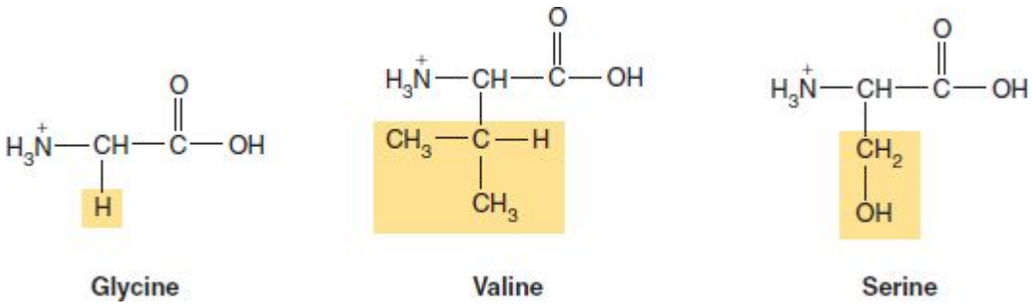


FIGURE 1.4

Structural Formulas for Several *α*-Amino Acids

An R group (highlighted) in an amino acid structure can be a hydrogen atom (e.g., in glycine), a hydrocarbon group (e.g., the isopropyl group in valine), or a hydrocarbon derivative (e.g., the hydroxy methyl group in serine).

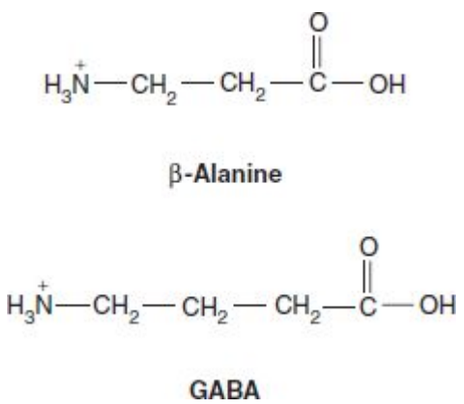


FIGURE 1.5

Select Examples of Naturally Occurring Amino Acids That Are Not *α*-Amino Acids: *β*-Alanine and *γ*-Aminobutyric Acid (GABA)

WORKED PROBLEM 1.1

Living organisms generate a vast number of different biopolymers by linking monomers in different sequences. A set of tripeptides, each containing three amino acid residues, contains only two types of amino acids: A and B. How many possible tripeptides are in this set?

SOLUTION

The number of possible tripeptides is given by the formula $\mathbf{X}^n$, where

$\mathbf{X}$ = the number of types of constituent amino acid residues

n = length of the peptide.

Substituting these values into the formula yields $2^3 = 8$. The eight tripeptides are as follows: AAA, AAB, ABA, BAA, ABB, BAB, BBA, and BBB. ■

SUGARS AND CARBOHYDRATES **Sugars**, the smallest carbohydrates, contain alcohol and carbonyl functional groups. They are described in terms of both carbon number and the type of carbonyl group they contain. Sugars that possess an aldehyde group are called *aldoses*, and those that possess a ketone group are called *ketoses*. For example, the six-carbon sugar glucose (an important energy source in most living organisms) is an aldohexose; fructose (fruit sugar) is a ketohexose (Figure 1.8).

Sugars are the basic units of carbohydrates, the most abundant organic molecules found in nature. Carbohydrates range from the simple sugars, or **monosaccharides**, such as glucose and fructose, to the **polysaccharides**, polymers that contain thousands of sugar units. Examples of polysaccharides include starch and cellulose in plants and glycogen in animals. Carbohydrates serve a variety of functions in living organisms. Certain sugars are important energy sources. Glucose is the principal carbohydrate energy source in animals and plants. Plants use sucrose as an efficient means of transporting energy throughout their tissues. Some carbohydrates serve as structural materials. Cellulose is the major structural component of wood and certain plant fibers. Chitin, another type of polysaccharide, is found in the exoskeletons of insects and crustaceans.

Some biomolecules contain carbohydrate components. Nucleotides, the building-block molecules of the nucleic acids, contain either of the pentoses ribose or deoxyribose. Certain proteins and lipids also contain carbohydrate. Glycoproteins and glycolipids occur on the external surface of cell membranes in multicellular organisms, where they play critical roles in the interactions between cells.

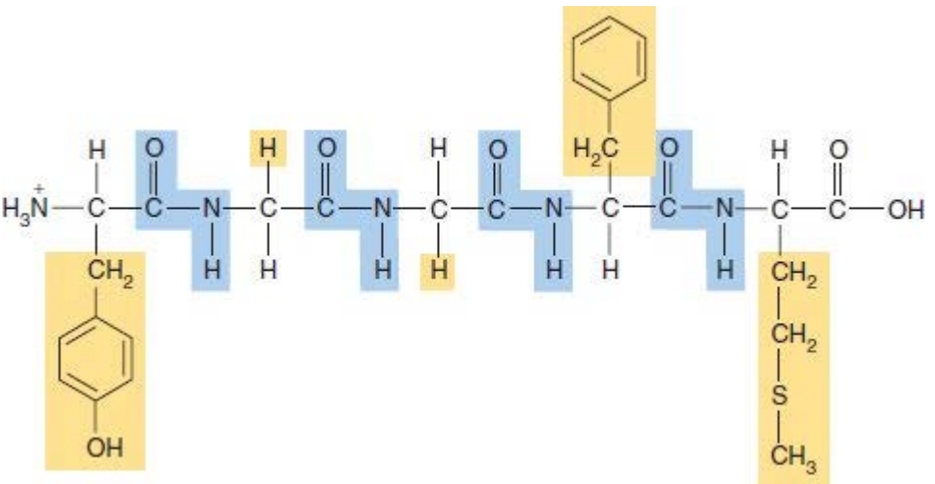


FIGURE 1.6

Structure of Met-Enkephalin, a Pentapeptide

Met-enkephalin is one of a class of molecules that have opiate-like activity. Found in the brain, met-enkephalin inhibits pain perception. (The peptide bonds are colored blue. The R groups are highlighted.)

3D animation of Met-Enkephalin

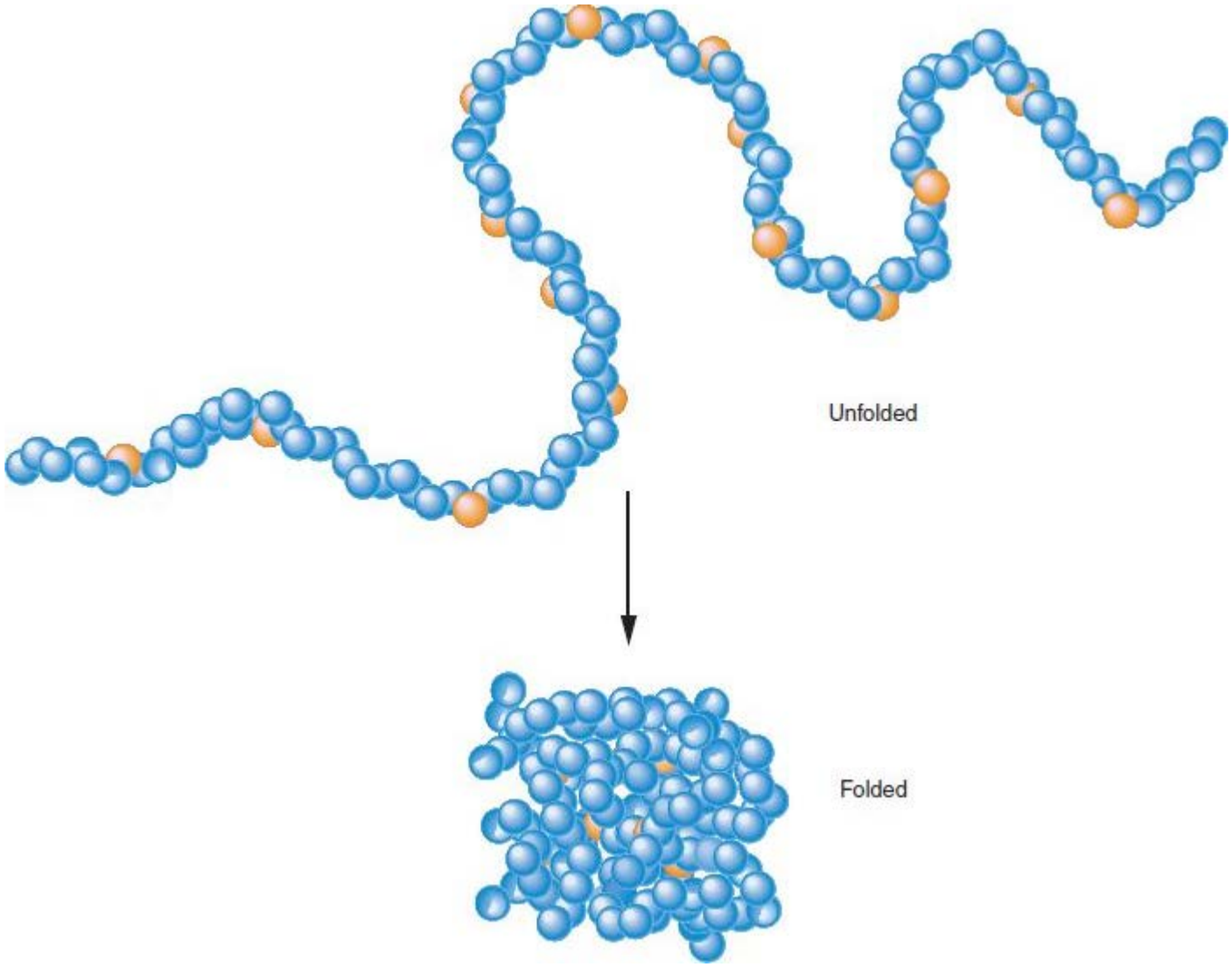


FIGURE 1.7

Polypeptide Structure

As a polypeptide folds into its unique three-dimensional form, hydrophobic R groups (yellow spheres) become buried in the interior away from water. Hydrophilic groups usually occur on the surface.

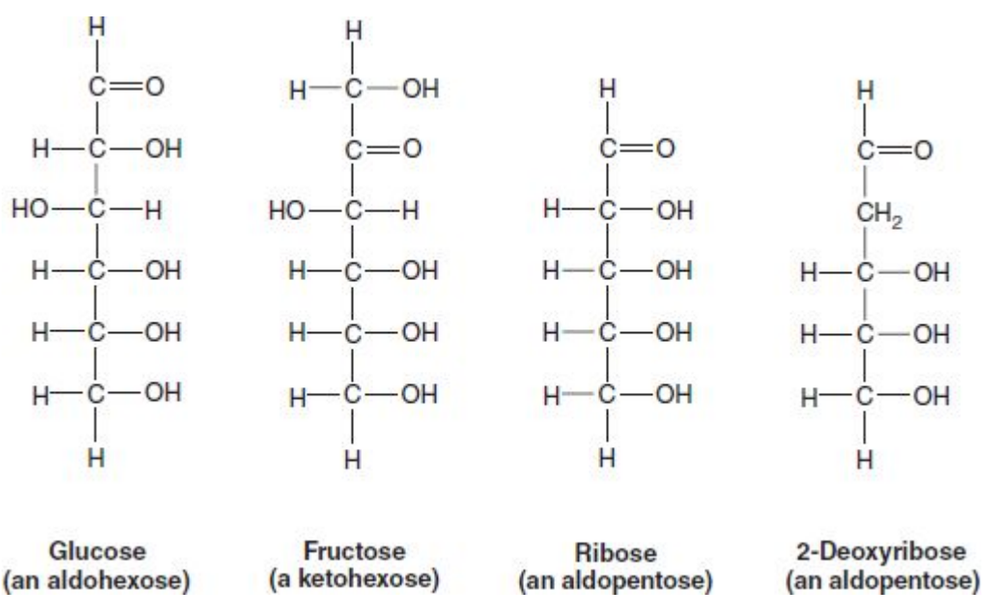


FIGURE 1.8

Some Biologically Important Monosaccharides

Glucose and fructose are important sources of energy in plants and animals. Ribose and deoxyribose are components of nucleic acids. These monosaccharides occur as ring structures in nature.

3D animation of fructose, d-

3D animation of ribose

3D animation of 2-deoxyribose

3D animation of fructose, l-

FATTY ACIDS **Fatty acids** are monocarboxylic acids that usually contain an even number of carbon atoms. Fatty acids are represented by the chemical formula R—COOH, in which R is an alkyl group that contains carbon and hydrogen atoms. There are two types of fatty acids: **saturated** fatty acids, which contain no carbon–carbon double bonds, and **unsaturated** fatty acids, which have one or more double bonds (**Figure 1.9**). Under physiological conditions, the carboxyl group of fatty acids exists in the ionized state, $R—COO^-$. For example, the 16-carbon saturated fatty acid called palmitic acid usually exists as palmitate, $CH_3(CH_2)_{14}COO^-$. Although the charged carboxyl group has an affinity for water, the long nonpolar hydrocarbon chains render most fatty acids insoluble in water.

Fatty acids occur as independent (free) molecules in only trace amounts in living organisms. Most often they are components of several types of **lipid** molecules (**Figure 1.10**). Lipids are a diverse group of substances that are soluble in organic solvents such as chloroform or acetone, but they are not soluble in water. For example, triacylglycerols (fats and oils) are esters containing glycerol (a three-carbon alcohol with three hydroxyl groups) and three fatty acids. Phosphoglycerides contain two fatty acids. In these molecules, the third hydroxyl group of glycerol is coupled with phosphate, which is in turn attached to small polar compounds such as choline. Phosphoglycerides are an important structural component of cell membranes.

NUCLEOTIDES AND NUCLEIC ACIDS Each **nucleotide** contains three components: a five-carbon sugar (either ribose or deoxyribose), a nitrogenous base, and one or more phosphate groups (**Figure 1.11**). The bases in nucleotides are heterocyclic aromatic rings with a variety of substituents. There are two classes of base: the bicyclic purines and the monocyclic pyrimidines (Figure 1.12).

Nucleotides participate in a wide variety of biosynthetic and energy-generating reactions. For example, the energy obtained from food molecules is used to form the high-energy phosphate bonds of ATP. The energy, released when the phosphoanhydride bonds are hydrolyzed, drives

cellular processes. Nucleotides also have an important role as the building-block molecules of the nucleic acids. In a **nucleic acid**, dozens to millions of nucleotides are linked by phosphodiester linkages to form long polynucleotide chains or strands. There are two types of nucleic acid: DNA and RNA.

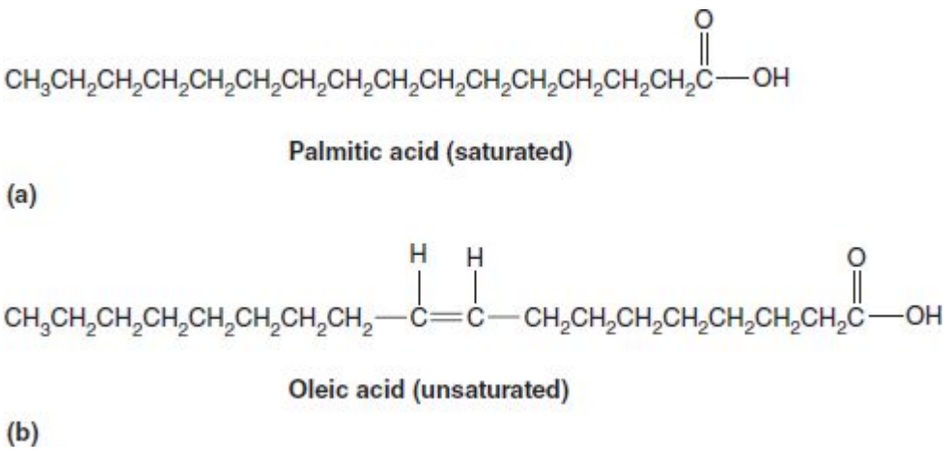


FIGURE 1.9

Fatty Acid Structure

(a) A saturated fatty acid. (b) An unsaturated fatty acid.

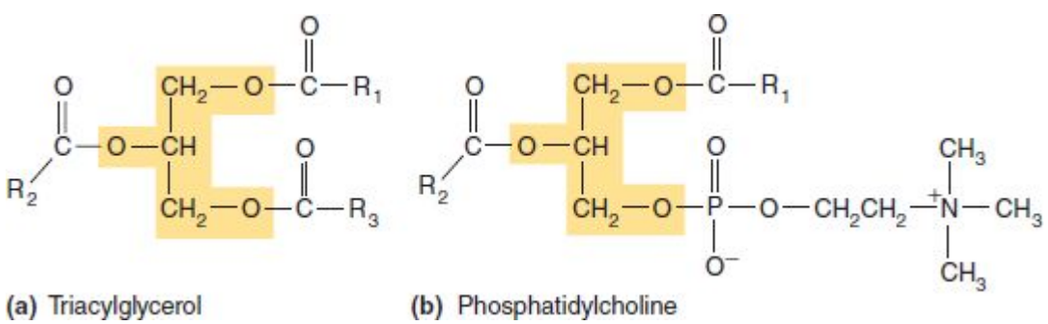


FIGURE 1.10

Lipid Molecules That Contain Fatty Acids

(a) Triacylglycerol. (b) Phosphatidylcholine, a type of phosphoglyceride.

3D animation of triacylglycerol

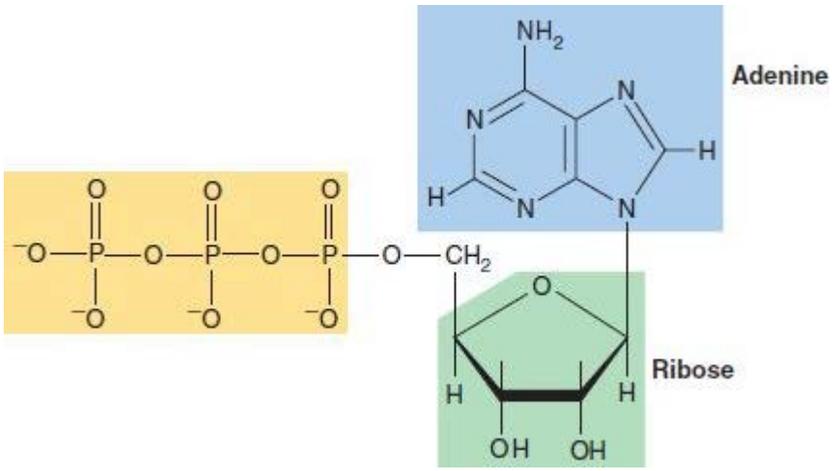


FIGURE 1.11

Nucleotide Structure

Each nucleotide contains a nitrogenous base (in this case, adenine), a pentose sugar (ribose), and one or more phosphates. This nucleotide is adenosine triphosphate.

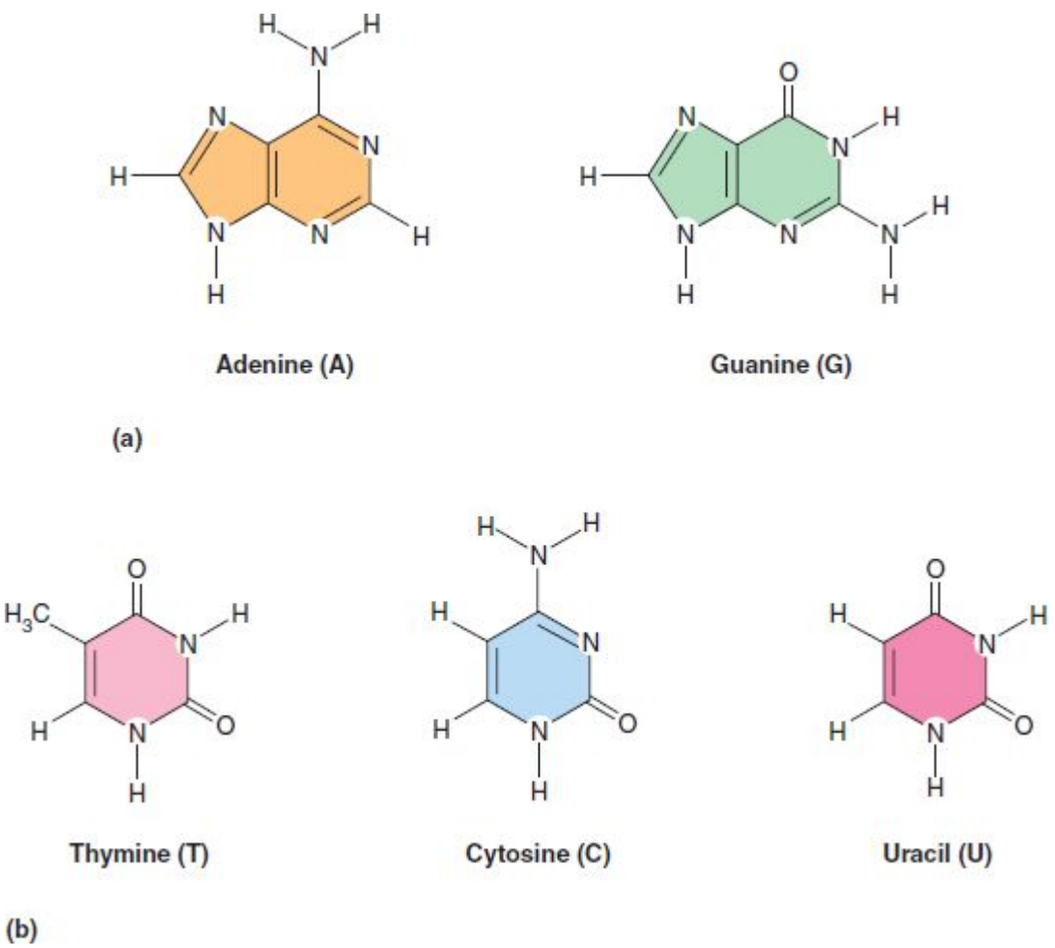


FIGURE 1.12

The Nitrogenous Bases

(a) The purines. (b) The pyrimidines.

3D animation of adenine

3D animation of Thymine

3D animation of Cytosine

3D animation of uracil

DNA DNA is the repository of genetic information. Its structure consists of two antiparallel polynucleotide strands wound around each other to form a right-handed double helix (**Figure 1.13**). In addition to the pentose sugar deoxyribose and phosphate, DNA contains bases of four types: the **purines** adenine and guanine and the **pyrimidines** thymine and cytosine; adenine pairs with thymine and guanine pairs with cytosine. The double helix forms because of complementary pairing between the bases due to hydrogen bonding. A hydrogen bond is a force of attraction between a polarized hydrogen of one molecular group and the electronegative oxygen or nitrogen atoms of nearby aligned molecular groups.

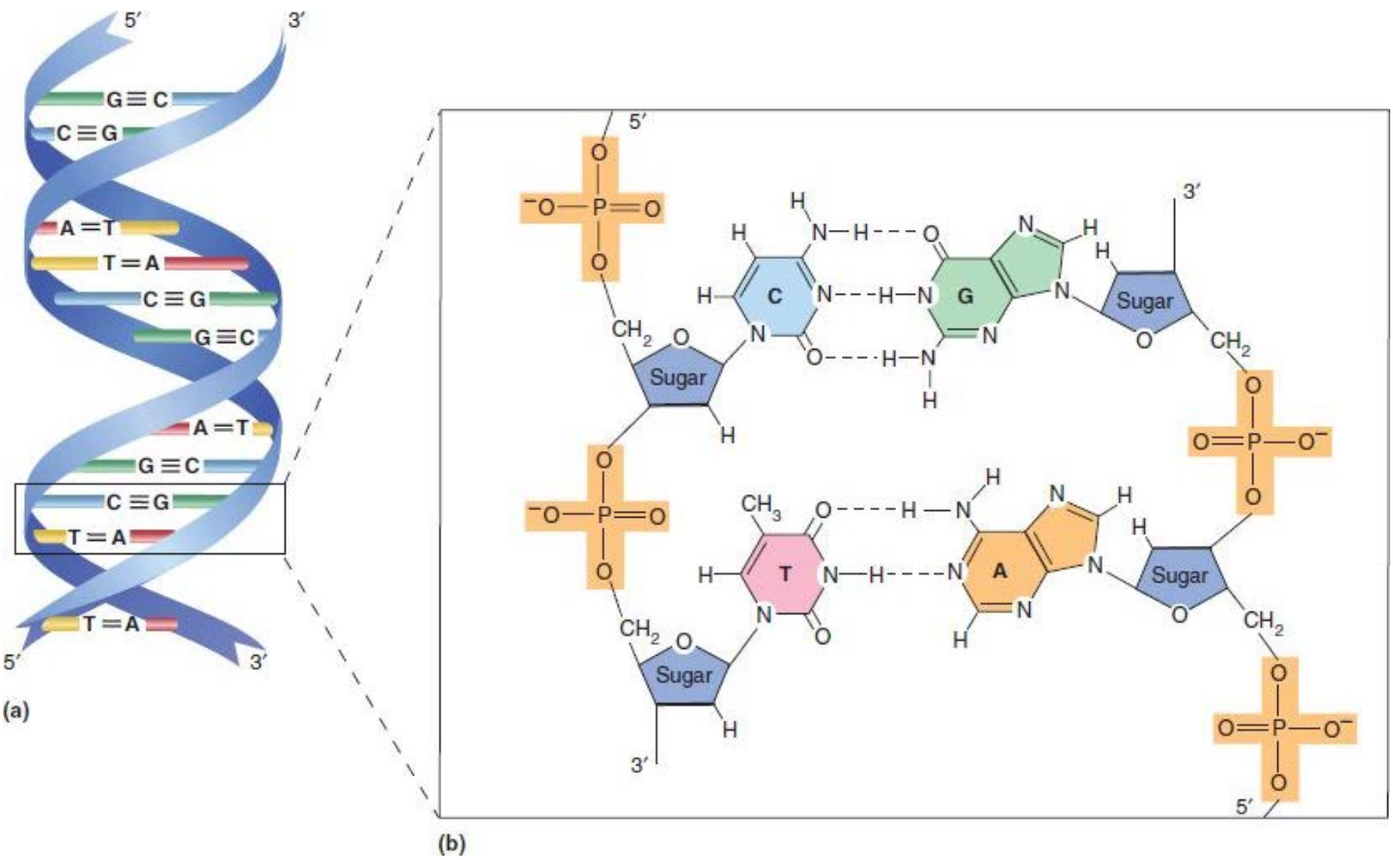


FIGURE 1.13

DNA

(a) A diagrammatic view of DNA. The sugar-phosphate backbones of the double helix are represented by colored ribbons. The bases attached to the sugar deoxyribose are on the inside of the helix. (b) An enlarged view of two base pairs. Note that the two DNA strands run in opposite directions defined by the 5′ and 3′ groups of deoxyribose. The bases on opposite strands form pairs because of hydrogen bonds. Cytosine always pairs with guanine; thymine always pairs with adenine.

An organism's entire set of DNA sequences is called its **genome**. DNA consists of both coding and noncoding sequences. Coding sequences, called **genes**, specify the structure of functional gene products such as polypeptides and RNA molecules. Some noncoding sequences have regulatory functions (e.g., controlling the synthesis of certain proteins), whereas the functions of others are as yet undetermined.

RNA Ribonucleic acid (RNA) is a polynucleotide that differs from DNA in that it contains the sugar ribose instead of deoxyribose and the base uracil instead of thymine. In RNA, as in DNA, the nucleotides are linked by phosphodiester linkages. In contrast to the double helix of DNA, RNA is single-stranded. RNA molecules fold into complex three-dimensional structures created by local regions of complementary base pairing. When the DNA double helix unwinds, one strand can serve as a template. RNA molecules are synthesized in a process called **transcription**. Complementary base pairing between DNA bases and the bases of incoming ribonucleotides specifies the base sequence of the RNA molecule. There are three major types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). Each unique sequence or molecule of mRNA possesses the information that codes directly for the amino acid sequence in a specific polypeptide. Ribosomes, large, complex, supramolecular structures composed of rRNA and protein molecules, convert the mRNA base sequence into the amino acid sequence of a polypeptide. Transfer RNA molecules deliver activated amino acids to the ribosome during protein synthesis.

In recent years, large numbers of RNA molecules have been discovered that are not directly involved in protein synthesis. These molecules, called *noncoding RNAs* (ncRNA), have roles in a great variety of cellular processes. Examples include short interfering RNAs (siRNAs), micro RNAs (miRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and long noncoding RNAs (lncRNAs). siRNAs are important components in *RNA interference*, an antiviral defense mechanism. miRNAs have diverse roles in gene regulation including binding to and silencing specific mRNAs. snRNAs facilitate the process by which mRNA precursor molecules are transformed into functional mRNA. snoRNAs assist in the maturation of ribosomal RNA during ribosome formation. lncRNAs play vital roles in gene expression.

KEY CONCEPTS

- Most molecules in living organisms are organic. The chemical properties of organic molecules are determined by specific arrangements of atoms called functional groups.
- Cells contain four families of small molecules: amino acids, sugars, fatty acids, and nucleotides.
- Proteins, polysaccharides, and the nucleic acids are biopolymers composed of amino acids, sugars, and nucleotides, respectively.

GENE EXPRESSION **Gene expression** controls when or if the information encoded in a gene will be accessed. The process begins with transcription, the mechanism whereby the base sequence of a DNA segment is used to synthesize a gene product. A class of proteins called **transcription factors** regulates the expression of protein-coding genes when they bind to specific regulatory DNA sequences referred to as **response elements**. Transcription factors are synthesized and/or

regulated in response to an information-processing mechanism initiated by a signal molecule (e.g., insulin, a protein that regulates several metabolic processes) or an abiotic factor such as light.

1.3 IS THE LIVING CELL A CHEMICAL FACTORY?

Even the simplest cells are so remarkable that they have often been characterized as chemical factories. Like factories, living organisms acquire raw materials, energy, and information from their environment. Components are manufactured, and waste products and heat are discharged back into the environment. However, for this analogy to hold true, human-made factories would not only manufacture and repair all their structural and functional components, but also clone themselves, that is, manufacture new factories. The term **autopoiesis** has been created to describe the remarkable properties of living organisms. In this view, each living organism is considered an autonomous, self-organizing, and self-maintaining entity. Life emerges from a self-regulating network of thousands of biochemical reactions.

The constant flow of energy and nutrients through organisms and the functional properties of thousands of enzymes make possible the process of metabolism. The primary functions of metabolism are (1) acquisition and utilization of energy, (2) synthesis of molecules needed for cell structure and functioning (i.e., proteins, carbohydrates, lipids, and nucleic acids), (3) growth and development, and (4) removal of waste products. Metabolic processes require significant amounts of useful energy. This section begins with a review of the primary chemical reaction types and the essential features of energy-generating strategies observed in living organisms. A brief outline of metabolic processes and the means by which living organisms maintain ordered systems follows.

Biochemical Reactions

At first glance, the thousands of reactions that occur in cells appear overwhelmingly complex. However, several characteristics of metabolism allow us to simplify this picture:

1. Although the number of reactions is large, the number of reaction types is relatively small.
2. Biochemical reactions have simple organic reaction mechanisms.
3. Reactions of central importance in biochemistry (i.e., those used in energy production and the synthesis and degradation of major cell components) are relatively few.

Among the most common reaction types encountered in biochemical processes are nucleophilic substitution, elimination, addition, isomerization, and oxidation-reduction.

NUCLEOPHILIC SUBSTITUTION REACTIONS In **nucleophilic substitution** reactions, as the name suggests, one atom or group is substituted for another:

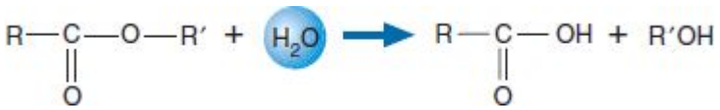
In the general reaction shown, the attacking species (A) is called a **nucleophile** (“nucleus lover”). Nucleophiles are anions (negatively charged atoms or groups) or neutral species possessing nonbonding electron pairs. **Electrophiles** (“electron lovers”) are deficient in electron density and are therefore easily attacked by a nucleophile. As the new bond forms between A and B, the old one between B and X breaks. The outgoing nucleophile (in this case, X), called a **leaving group**, leaves with its electron pair. Several types of nucleophilic substitution reactions

occur in living organisms. Examples include S_N2 reactions (e.g., the methylation of epinephrine; refer to p. P-33), acyl group transfers, and phosphoryl group transfers.

In nucleophilic substitution reactions involving acyl transfer, a nucleophile attacks the carbonyl carbon of a carboxylic acid derivative, forming a tetrahedral intermediate. The carbonyl group reforms as the tetrahedral intermediate collapses and the leaving group is ejected. Biologically important examples of carboxylic acid derivatives include carboxylates (deprotonated carboxylic acids), esters, amides, thioesters, and acyl phosphates. These derivatives vary in their reactivity with nucleophiles. Acyl phosphates are the most reactive, followed by thioesters, esters, amides, and finally carboxylates.

The biologically active form of fatty acids is the thioester of coenzyme A (p. 341). Carboxylates are not good substrates for nucleophilic substitution reactions because the carbonyl carbon is not sufficiently electrophilic. As a result, fatty acids must first be activated by the formation of an acyl adenosyl monophosphate derivative (**Figure 1.14**) at the expense of ATP bond energy. Once the activated fatty acyl-AMP is formed, its carbonyl carbon is easily attacked by the thiol sulfur of coenzyme A (CoASH) to yield the fatty acyl-SCoA product.

Hydrolysis reactions are nucleophilic acyl substitution reactions in which the oxygen of a water molecule serves as the nucleophile.



The electrophile is usually the carbonyl carbon of an ester, amide, or anhydride. (An **anhydride** is a molecule containing two carbonyl groups linked through an oxygen atom. A **mixed anhydride** is an anhydride formed from two different acids. For example, glycerate-1,3-bisphosphate is an important metabolic molecule.) The digestion of many food molecules involves hydrolysis. For example, the amide linkages of proteins are hydrolyzed in the stomach in an acid-catalyzed reaction that yields amino acids.

The hydrolysis of ATP to yield ADP and inorganic phosphate (P_i) and the reaction of glucose with ATP provide two examples of nucleophilic substitution involving phosphoryl group transfer. The attack by the OH of water on the terminal phosphate of ATP (**Figure 1.15**) breaks the phosphoanhydride bond, thereby releasing energy that is used to drive many cellular processes.

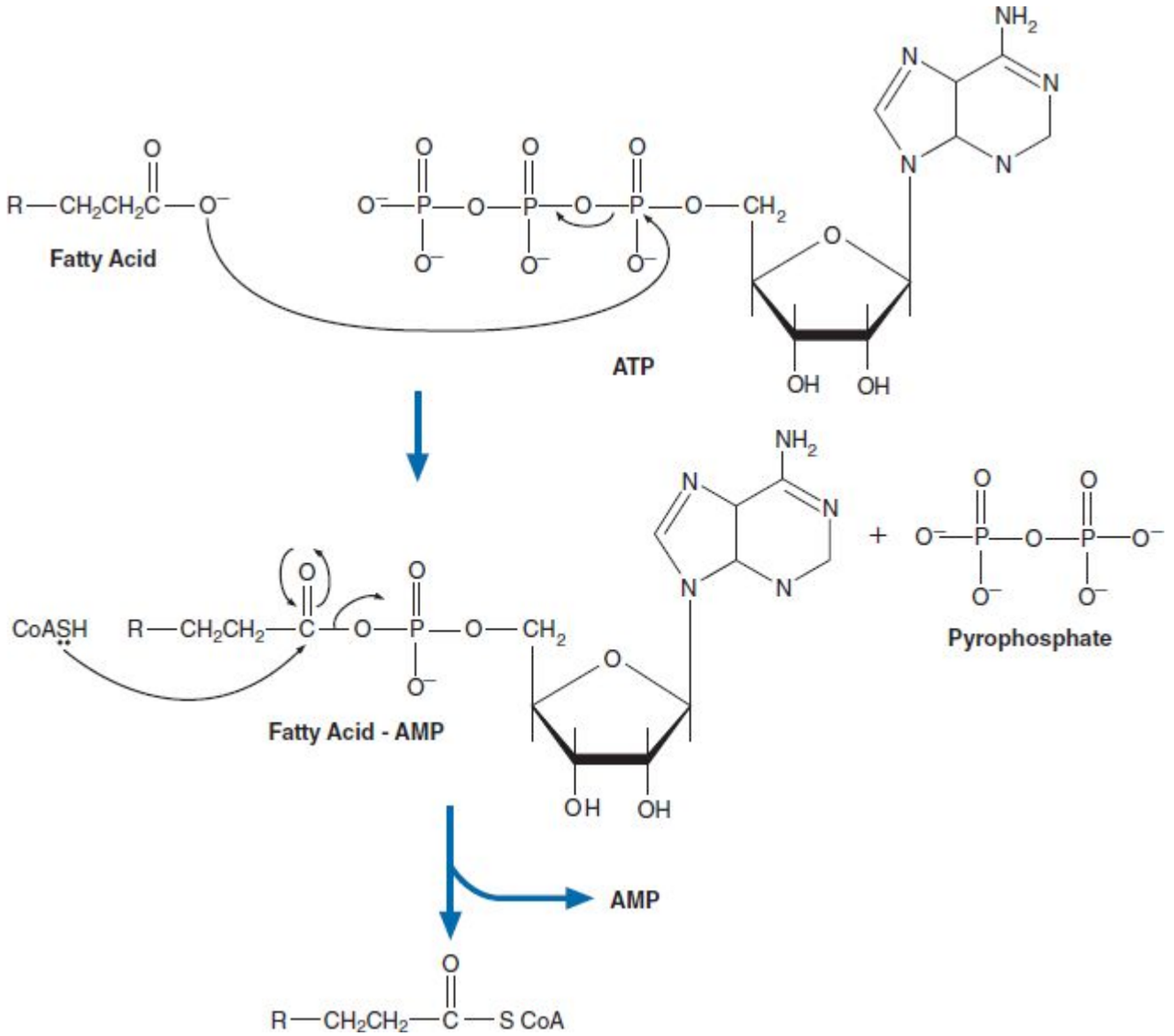


FIGURE 1.14

Activation of a Fatty Acid

Before a fatty acid can be degraded to yield energy or used in the synthesis of a triacylglycerol, it must first be activated. In the first step, the carboxylate ion attacks a phosphate of ATP to form a fatty acyl-AMP intermediate and pyrophosphate (PP_i). In the second step, the fatty acyl-AMP is attacked by the thiol group of coenzyme A (CoASH) to form the thioester fatty acyl-SCoA and AMP. The rapid hydrolysis of PP_i to form two phosphates (P_i) drives the reaction forward.

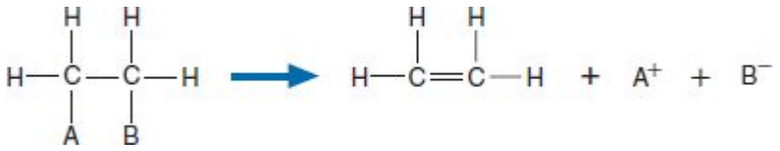
3D animation of ATP

3D animation of AMP

The reaction of glucose with ATP, yielding glucose-6-phosphate and ADP, is the first step in the utilization of glucose as an energy source (**Figure 1.16**). The hydroxyl oxygen on carbon 6 of the

sugar molecule is the nucleophile, and phosphorus is the electrophile. Adenosine diphosphate is the leaving group.

ELIMINATION REACTIONS In **elimination reactions,** a double bond is formed when atoms in a molecule are removed.



The removal of H_2O from biomolecules containing alcohol functional groups is a commonly encountered reaction. A prominent example is the dehydration of 2-phosphoglycerate, a reaction in *glycolysis*, which is a biochemical pathway in carbohydrate metabolism (**Figure 1.17**). As illustrated on pp. P-33–P-34, this reaction occurs via an E1cB mechanism. Other products of elimination reactions include ammonia (NH_3), amines (RNH_2), and alcohols (ROH).

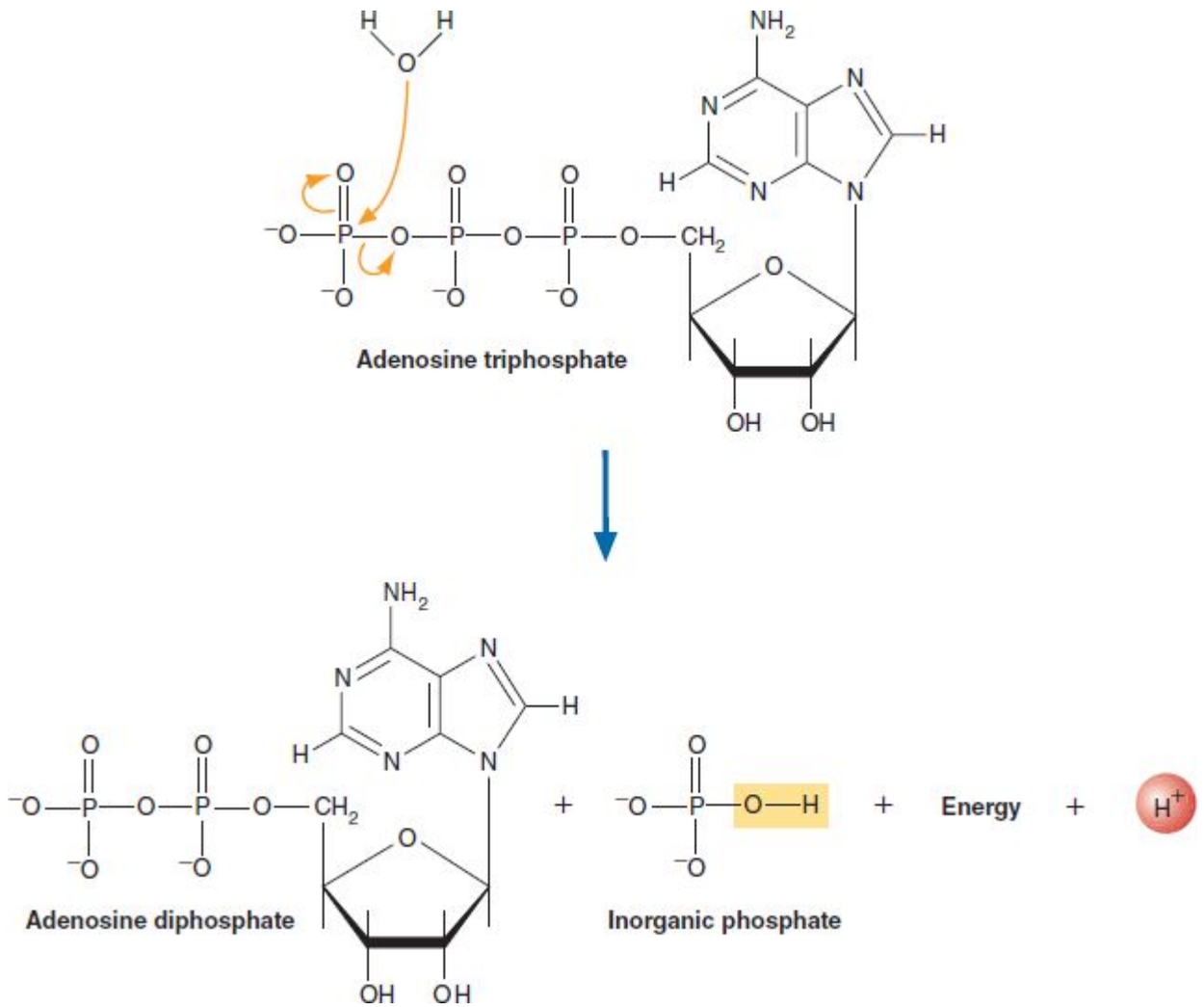


FIGURE 1.15

A Hydrolysis Reaction

The hydrolysis of ATP, a nucleophilic substitution reaction involving phosphoryl transfer, is used to drive an astonishing diversity of energy-requiring biochemical reactions.

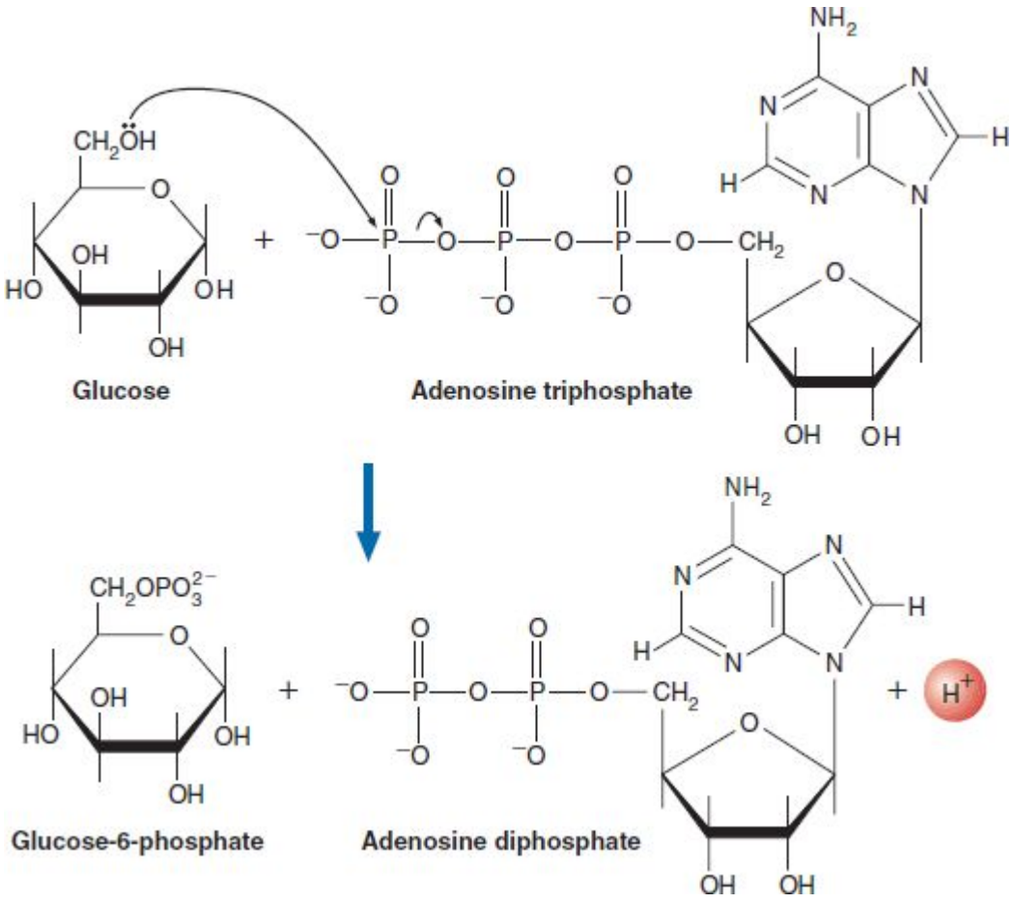


FIGURE 1.16

Example of Nucleophilic Substitution

In the reaction of glucose with ATP, the hydroxyl oxygen of glucose is the nucleophile. The phosphorus atom (the electrophile) is polarized by the oxygens bonded to it so that it bears a partial positive charge. As the reaction occurs, the unshared pair of electrons on the hydroxyl oxygen of CH_2OH of the sugar attacks the phosphorus, resulting in the expulsion of ADP, the leaving group.

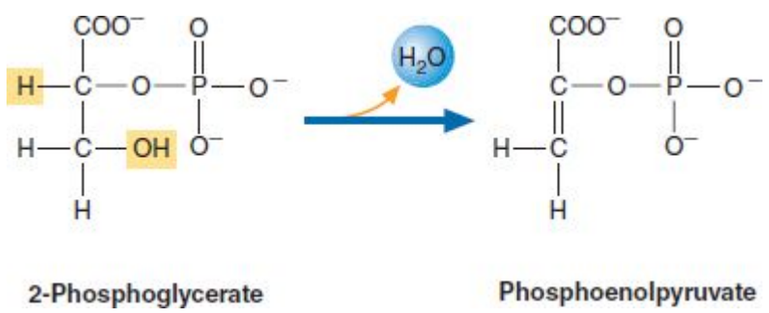
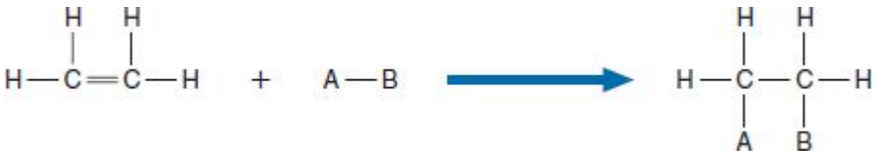


FIGURE 1.17

An Elimination Reaction

When 2-phosphoglycerate is dehydrated, a double bond is formed. This reaction involves an E1cB mechanism, which is illustrated on p. P-34.

ADDITION REACTIONS In **addition reactions,** two molecules combine to form a single product.



Hydration is one of the most common addition reactions. When water is added to an alkene, an alcohol results. The hydration of the metabolic intermediate fumarate to form malate is a typical example (**Figure 1.18**).

ISOMERIZATION REACTIONS In **isomerization reactions**, atoms or groups undergo intramolecular shifts. One of the most common biochemical isomerizations is the interconversion between aldose and ketose sugars (**Figure 1.19**). The isomerization of dihydroxyacetone phosphate to glyceraldehyde-3-phosphate (**Figure 1.19b**) is a reaction in glycolysis.

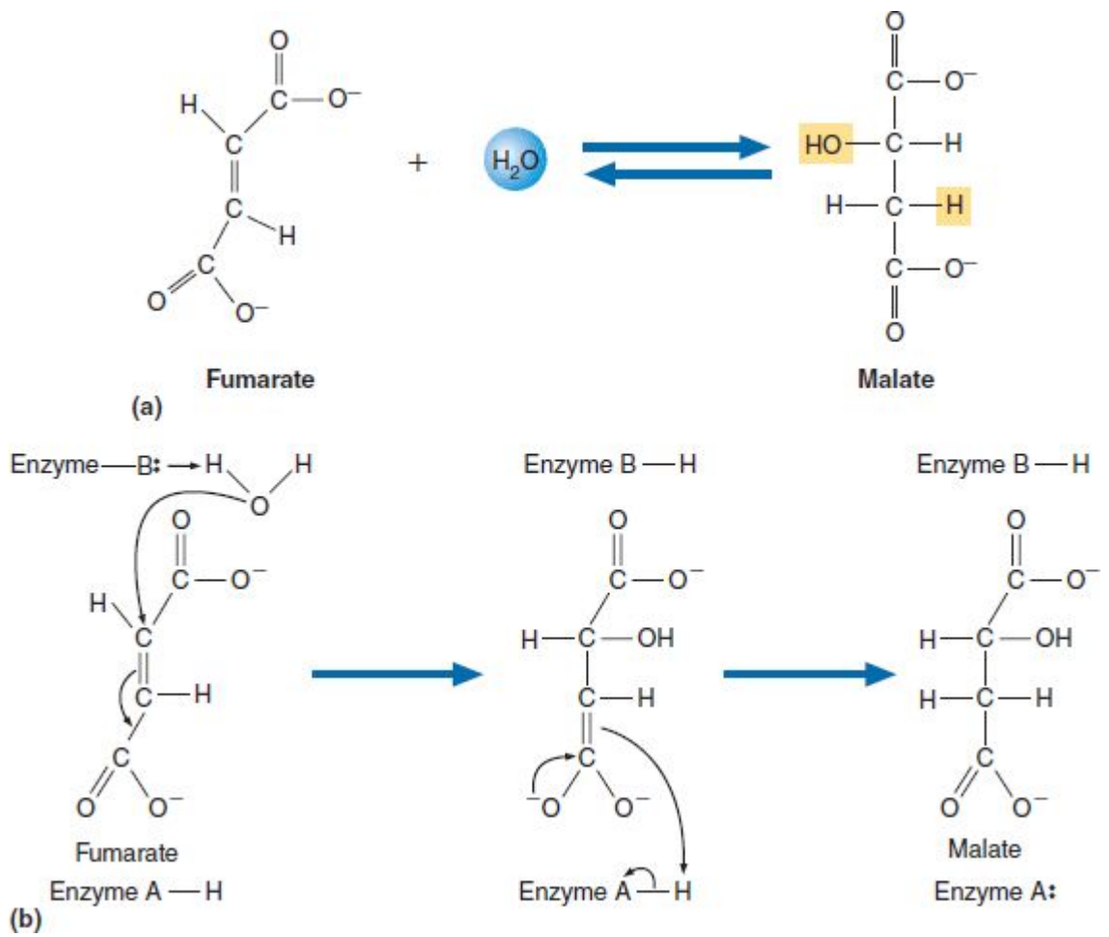


FIGURE 1.18

An Addition Reaction

(a) When water is added to a molecule that contains a double bond, such as fumarate, an alcohol results. (b) The hydration of fumarate, catalyzed by the enzyme fumarase, begins with the removal of a proton from a water molecule by an amino acid side chain acting as a base. The resulting nucleophile attacks the carbon–carbon double bond. The initial product, a resonance-stabilized ion, is then protonated by an acidic side chain of the enzyme to yield the product malate.

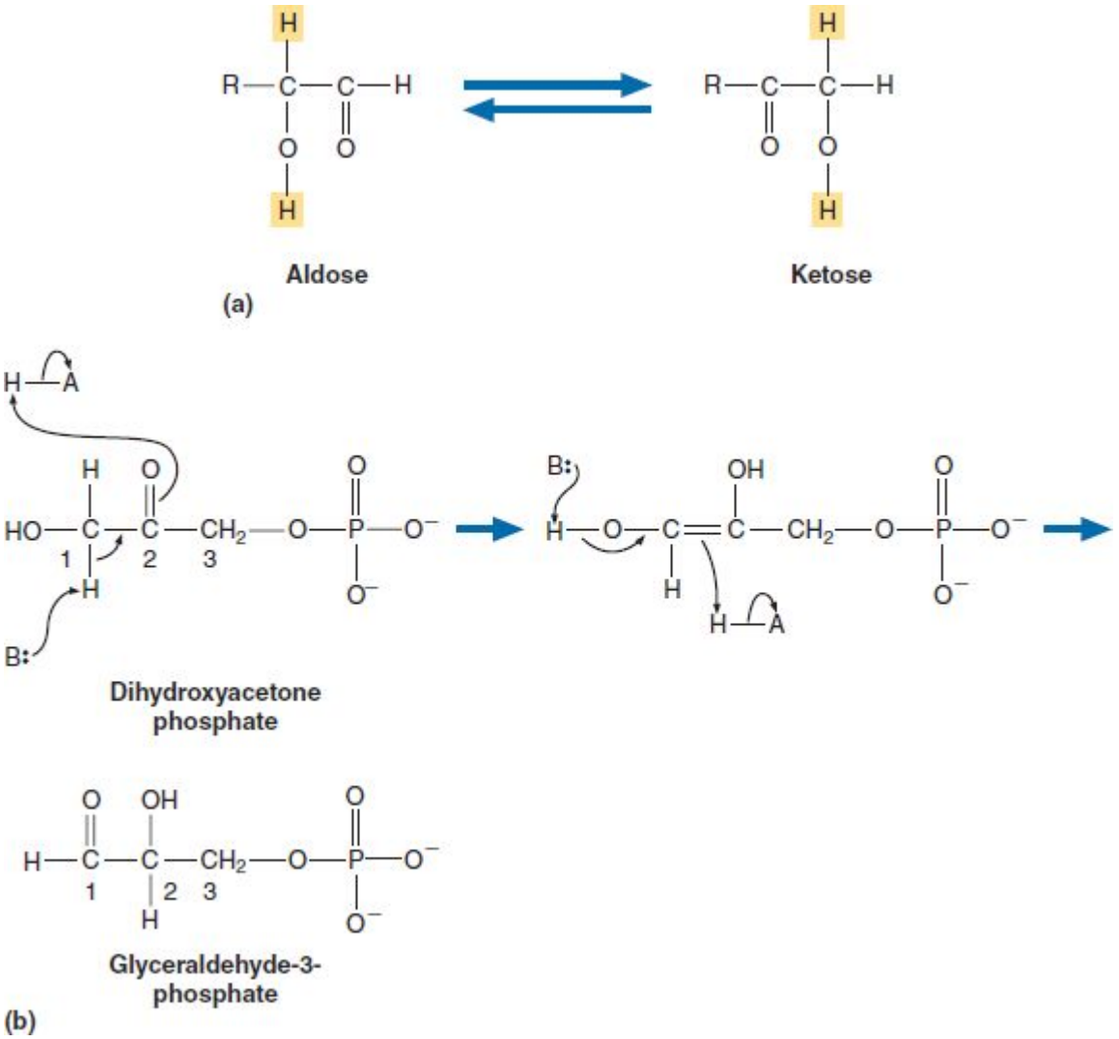


FIGURE 1.19

An Isomerization Reaction

(a) The reversible interconversion of aldose and ketose isomers is a commonly observed biochemical reaction type. (b) The isomerization of dihydroxacetone phosphate to form glyceraldehyde-3-phosphate begins when a basic side chain of the enzyme, triose phosphate isomerase, removes a proton from carbon 1 and an acidic side chain donates a proton to the carbonyl oxygen. The intermediate product is an *enediol* (a molecule in which a hydroxyl group is attached to each of the carbon atoms in a carbon–carbon double bond). In the second step the enediol is deprotonated by a basic side chain and an acidic side chain adds a proton to carbon 2, yielding the product glyceraldehyde-3-phosphate.

OXIDATION-REDUCTION REACTIONS Oxidation-reduction (redox) reactions occur when there is a transfer of electrons from a donor (called the **reducing agent**) to an electron acceptor (called the **oxidizing agent**). When reducing agents donate their electrons, they become **oxidized**. As oxidizing agents accept electrons, they become **reduced**. The two processes always occur simultaneously.

It is not always easy to determine whether biomolecules have gained or lost electrons. However, two simple rules may be used to ascertain whether a carbon atom in a molecule has been oxidized or reduced:

1. Oxidation has occurred if a carbon atom gains oxygen or loses hydrogen:

2. Reduction has occurred if a carbon atom loses oxygen or gains hydrogen:

KEY CONCEPTS

The most common reaction types encountered in biochemical processes are nucleophilic substitution, elimination, addition, isomerization, and oxidation-reduction.

Numerous biological processes involve redox reactions. Energy transformations, for example, involve electron transfers. In photosynthesis, light energy capture drives CO_2 reduction and H_2O oxidation to yield sugar molecules and O_2 synthesis, respectively. In the reverse process called cell respiration, sugar molecules are oxidized to form CO_2 and O_2 is reduced to form H_2O. In the intermediate steps in these redox processes, high-energy electrons are transferred to electron acceptors such as the nucleotide NAD^+/NADH (nicotinamide adenine dinucleotide in its oxidized/reduced form).

Energy

Energy is defined as the capacity to do work, that is, to move matter. In contrast to human-made machines, which generate and use energy under harsh conditions such as high temperature, high pressure, and electrical currents, the relatively fragile molecular machines within living organisms must use more subtle mechanisms. Cells generate most of their energy by using redox reactions in which electrons are transferred from an oxidizable molecule to an electron-deficient molecule. In these reactions, electrons are removed or added as hydrogen atoms (H•) or hydride ions ($H^{:-}$, i.e., an anion with one proton and two electrons). The more reduced a molecule is—that is, the more hydrogen atoms it possesses—the more energy it contains. For example, fatty acids contain proportionately more hydrogen atoms than sugars do and therefore yield more energy upon oxidation. When fatty acids and sugars are oxidized, their hydrogen atoms are removed by the redox coenzymes FAD (flavin adenine dinucleotide) or NAD^+. (Coenzymes are small molecules that function in association with enzymes by serving as carriers of small molecular groups, or in this case electrons.) The reduced products of this process ($FADH_2$ or NADH, respectively) can then transfer the electrons to another electron acceptor.

Whenever an electron is transferred, energy is lost. Cells have complex mechanisms for exploiting this phenomenon in a way that permits some of the released energy to be captured for cellular work. The most prominent feature of energy generation in most cells is the electron transport pathway, a series of linked membrane-embedded electron carrier molecules. During a regulated process, energy is released as electrons are transferred from one electron carrier molecule to another. During several of these redox reactions, the energy released is sufficient to

drive the synthesis of ATP, the energy carrier molecule that directly supplies the energy used to maintain highly organized cellular structures and functions.

Despite their many similarities, groups of living organisms differ in the precise strategies they use to acquire energy from their environment. **Autotrophs** are organisms that transform the energy of the sun (**photosynthesis**) or various chemicals (**chemosynthesis**) into chemical bond energy; they are called, respectively, **photoautotrophs** and **chemoautotrophs**. The **heterotrophs** obtain energy by degrading preformed food molecules obtained by consuming other organisms. **Chemoheterotrophs** use preformed food molecules as their sole source of energy. Some prokaryotes and a small number of plants (e.g., the pitcher plant, which digests captured insects) are **photoheterotrophs**; that is, they use both light and organic biomolecules as energy sources.

KEY CONCEPTS

In living organisms, energy, the capacity to move matter, is usually generated by redox reactions.

The ultimate source of the energy used by most life-forms on Earth is the sun. Photosynthetic organisms such as plants, certain prokaryotes, and algae capture light energy and use it to transform carbon dioxide (CO_2) into sugar and other biomolecules. Chemotrophic species derive the energy required to incorporate CO_2 into organic biomolecules by oxidizing inorganic substances such as hydrogen sulfide (H_2S), nitrite (NO_2^-), or hydrogen gas (H_2). The biomass produced in both types of process is, in turn, consumed by heterotrophic organisms that use it as a source of energy and structural materials. At each step, as molecular bonds are rearranged, some energy is captured and used to maintain the organism's complex structures and activities. Eventually, energy becomes disorganized and is released in the form of heat. The metabolic pathways by which energy is generated and used by living organisms are briefly outlined next, followed by the basic mechanisms by which cellular order is maintained.

Overview of Metabolism

Metabolism is the sum of all the enzyme-catalyzed reactions in a living organism. These reactions are organized into pathways (**Figure 1.20**) in which an initial reactant molecule is modified in a step-by-step sequence into a product that can be used by the cell for a specific purpose. For example, glycolysis, the energy-generating pathway that degrades the six-carbon sugar glucose, is composed of 10 reactions. All of an individual organism's metabolic processes consist of a vast web-like pattern of interconnected biochemical reactions that are regulated such that resources are conserved and energy use is optimized. There are three classes of biochemical pathways: metabolic, energy transfer, and signal transduction.

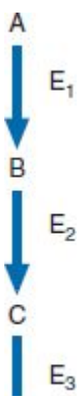


FIGURE 1.20

A Biochemical Pathway

In this three-step biochemical pathway, biomolecule A is converted into biomolecule D in three sequential reactions. Each reaction is catalyzed by a specific enzyme (E).

METABOLIC PATHWAYS There are two types of metabolic pathway: anabolic and catabolic. In **anabolic** (biosynthetic) **pathways**, larger molecules are synthesized from smaller precursors. Building-block molecules (e.g., amino acids, sugars, and fatty acids), either produced or acquired from the diet, are incorporated into larger, more complex molecules. Anabolic processes include the synthesis of polysaccharides and proteins from sugars and amino acids, respectively. Because biosynthesis increases order and complexity, anabolic pathways require an input of energy. During **catabolic pathways** large complex molecules are degraded into smaller, simpler products. Some catabolic pathways release energy. A fraction of this energy is captured and used to drive anabolic reactions.

The relationship between anabolic and catabolic processes is illustrated in **Figure 1.21**. As nutrient molecules are degraded, energy and reducing power (high-energy electrons) are conserved in ATP and NADH molecules, respectively. Biosynthetic processes use metabolites of catabolism, synthesized ATP and NADPH (reduced nicotinamide adenine dinucleotide phosphate, a source of reducing power), to create complex structure and function.

ENERGY TRANSFER PATHWAYS Energy transfer pathways capture energy and transform it into forms that organisms can use to drive biomolecular processes. The absorption of light energy by chlorophyll molecules and the energy-releasing redox reactions required for its conversion to chemical bond energy in a sugar molecule is a prominent example.

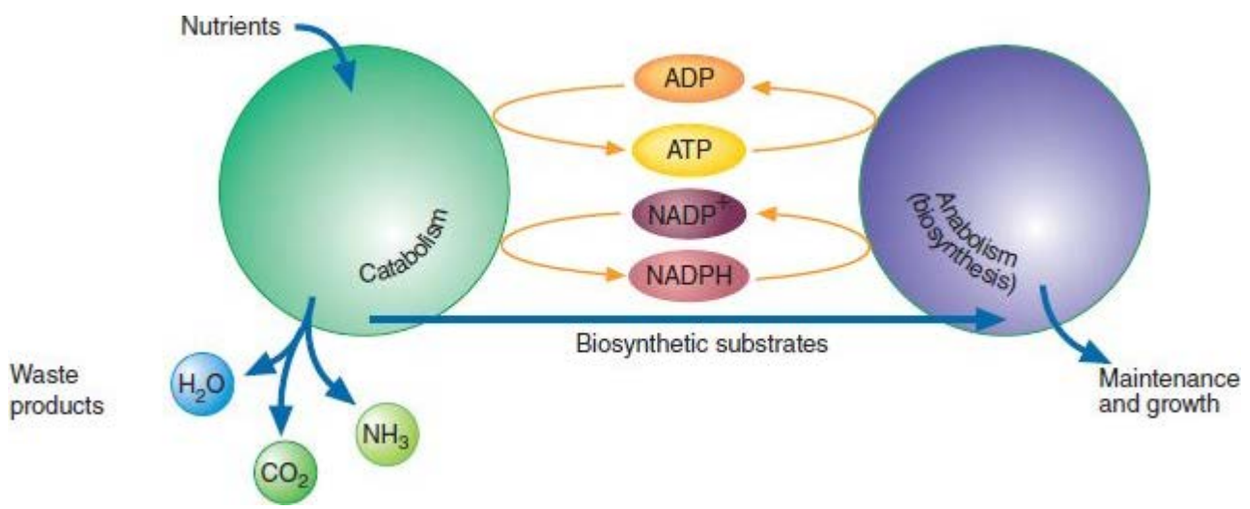


FIGURE 1.21

Anabolism and Catabolism

In organisms that use oxygen to generate energy, catabolic pathways convert nutrients to small-molecule starting materials. The energy (ATP) and reducing power (NADPH) that drive biosynthetic reactions are generated during catabolic processes as certain nutrient molecules are converted to waste products such as carbon dioxide, ammonia, and water.

SIGNAL TRANSDUCTION **Signal transduction** pathways allow cells to receive and respond to signals from their surroundings. In the initial or *reception phase*, a signal molecule such as a hormone or a nutrient molecule binds to a receptor protein. This binding event initiates the *transduction phase*, a cascade of intracellular reactions that triggers the cell's response to the original signal. For example, glucose binds to its receptor on pancreatic insulin-secreting cells, whereupon insulin is released into the blood. Most commonly, such responses are an increase or a decrease in the activity of already existing enzymes or the synthesis of new enzyme molecules.

KEY CONCEPTS

- Metabolism is the sum of all the enzyme-catalyzed reactions in a living organism.
- There are three classes of biochemical pathway: metabolic (anabolic and catabolic), energy transfer, and signal transduction.

Biological Order

The coherent unity that is observed in all living organisms involves the functional integration of millions of molecules. In other words, life is highly organized complexity. Despite the rich diversity of living processes that contribute to generating and maintaining biological order, most can be classified into the following categories: (1) synthesis and degradation of biomolecules, (2) transport of ions and molecules across cell membranes, (3) production of force and movement, and (4) removal of metabolic waste products and other toxic substances.

SYNTHESIS OF BIOMOLECULES Cellular components are synthesized in a vast array of chemical reactions, many of which require energy, supplied directly or indirectly by ATP molecules. The molecules formed in biosynthetic reactions perform several functions. They can be assembled into supramolecular structures (e.g., the proteins and lipids that constitute membranes) or serve as informational molecules (e.g., DNA and RNA) or catalyze chemical reactions (i.e., the enzymes).

TRANSPORT ACROSS MEMBRANES Cell membranes regulate the passage of ions and molecules from one compartment to another. For example, the plasma membrane (the animal cell's outer membrane) is a selective barrier. It is responsible for the transport of certain substances such as nutrients from a relatively disorganized environment into the more orderly cellular interior. Similarly, ions and molecules are transported into and out of organelles (p. 59) during biochemical processes. For example, fatty acids are transported into organelles known as mitochondria so that they may be broken down to generate energy.

CELL MOVEMENT Organized movement is one of the most obvious characteristics of living organisms. The intricate and coordinated activities required to sustain life require the movement of cell components. Examples in eukaryotic cells include cell division and organelle movement, two processes that depend to a large extent on the structure and function of a complex network of protein filaments known as the *cytoskeleton*. The forms of cellular motion profoundly influence the ability of all organisms to grow, reproduce, and compete for limited resources. As examples, consider the movement of protist cells as they search for food in a pond or the migration of human white blood cells as they search for infectious foreign cells throughout the body. More subtle examples include the movement of specific enzymes along a DNA molecule during the chromosome replication that precedes cell division and the secretion of insulin by certain

pancreatic cells.

WASTE REMOVAL All living cells produce waste products. For example, animal cells ultimately convert food molecules, such as sugars and amino acids, into CO_2, H_2O, and NH_3. These molecules, if not disposed of properly, can be toxic. Some substances are readily removed. In animals, for example, CO_2 diffuses out of cells and (after a brief and reversible conversion to bicarbonate by red blood cells) is quickly exhaled through the respiratory system. Excess H_2O is excreted through the kidneys. Other molecules, however, are so toxic that specific processes have evolved to provide for their disposal. The urea cycle (described in Chapter 15) provides a mechanism for converting free ammonia and excess amino nitrogen into urea, a less toxic molecule. The urea molecule is then removed from the body through the kidney as a major component of the urine.

Living cells also contain a wide variety of potentially toxic molecules that must be disposed of. Plant cells solve this problem by transporting such molecules into a vacuole, where they are either broken down or stored. Animals, however, must use disposal mechanisms that depend on water solubility (e.g., the formation of urine by the kidney). Hydrophobic substances such as steroid hormones, which cannot be broken down into simpler molecules, are converted during a series of reactions into water-soluble derivatives. This mechanism is also used to solubilize some exogenous organic molecules such as drugs and environmental contaminants.

KEY CONCEPTS

In living organisms, processes of highly ordered complexity are sustained by a constant input of energy.

1.4 SYSTEMS BIOLOGY

Information in the overview of biochemical processes that you have just read about was discovered using a method of inquiry based on *reductionism*. In this powerful, mechanistic strategy a complex, living “whole” is studied by “reducing” it to its component parts. Each individual part is then further broken down so that the chemical and physical properties of its molecules and the connections between them can be determined. Most of the accomplishments of the modern life sciences would have been impossible without the reductionist philosophy. However, reductionism has its limitations because of the assumption that detailed knowledge of all the properties of the parts will of itself ultimately provide a complete understanding of the functioning of the whole. Despite intense efforts, a coherent understanding of dynamic living processes continues to elude investigators.

In recent decades, a new approach called systems biology has been utilized to achieve a deeper understanding of living organisms. Based on the engineering principles originally developed to build jet aircraft, **systems biology** regards living organisms as integrated systems. Each system allows certain functions to be performed. One such system in animals is the digestive system, which comprises a group of organs that is tasked to break down food into molecules that can be absorbed by the body’s cells.

Although human-engineered systems and living systems are remarkably similar in some respects, they are significantly different in others. The most important difference is the design issue. When engineers plan a complex mechanical or electrical system, each component is designed to fulfill a precise function, and there are no unnecessary or unforeseen interactions between network components. For example, the individual electrical wires in the cables that

control aircraft are insulated to prevent damage caused by short circuits. In contrast, biological systems have evolved by trial and error over several billion years. Evolution, the adaptation of populations of living organisms in response to selection pressure, is made possible by the capacity to generate genetic diversity through various forms of mutation, gene duplications, or the acquisition of new genes from other organisms. The components of living organisms, unlike engineer-designed parts, have no fixed functions, and overlapping functions are permissible. Living systems have become increasingly more complex, in part because of the unavoidability of interactions among established system components and potentially useful new parts (e.g., derived from gene duplications followed by mutations).

The systems approach is especially useful because the human mind cannot analyze the hundreds of biochemical reactions that are taking place at once in a living organism. To tackle this problem, systems biologists have invented mathematical and computer models to derive from biochemical reaction pathways an understanding of how these processes operate over time and under varying conditions. The success of these models is reliant on huge data sets containing accurate information about cellular concentrations of biomolecules and the rates of biochemical reactions as they occur in living, functioning cells. Although these data sets are incomplete, this analytic method has produced some notable successes. The technology required to identify and quantify biomolecules of all types continues to be refined. System biologists have identified two core principles that underpin the complex and diverse biochemical pathways described in this textbook: emergence and robustness. In addition, systems biologists organize the vast complexities of living cells with concepts such as systems, networks, modules, and motifs, which are also briefly described.

Emergence

As we have discovered, the behavior of complex systems cannot be understood simply by knowing the properties of constituent parts. At each level of organization of the system, new and unanticipated properties emerge from interactions among parts. In other words, emergent properties of a complex system have different characteristics than those of its component parts. For example, hemoglobin (the protein that transports oxygen in the blood to the body's cells) requires ferrous iron (Fe^{2+}) to be functional. Whereas iron easily oxidizes in the inanimate world, the iron in hemoglobin does not usually oxidize, even though it is linked directly to oxygen during the transport process. The amino acid residues that line the iron-binding site protect Fe^{2+} from oxidation. The protection of ferrous iron in hemoglobin is an **emergent property**, that is, a property conferred by the complexity and dynamics of the system and not anticipated by knowledge of the chemical properties of hemoglobin's amino acids.

Robustness

Systems that remain stable despite diverse perturbations such as fluctuations or damage are described as *robust*. Autopilot systems in aircraft, for example, maintain a designated flight path despite expected fluctuations in conditions such as wind speed or the plane's mechanical functions. All robust systems are necessarily complex because failure prevention requires an integrated set of automatic fail-safe mechanisms. The robust (fail-safe) properties of human-made mechanical systems are created by *redundancy*, the use of duplicate parts (e.g., backup electric generators in an airplane). Although the design of living organisms does include some redundant parts, the robust properties of living systems are largely the result of **degeneracy**, the capacity of structurally different parts to perform the same or similar functions. The genetic code is a simple,

well-recognized example. Of the 64 possible three-base sequences (called codons) on an mRNA molecule, 61 base triplets code for 20 amino acids during protein synthesis. Since most amino acids have more than one codon, degeneracy of the code provides a measure of protection against base substitution mutations. Similarly, the inactivation of a specific type of hormone receptor may be compensated for by similar receptors with partial overlapping functions.

Systems Biology Model Concepts

The research efforts of systems biologists have resulted in the development of simplifying models, which facilitate the efforts of life science researchers, as well as students, to understand the vast complexities of living organisms. Terminology used in systems biology includes system, network, module, and motif.

SYSTEM A *system* is defined as an interconnected and interacting assembly of biomolecules. Systems under investigation can be organisms, organs, cells, or organelles. For example, the mitochondrion is an organelle (a type of cell compartment in the cells of organisms such as animals and plants). It possesses structural features and biochemical pathways that convert the energy in food molecules into the chemical energy required to drive cell processes and synthesize numerous biomolecules, among other functions.

NETWORK Systems can be thought of as the dynamic interaction of networks, each of which is a group of interconnected molecules that performs one or more functions. Living organisms possess metabolic, signaling, and regulatory networks. A *metabolic network* consists of interconnected biochemical reaction pathways that synthesize and degrade biomolecules. Reactant and product molecules connect these pathways to each other. For example, glycolysis, the pathway that degrades the sugar glucose, is linked to energy capture pathways within mitochondria by pyruvate, the product of glycolysis. Pyruvate is transported into the mitochondrion, where biochemical reactions in another pathway begin the process of capturing the energy in its hydrogen atoms. Glycolysis is also linked to amino acid biosynthetic pathways because certain glycolytic intermediates serve as precursor molecules.

Living organisms must perceive and correctly respond to both their internal and external environments. Cells acquire and process information through vast, intricate *signaling networks* composed of receptor proteins that receive information and signaling pathways, whose components process it. For example, the binding of epinephrine (p. 319) to its receptor on the surface of liver cells initiates a signaling mechanism that results in the activation of enzymes that degrade glycogen.

Living organisms have elaborate, robust mechanisms that tightly control metabolic pathways. This control is accomplished by *regulatory networks* that switch on and off the genes that code for the synthesis of enzymes and all other biomolecules. For example, the binding of the hormone insulin to its receptor on the surface of its target cells sets in motion a signaling mechanism that alters the expression of numerous genes (e.g., enzymes in glycogen and triacylglycerol synthesis). However, gene regulatory networks in living organisms are inextricably integrated with other networks. For example, insulin receptor binding also triggers a signaling pathway that quickly modifies the activity of several biochemical pathways by stimulating the activity of certain enzymes while inhibiting others.

MODULE Complex systems are composed of *modules*, components or subsystems that perform specific functions. Living organisms utilize modules because they are easily assembled,

rearranged, and repaired, as well as eliminated when necessary. Although modules (e.g., enzymes extracted from cells in the lab) can often be isolated with some or even most of their functional properties, their function is meaningful only within the context of the larger system. In living organisms, modularity occurs at all system levels. Examples within a cell include amino acids, proteins, and biochemical pathways. Modularity is especially important because it provides the capacity to limit damage to components that can be easily removed and replaced. For example, glycolysis can be considered a module. Functional relationships between modules in a system are managed by *protocols*, or sets of rules that specify how and whether modules will interact. The mechanism that facilitates pyruvate transport into a mitochondrion is an example of a protocol.

MOTIF Network motifs are recurring regulatory circuits that have many different uses. In living organisms, the most common type is **feedback control** (**Figure 1.22**), a self-regulating mechanism in which the product of a process acts to modify the process, either negatively or positively. In *negative feedback*, the most common form, an accumulating product slows its own production. Many biochemical pathways are regulated by negative feedback. Typically, a pathway product inhibits an enzyme near the beginning of the pathway. In *positive feedback* control, a product increases its own production. Positive feedback control is found less often in living organisms because the mechanism is potentially destabilizing. If not carefully controlled, the amplifying effect of a positive feedback loop can result in the collapse of the system. In blood clotting, for example, the platelet plug that seals a damaged blood vessel does not expand continuously because inhibitors are released by nearby undamaged blood vessel cells.

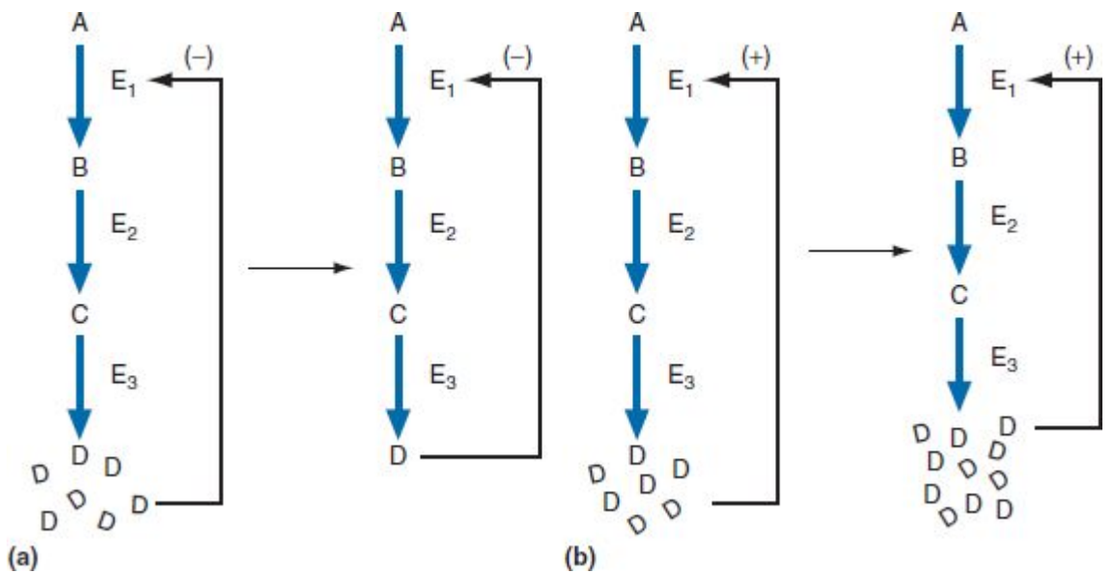


FIGURE 1.22

Feedback Mechanisms

(a) **Negative Feedback**. As a product molecule accumulates, it binds to and inhibits the activity of an enzyme in the pathway. The result is the decreased production of the product. (b) **Positive Feedback**. As product molecules accumulate, they stimulate an enzyme in the pathway, thereby causing an increased rate of product synthesis.

Living Organisms and Robustness

Fail-safe control mechanisms, whether in human-made systems or in living organisms, are expensive. Constraints such as energy availability require priorities in resource allocation. As a result, systems are generally protected from commonly encountered environmental changes, but

they are vulnerable to unusual or rare events that cause damage. This vulnerability, referred to as *fragility*, is an inescapable feature of robust systems. Cancer, a group of diseases in which cell cycle control is disrupted, is an example of the "robust, yet fragile" nature of robust systems. Despite the elaborate controls on cell division in animal bodies, mutations in just a few of the genes that code for cell cycle regulatory proteins can result in a robust uncontrolled proliferation of the affected cell.

KEY CONCEPTS

- Systems biology is an attempt to reveal the functional properties of living organisms by developing mathematical models of interactions from available data sets.
- The systems approach has provided insights into the emergent properties, robustness, and modularity of living organisms.

Biochemistry IN THE LAB

An Introduction

Biochemical technologies exploit the chemical and physical properties of biomolecules: chemical reactivity, size, solubility, net electrical charge, movement in an electric field, and absorption of electromagnetic radiation. As life science research has become increasingly more sophisticated, scientists have provided a progressively more coherent view of the living state. The Human Genome Project was a landmark event in this process. The goal of this international research effort, begun in the late 1980s, was to determine the nucleotide base sequence of human DNA. The subsequent development of automated DNA sequencing technology revolutionized life science research because it provided scientists with a "high-throughput" (i.e., rapid, high-volume, relatively inexpensive) means of investigating the information content of genomes, a field now referred to as **genomics**.

Genomics has been especially useful in medical research. A large number of human diseases have been linked to errors in one or more gene sequences or to faulty regulation of gene expression. Among the early benefits of this work are rapid, accurate tests for predisposition to pathological conditions such as cystic fibrosis, breast cancer, and some liver diseases. Several recently developed technologies have created additional opportunities to investigate the molecular basis of disease. For example, DNA microchips (thousands of DNA molecules arrayed on a solid surface) are now routinely used to monitor gene expression of cells. Proteins can also be rapidly analyzed by mass spectrometry in combination with new technologies. Among the new fields created by high-throughput methods are **functional genomics** (the investigation of gene expression patterns) and **proteomics** (the investigation of protein synthesis patterns and protein–protein interactions). The science of **bioinformatics** is the computer-based field that facilitates analysis of the massive amounts of protein and nucleic acid sequence data that are being generated. System biologists take advantage of all of these methods to decipher biological networks such as metabolic process control, gene regulation, and signal transduction (information-processing) mechanisms.

In the past, biochemists and other scientists have often benefited from each other's work. For

example, technologies created by physicists such as X-ray diffraction, electron microscopy, and radioisotope labeling made biomolecular structure and function investigations possible. In recent years, the life sciences have also benefited from the services provided by computer scientists, mathematicians, chemists, and engineers. As the biological knowledge base has continued to expand, it has become increasingly obvious that future advances in life science and medical research will require the efforts of multidisciplinary teams of scientists.

Chapter Summary

1. Biochemistry may be defined as the study of the molecular basis of life. Biochemists have contributed to the following insights into life: (1) life is complex and dynamic, (2) life is organized and self-sustaining, (3) life is cellular, (4) life is information-based, and (5) life adapts and evolves.
2. Animal and plant cells contain thousands of different types of molecules. Water constitutes 50 to 90% of a cell's content by weight, and ions such as Na^+, K^+, and Ca^{2+} may account for another 1%. Almost all the other kinds of biomolecules are organic. Many biomolecules are proteins, which play a variety of roles in living organisms: transport proteins, structural proteins, and catalytic proteins (enzymes).
3. Many of the biomolecules found in cells are relatively small, with molecular weights of less than 1000 daltons. Cells contain four families of small molecules: amino acids, sugars, fatty acids, and nucleotides.
4. DNA, consisting of two antiparallel polynucleotide strands, is the repository of genetic information in living organisms. DNA contains coding sequences, referred to as genes, and noncoding sequences, some of which have regulatory functions. RNA is a single-stranded polynucleotide that differs from DNA in that it contains the sugar ribose instead of deoxyribose and the base uracil instead of thymine. RNAs have numerous functions. Examples include protein synthesis and transcription regulation. Gene expression, the process that controls if or when a gene will be transcribed, involves the binding of transcription factors to specific regulatory DNA sequences called response elements.
5. All life processes consist of chemical reactions catalyzed by enzymes. Among the most common reaction types encountered in biochemical processes are nucleophilic substitution, elimination, addition, isomerization, and oxidation-reduction. Biochemical reactions are organized into pathways where a reactant is converted to a product in a step-by-step sequence where each reaction is catalyzed by a separate enzyme.
6. Living organisms require a constant flow of energy to prevent disorganization. The principal means by which cells obtain energy is oxidation of biomolecules or certain minerals.
7. Metabolism is the sum of all the reactions in a living organism. There are two types of metabolic pathway: anabolic and catabolic. Energy transfer pathways capture energy and transform it into forms that organisms can use to drive biomolecular processes. Signal transduction pathways, which allow cells to receive and respond to signals from their environment, consist of three phases: reception, transduction, and response.
8. The complex structure of cells requires a high degree of internal order. This is accomplished by four primary means: synthesis of biomolecules, transport of ions and molecules across cell membranes, production of movement, and removal of metabolic waste products and other toxic substances.
9. Systems biology is a new field that attempts to provide understanding of the functional properties of living organisms by applying mathematical modeling strategies to amassed biological data. Among the early benefits of the systems approach are the insights associated with emergence, robustness, and modularity.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee**, where you can complete a multiple-choice quiz on this introductory chapter to help you prepare for exams.

Suggested Readings

Cai L, Fisher AL, Huang H, Xie Z. 2016. CRISPR-mediated genome editing and human diseases. *Genes and Diseases* 3:244–51.

Collins FS, Varmus HA. 2015. New initiative on precision medicine. *New Eng J Med* 372(9):793–5.

Goodsell DS. 2009. The machinery of life, 2nd ed. New York (NY): Springer.

Rothman S. 2002. Lessons from the living cell: the limits of reductionism. New York (NY): McGraw-Hill.

Tudge C. 2000. The variety of life: a survey and a celebration of all the creatures that have ever lived. New York (NY): Oxford.

Key Words

addition reaction, *17*
amino acid, *7*
anabolic pathway, *20*
anhydride, *14*
autopoiesis, *13*
autotroph, *19*
bioinformatics, *26*
biomolecule, *3*
catabolic pathway, *20*
chemoautotroph, *19*
chemoheterotroph, *19*
chemosynthesis, *19*
degeneracy, *23*
electrophile, *14*
elimination reaction, *15*
emergent property, *23*
energy, *19*
enzyme, *3*
fatty acid, *10*
feedback control, *24*
functional genomics, *26*
functional group, *5*
gene, *12*
gene expression, *13*
genome, *12*
genomics, *26*

heterotroph, *19*
homeostasis, *3*
hydration reaction, *17*
hydrocarbon, *5*
hydrolysis, *14*
hydrophilic, *7*
hydrophobic, *5*
isomerization, *17*
leaving group, *14*
lipid, *10*
macromolecule, *3*
metabolism, *3*
mixed anhydride, 14
modules, *24*
monosaccharide, *8*
mutation, *5*
negative feedback, *25*
neurotransmitter, *7*
noncoding RNA, *13*
nucleic acid, *11*
nucleophile, *14*
nucleophilic substitution, *14*
nucleotide, *10*
oxidation-reduction (redox) reaction, *18*
oxidize, *18*
oxidizing agent, *18*
peptide, *7*
peptide bond, *7*
photoautotroph, *19*
photoheterotroph, *19*
photosynthesis, *19*
polypeptide, *7*
polysaccharide, *8*
positive feedback, *25*
protein, *7*
proteomics, *26*
purine, *11*
pyrimidine, *11*
reduce, *18*
reducing agent, *18*
reductionism, *22*
response element, *13*

robust, *23*
saturated, *10*
signal transduction, *21*
sugar, *8*
systems biology, *22*
transcription, *12*
transcription factor, *13*
unsaturated, *10*

Review Questions

SECTION 1.1

Comprehension Questions

1. Define the following terms:
    a. biomolecules
    b. macromolecules
    c. metabolism
    d. homeostasis
    e. enzymes
2. Define the following terms:
    a. deoxyribonucleic acid
    b. genetic information
    c. insulin receptor
    d. mutation
    e. hierarchically organized system

Fill in the Blanks

3. The sum of all the reactions in a living organism is called its _____.
4. _____ are the linear sequences of nucleotides in an organism's genetic information.
5. In multicellular organisms, the levels of organization are _____, _____, _____ and _____.
6. _____ is a series of reactions in which a specific biomolecule is converted into a product molecule.

Short-Answer Questions

7. Distinguish between silent and harmful mutations.
8. Describe why insulin is considered a signal molecule.
9. List three life science fields that require a solid understanding of biochemical principles.

Critical-Thinking Questions

10. Describe in general terms how mutations are involved in the evolution of species.
11. Describe the properties that all cells have in common.

SECTION 1.2

Comprehension Questions

12. Define the following terms:
    a. hydrocarbon
    b. hydrophilic
    c. hydrophobic
    d. functional group
    e. R group
13. Define the following terms:
    a. carbonyl group
    b. carboxyl group
    c. amino group
    d. hydroxyl group
    e. peptide bond
14. Define the following terms:
    a. polypeptide
    b. peptide
    c. protein
    d. standard amino acids
    e. neurotransmitter
15. Define the following terms:
    a. sugar
    b. glucose
    c. monosaccharide
    d. polysaccharide
    e. cellulose
16. Define the following terms:
    a. triacylglycerol
    b. phosphoglyceride
    c. fatty acid
    d. saturated fatty acid
    e. unsaturated fatty acid
17. Define the following terms:
    a. nucleic acid
    b. nucleotide
    c. purine
    d. pyrimidine
    e. deoxyribose

18. Define the following terms:

    a. DNA
    b. RNA
    c. genome
    d. transcription
    e. fructose

19. Define the following terms:

    a. rRNA
    b. tRNA
    c. mRNA
    d. siRNA
    e. miRNA

20. Define the following terms:

    a. ribosome
    b. transcription factor
    c. signal molecule
    d. response element
    e. RNA interference

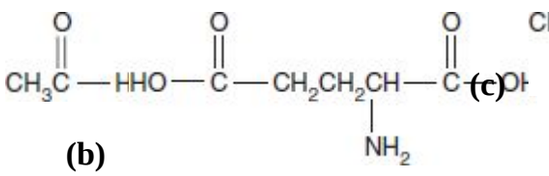
Fill in the Blanks

21. The biologically active form of a fatty acid is the thioester of _____.
22. The genome's functional gene products are either _____ or _____.
23. _____ is an antiviral defense mechanism involving siRNA molecules.

Short-Answer Questions

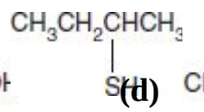
24. Identify the functional groups in the following molecules:

(a)

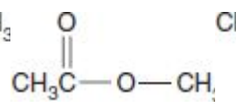


(b)

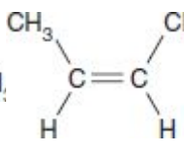
(c)



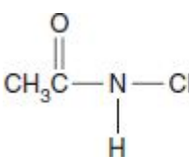
(d)



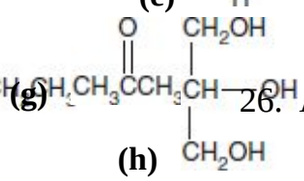
(e)



(f)



(g)

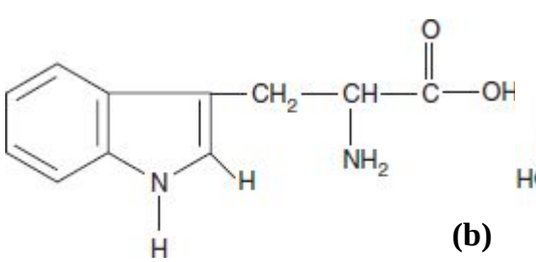


(h)

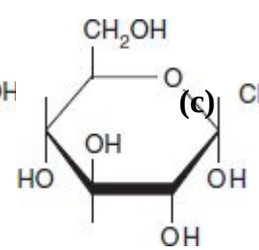
25. What are the six major elements present in living organisms?

26. Assign each of the following molecules to one of the major classes of biomolecule:

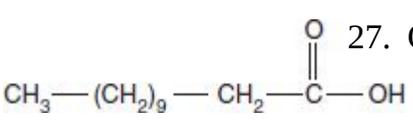
(a)



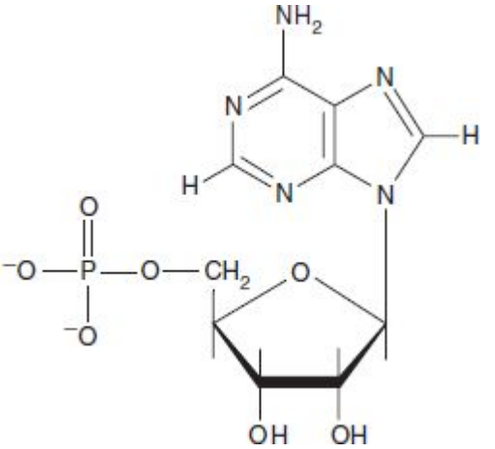
(b)



(c)



27. Carbohydrates are widely



(d)

(e) Name four classes of small biomolecules. In what larger molecules are they found?

recognized as sources of metabolic energy. What are the two other critical roles that carbohydrates play in living organisms?

28. Nucleotides have roles in addition to being components of DNA and RNA. Give an example.
29. List two functions of each of the following biomolecules:

    a. fatty acids
    b. sugars
    c. nucleotides
    d. amino acids

Critical-Thinking Questions

31. Elements such as carbon, hydrogen, and oxygen that occur in biomolecules form stable covalent bonds. What would be the result if the bonds between these atoms were either slightly less or more stable than the naturally occurring bonds?
32. Why are fatty acids the principal long-term energy reserve of the body?
33. Hundreds of thousands of proteins have been discovered in living organisms. Yet, as astonishing as this diversity is, these molecules constitute only a small fraction of the possible protein molecules. Calculate the total number of possible decapeptides (molecules with 10 amino acid residues linked by peptide bonds) that can be synthesized from the 20 standard amino acids. If you were to spend 5 minutes writing out the molecular structure of each decapeptide, how long would the task take?
34. When a substance such as sodium chloride is dissolved in water, the ions become completely surrounded by water molecules, forming structures called hydration spheres. When the sodium salt of a fatty acid is mixed with water, the carboxylate group of the molecule becomes hydrated, but the hydrophobic portion of the molecule is poorly hydrated, if at all. Using a circle to represent the carboxylate group and an attached squiggly line to represent the hydrocarbon chain of a fatty acid, draw a picture of how fatty acids interact with water.

SECTION 1.3

Comprehension Questions

35. Define the following terms:

    a. hydrolysis

b. nucleophilic substitution
c. elimination reaction
d. hydration reaction
e. isomerization reaction

36. Define the following terms:

a. nucleophile
b. electrophile
c. leaving group
d. addition reaction
e. anhydride

37. Define the following terms:

a. adenosine triphosphate
b. redox reaction
c. oxidizing agent
d. reducing agent
e. NADH

38. Define the following terms:

a. hydride ion
b. oxidation reaction
c. energy
d. FAD
e. electron transport pathway

39. Define the following terms:

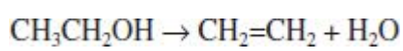
a. coenzyme
b. anabolic pathway
c. catabolic pathway
d. signal transduction pathway
e. glycolysis

40. Define the following terms:

a. autotroph
b. chemoautotroph
c. photoautotroph
d. photoheterotroph
e. chemoheterotroph

Fill in the Blanks

41. The following is an example of a _____ reaction.



$$CH_3CH_2OH \rightarrow CH_2{=}CH_2 + H_2O$$

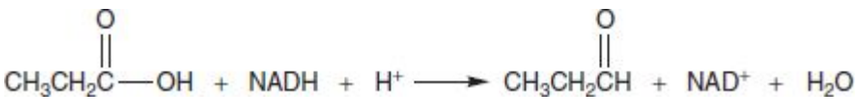
42. Each organism is considered an autonomous self-organizing self-maintaining entity. This set of properties is referred to as _____.
43. _____ are organisms that transform the energy of the sun into chemical bond

energy.

44. The sum of all reactions in a living organism is called its _____.

Short-Answer Questions

45. Distinguish between living organisms and human-made factories.
46. Compare anabolic and catabolic pathways. What molecules link these two processes?
47. How do cells obtain energy from chemical bonds?
48. List several important ions that are found in living organisms.
49. List three types of biochemical reactions involving acyl nucleophilic substitution.
50. Name three waste products that animal cells produce. How do plants dispose of waste products?
51. Identify the oxidizing and reducing agents in the following reaction:



52. What reaction is the first step in utilizing glucose as an energy source?
53. What are the primary functions of metabolism?

Critical-Thinking Questions

54. The order of reactivity in nucleophilic substitution reactions is as follows:

phosphate > thiols > esters > amides.

Explain this order on the basis of their pK_a values: phosphoric acid (1×10^{-3}), hydrogen sulfide (1×10^{-7}), alcohols (1×10^{-16}), and ammonia (1×10^{-36}). (pK_a is the negative log of the acid dissociation constant, which is a quantitative measure of the strength of an acid in a solution, i.e., the tendency of an acid to lose a proton.)

55. Carboxylic acids that undergo nucleophilic acyl substitution reactions are often first converted to thioesters, For example, acetic acid forms a thioester with a molecule called coenzyme A, which has a sulfhydryl group.

What is the leaving group in these reactions?

SECTION 1.4

Comprehension Questions

56. Define the following terms:

a. systems biology
b. emergence

c. robustness
d. degeneracy
e. system

57. Define the following terms:

a. network
b. metabolic network
c. signaling network
d. module
e. motif

58. Define the following terms:

a. emergent property
b. negative feedback control
c. positive feedback control
d. bioinformatics
e. proteomics

Fill in the Blanks

59. A property conferred by the complexity and dynamics of a system is called a(n) _____ property.
60. A _____ is an interconnected and interacting assembly of biomolecules.
61. _____ are components or subsystems that perform specific functions in complex systems.
62. _____ are recurring regulatory circuits that have many different uses in complex systems.

Short-Answer Questions

63. Why is cancer, which occurs as the result of disrupted cell cycle control, an example of the fragility of a robust system?
64. Compare and contrast the feature of an airplane autopilot system with a biological system.
65. Provide several examples of emergent properties.

Critical-Thinking Questions

66. How does the statement “the whole is more than the sum of its parts” apply to living organisms? Give an example.
67. Tay–Sachs disease is a devastating genetic neurological disorder caused by the lack of an enzyme that degrades a specific lipid molecule. When this molecule accumulates in brain cells, an otherwise healthy child undergoes motor and mental deterioration within months after birth and dies by the age of 3 years. In general terms, describe how a system biologist evaluates this phenomenon?
68. Humans synthesize most of the cholesterol required for cell membranes and for the synthesis of vitamin D and other steroid hormones. Keeping in mind that living organisms are self-regulating, what would you expect to happen if a person’s diet is high in cholesterol? Provide a reason for your response.

69. Unlike human-engineered system, the components of biological systems often have multiple functions. Is this phenomenon a strength or a weakness of biological systems?
70. The cancerous cells in a tumor proliferate uncontrollably, and treatment often involves the use of toxic drugs in attempts to kill them. Often, however, after initial success (i.e., shrinkage of the tumor), the cancer returns because resistance to the drugs develops. Biochemists have identified one of the major causes of this phenomenon called multidrug resistance. One or more cells in the tumor have expressed the gene for P-glycoprotein, a membrane transport protein that pumps the drugs out of the cells. In the absence of the toxic drug molecules, these cells again grow uncontrollably and eventually become the dominant cells in the tumor. What features of living organisms does this process illustrate?

MCAT Study Questions

71. What are the possible products of a hydrolysis reaction of an amide?

    a. carboxylic acid and alcohol
    b. amino acid and carboxylic acid
    c. ether and aldehyde
    d. ketone and amine

72. The major roles of RNA include:

    i gene expression; ii. protein synthesis; iii. carbohydrate synthesis

    a. i
    b. ii
    c. i and ii
    d. ii and iii

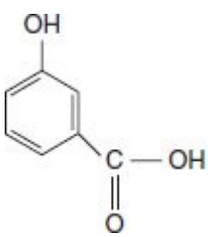
73. When an carboxylic acid reacts with an alcohol, the product is called a(n)

    a. amide
    b. diester
    c. ester
    d. fatty acid

74. Which of the following molecules is the most polar?

    a. propane
    b. propanol
    c. acetic acid
    d. formaldehyde

75. The following compound has which functional groups?

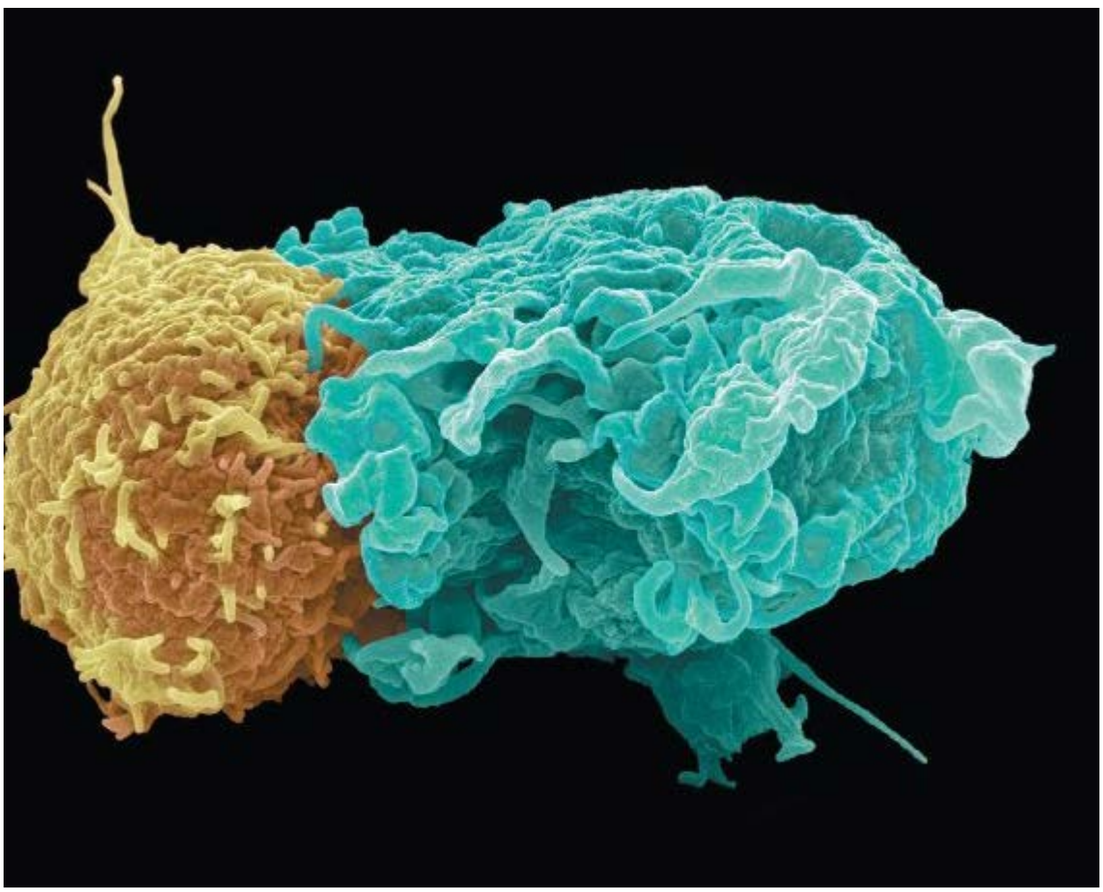


    i = alcohol ii = ester iii = aromatic iv = carboxylic acid

    a. i, ii, and iv

b. i, iii, and iv
c. i and ii
d. iii and iv

CHAPTER 2 Living Cells



The Immune System at Work: Antigen Presentation In this colored scanning electron micrograph, two white blood cells are performing a vital process that can result in the destruction of foreign cells or host-infected cells. After a macrophage (blue) has engulfed a foreign cell such as a bacterium, it proceeds to initiate a multifaceted process that will eliminate the threat. In the first step, the macrophage uses enzymes and toxic chemicals to destroy the bacterium. As it does so, the macrophage inserts bacterial protein fragments (antigens) into its own plasma membrane in a phenomenon referred to as antigen presentation. T cell activation occurs when a T helper lymphocyte (yellow) with surface proteins capable of binding to a specific foreign antigen interacts with the macrophage. The subsequent activation of the T cell leads to its proliferation, followed by activation of other immune system cells. The end result is the destruction of the invading bacteria.

OUTLINE

OUR BODIES, OUR SELVES

2.1 CORE BIOCHEMISTRY CONCEPTS

Biochemistry and Water

Biological Membranes
Self-Assembly
Molecular Machines
Macromolecular Crowding
Proteostasis
Signal Transduction

2.2 STRUCTURE OF PROKARYOTIC CELLS

Cell Wall
Plasma Membrane
Cytoplasm
Pili and Flagella

2.3 STRUCTURE OF EUKARYOTIC CELLS

Plasma Membrane
Endoplasmic Reticulum
Golgi Apparatus
Vesicular Organelles and Lysosomes: The Endocytic Pathway
Nucleus
Mitochondria
Peroxisomes
Chloroplasts
Cytoskeleton

Biochemistry in Perspective

Primary Cilia and Human Disease

Biochemistry in the Lab

Cell Technology

AVAILABLE ONLINE

Biochemistry in Perspective
Organelles and Human Disease
Biochemistry in Perspective
Caveolar Endocytosis

Our Bodies, Our Selves

It would surprise most humans that we are colonized by a vast number and diversity of microorganisms. Current estimates of microorganisms to human cell ratios range from 3:1 to 1:1. Most of these organisms, referred to as an indigenous flora or **microbiota** (**Figure 2.1**), are bacteria with smaller numbers of *archaeans* (another type of prokaryote), fungi, and viruses. Humans and their microbiota have evolved together into a dynamic, interdependent superorganism. This relationship is usually symbiotic (mutually beneficial in some way) or commensal (nonharmful). However, a few species in the normal human microbiota are pathogens that can cause disease if conditions permit (e.g., if the immune system is depressed).

Our bodies begin acquiring microbes as soon as the amniotic sac ruptures. As babies proceed down the

birth canal, colonization begins as they are exposed to their mother's microbiota. Within a short time, a diverse array of microbes has taken up residence in all body surfaces that are exposed to the external environment: skin and certain parts of the respiratory, gastrointestinal (GI), and urogenital tracts. These ecosystems, each with its own set of environmental conditions (e.g., temperature, pH, and O_2 availability), eventually possess their own characteristic communities of microorganisms. The majority of human microbiotic organisms occur in the intestines (between 500 and 1000 species by some estimates) where they provide a spectrum of beneficial services in exchange for a stable nutrient supply and favorable environmental conditions. For example, numerous bacteria contribute to dietary fiber digestion, a process that contributes about 5% of human energy requirements and produces a variety of bioactive molecules. Among these are the *short-chain fatty acids* (SCFAs) acetate, propionate, and butyrate, which are responsible for some of the beneficial effects of dietary fiber in human health (e.g., reduction in risk of colon cancer and metabolic disorders such as type 2 diabetes (pp. 615–17), cardiovascular disease, and obesity). Other examples of microbiotic roles include vitamin synthesis (vitamin K and several B vitamins), repression of pathogens (organisms that produce toxins or cause severe diarrhea), and robust immune system development.

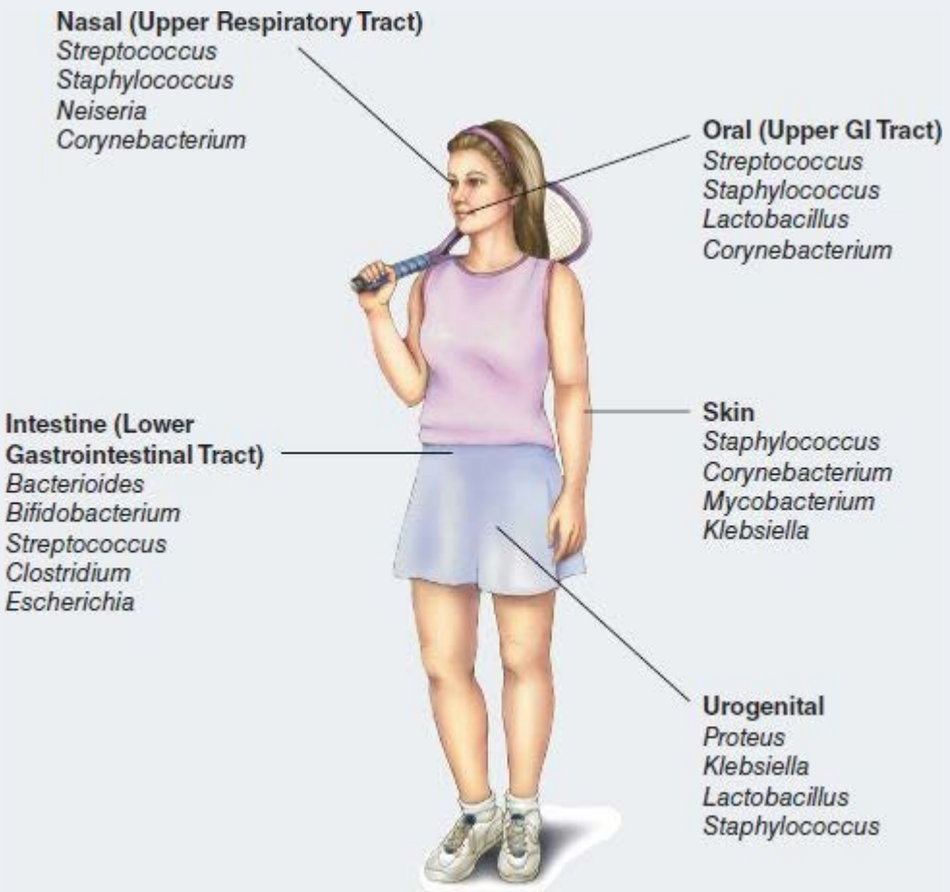


FIGURE 2.1

The Human Microbiota

Each healthy human is host to an exceptionally large and unique set of highly adapted microorganisms. Examples of the major groups of organisms observed in the normal flora are given for each ecosystem. Note that although some bacterial groups occur at different body locations, the species often differ.

Experiments with gnotobiotic (germ-free) mice have revealed that the absence of a microbiota has a profound effect on health. In addition to requiring 30% more calories to maintain body weight than conventional mice, germ-free mice were observed to have underdeveloped intestinal architecture and fat storage capacity, as well as smaller organs (heart, lungs, and liver). These mice are also highly susceptible to numerous infections caused by transient pathogens (e.g., respiratory or intestinal viruses) because their immune systems are immature.

Defense Mechanisms

Despite the many benefits of the human indigenous microbiota, the body must constantly protect itself from the microbiota's potentially unrestricted growth. Strategies used to prevent damage include impenetrable epithelial tissue barriers and immune system cells. Epithelial barriers in the GI tract lining protect internal organs from invasion of microorganisms. Epithelial cells also produce antimicrobial proteins. Examples include the α-defensins, released from cells of the small intestine that kill bacteria by inserting into membranes and forming pores that cause cell rupture.

The immune system normally strikes a fine balance of protection from pathogens and tolerance of nonpathogenic organisms and the body's own cells. Approximately 70% of immune system cells (e.g., lymphocytes and macrophages) are located in and around the GI tract (especially the small and large intestine), where maintaining this balance is vital.

Gut Flora and Human Health

Each person has a unique microbiota that is the result of genetic inheritance, birth circumstances (e.g., vaginal delivery vs. Caesarean section and the mother's microbiota), diet, and environment (e.g., exposure to antibiotics or acid-suppressing medications). Bacteria dominate the human gut microbiota. In most humans, more than 90% of the microorganisms in the intestines belong to two phyla: the Firmicutes and Bacteroidetes. A healthy gut flora is characterized as diverse, relatively stable, and resilient. The mechanism that maintains this condition, referred to as *colonization resistance*, protects the body from the challenge of exogenous and potentially pathogenic microbes. Colonization resistance is a constellation of tactics (e.g., competition for space in the mucus layer lining the intestinal lumen and intermicrobial chemical warfare).

In recent years, excessive antibiotic use, poor diets (e.g., those high in fat and sugar and low in fiber), and chronic stress have contributed to a rise in disrupted gut microbiota, characterized most notably by low species diversity. Antibiotic treatment causes major alterations of the gut microbiota. Frequent antibiotic use can result in *dysbiosis*, a condition in which a severely eroded colonization resistance allows the overgrowth of pathogens. Poor diets can also cause dysbiosis since dietary factors play a significant role in shaping the microbiota. For example, in both humans and lab animals, the switch from a balanced diet to a high-fat diet causes the loss of beneficial bacteria, some of which are known to protect cells in the gut lining. In addition to GI tract diseases such as inflammatory bowel disease (IBS, a stress-related disorder now believed to be a contributing factor in depression and anxiety), dysbiosis has also been linked to numerous systemic disease states, including obesity, type 2 diabetes (pp. 615–17), metabolic syndrome (pp. 626), and several autoimmune diseases. Low-level chronic systemic inflammation, caused by the leakage of bacterial molecules such as lipopolysaccharide (also called endotoxin) across a gut wall compromised by the loss of protective bacteria, is now believed to be a significant feature of most metabolic diseases. Recent evidence has revealed that consumption of artificial sweeteners in food and beverages has detrimental health effects because their effects on certain gut microbes result in destabilized blood glucose control in lab animals and some humans.

Overview

CELLS ARE THE STRUCTURAL UNITS OF ALL LIVING ORGANISMS. ONE REMARKABLE FEATURE OF CELLS IS THEIR DIVERSITY: THE HUMAN BODY contains about 200 types of cells. This great variation reflects the variety of functions that cells can perform. However, no matter what their shape, size, or species, cells are also amazingly similar. They are all surrounded by a membrane that separates them from their environment. They are all

composed of the same types of molecules.

The structural hierarchy of life on earth extends from the biosphere to biomolecules. Each level is inextricably linked to the levels above and below it. Cells, however, are considered the basic unit of life, since they are the smallest entities that are actually alive. Cells can sense and respond to their environment, transform matter and energy, and reproduce themselves.

Living cells are classified as either prokaryotic or eukaryotic. Prokaryotes are single-celled organisms that lack a nucleus (*pro* = "before," *karyon* = "nucleus" or "kernel"). Analysis of the RNA in prokaryotes has revealed that there are two distinct types of prokaryotes: the Bacteria and the Archaea. Some bacterial species cause disease (e.g., cholera, tuberculosis, syphilis, and tetanus), whereas others are of practical interest for humans (e.g., those used to make foods such as yogurt, cheese, and sourdough bread). A prominent feature of the Archaea is their unsurpassed capacity to occupy and even thrive in very challenging habitats. The eukaryotes (*eu* = "true") are composed of cells that possess a nucleus, a membrane-bound compartment that contains the cell's DNA. Animals, plants, fungi, and single-celled protists are examples of the eukaryotes. Eukaryotes also differ from prokaryotes in size and complexity. The volume of a typical eukaryotic cell such as a hepatocyte (liver cell) is between 6000 and 10,000 μm^3. The volume of the bacterium *Escherichia coli* is significantly smaller at 2 to 4 μm^3. Although the structural complexity of prokaryotes is significant, that of eukaryotes is greater by several orders of magnitude, largely because of subcellular compartments called **organelles**. Each organelle is specialized to perform specific tasks. The compartmentalization afforded by organelles creates microenvironments in which biochemical processes can be efficiently regulated. In multicellular eukaryotes, complexity is increased by cellular specialization and intercellular communication mechanisms.

The common features of prokaryotic and eukaryotic cells include their similar chemical composition and the universal use of DNA as genetic material. This chapter provides an overview of cell structure. This review is a valuable exercise because biochemical reactions do not occur in isolation. Our understanding of living processes is incomplete without knowledge of their cellular context. After a brief discussion of some basic themes in cellular structure and function, the essential structural features of prokaryotic and eukaryotic cells are described in relation to their biochemical roles.

2.1 CORE BIOCHEMISTRY CONCEPTS

Each living cell contains millions of densely packed biomolecules that perform at a frenetic pace the thousands of tasks that together constitute life. The application of biochemical techniques to investigations of living processes has provided significant insights into the unique chemical and structural properties of biomolecules that make their functional properties possible. Understanding biochemical processes can be significantly enhanced by examining the following core concepts: water, biological membranes, self-assembly, molecular machines, macromolecular crowding, proteostasis, signal transduction, calcium ions as a signaling device, and the relationship between signal transduction and metabolism.

KEY CONCEPTS

- The chemical and physical properties of water make it an indispensable component of living organisms.
- Hydrophilic molecules interact with water. Hydrophobic molecules do not interact with water.

Biochemistry and Water

Water dominates living processes. Water's chemical and physical properties (described in Chapter 3) that result from its unique polar structure and its high concentration make it an indispensable component of living organisms. Among water's most important properties is its capacity to interact with a wide range of substances. In fact, the behavior of all other molecules in living organisms is defined by the nature of their interactions with water. **Hydrophilic** molecules—those that possess positive or negative charges or contain relatively large numbers of electronegative oxygen or nitrogen atoms—interact easily with water. Examples of simple hydrophilic molecules include salts such as sodium chloride and sugars such as glucose. In contrast, **hydrophobic** molecules—such as the hydrocarbons, which possess few if any electronegative atoms—do not interact with water. Instead, when they are mixed with water, hydrophobic molecules spontaneously form clusters, minimizing contact between hydrocarbon chains and water molecules (**Figure 2.2**). Between the two extremes is an enormous group of both large and small biomolecules, each of which possesses its own unique pattern of hydrophilic and hydrophobic functional groups. Living organisms exploit the distinctive molecular structure of each of these biomolecules.

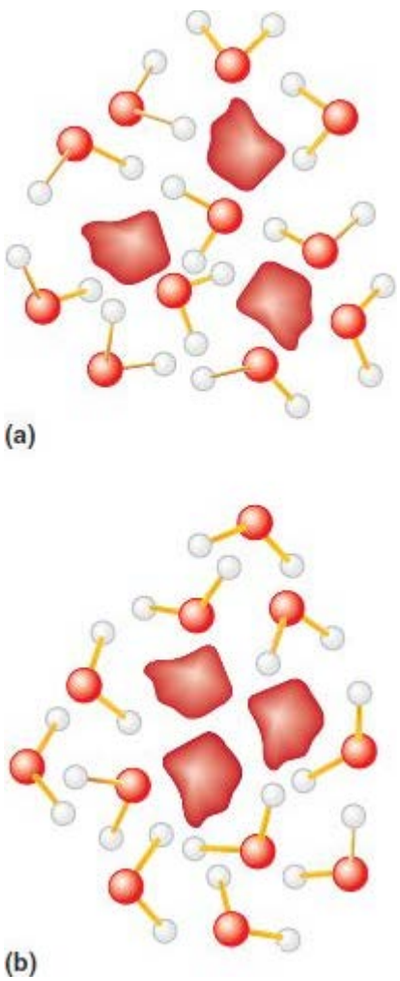


FIGURE 2.2

Hydrophobic Interactions between Water and a Nonpolar Substance

As soon as nonpolar substances (e.g., hydrocarbons) are mixed with water (a), they coalesce into droplets (b). Hydrophobic interactions between nonpolar molecules take effect only when the cohesiveness of water and other polar molecules forces nonpolar molecules or regions of molecules close together.

Biological Membranes

Biological membranes are thin, flexible, and relatively stable sheet-like structures that enclose all living cells and organelles. These membranes can be thought of as noncovalent two-dimensional supramolecular complexes (i.e., they are composed of molecules that are held together by noncovalent intermolecular forces; see pp. 80–82). Membranes provide chemically reactive surfaces and exhibit unique transport functions between the extracellular and intracellular compartments. They are also versatile and dynamic cellular components that are intricately integrated into all living processes. The most basic of the many crucial functions that membranes serve is to act as a selective physical barrier. Membranes prevent the indiscriminate leakage of molecules and ions out of cells or organelles into their surroundings and allow the timely intake of nutrients and export of waste products. In addition, membranes have significant roles in information processing and energy generation.

Most biological membranes have the same basic structure: a lipid bilayer composed of phospholipids and other lipid molecules, into which various proteins are embedded or attached indirectly (**Figure 2.3**). Phospholipids have two features that make them ideally suited to their structural role: a hydrophilic charged or uncharged polar group (referred to as a "head group") and a hydrophobic group composed of two fatty acid chains (often called hydrocarbon "tails").

There are two classes of membrane proteins: integral and peripheral. **Integral proteins** are embedded within the membrane because the amino acid residues in the membrane-spanning portions of these proteins have hydrophobic side chains. Because of the fluid nature of cellular membrane, integral proteins can diffuse laterally. **Peripheral proteins** are not embedded within the membrane. Rather, they are attached to it either by a covalent bond to a lipid molecule or by noncovalent interaction with a membrane protein or lipid. Membrane proteins perform a variety of functions. **Channel** and **carrier proteins** transport specific ions and molecules, respectively. **Receptors** are proteins with binding sites for extracellular ligands (signal molecules). The binding of a ligand to its cognate receptor triggers a cellular response.

Self-Assembly

Many of the working parts of living organisms are supramolecular structures. Prominent examples include *ribosomes* (the protein-synthesizing devices that are formed from several types of protein and RNA), large protein complexes such as the sarcomeres in muscle cells, and *proteosomes* (large protein complexes that degrade proteins). According to the principle of self-assembly, most molecules that interact to form stable and functional supramolecular complexes are able to do so spontaneously because they inherently possess the steric information required. They have or are predisposed to have intricately shaped surfaces with complementary structures, charge distributions, and/or hydrophobic regions that allow numerous, relatively weak noncovalent interactions (**Figure 2.4**). Self-assembly of such structures involves a balance between the tendency of hydrophilic groups to interact with water and the tendency for water to be excluded by hydrophobic groups. In some cases, self-assembly processes need assistance. For example, the folding of some proteins requires the aid of molecular chaperones, protein molecules that, among other functions, prevent inappropriate interactions during the folding process. The assembly of certain supramolecular structures (e.g., chromosomes and membranes) requires preexisting

information; that is, a new structure must be created on a template of an existing structure.

KEY CONCEPTS

- Each biological membrane is composed of a lipid bilayer into which proteins are inserted or attached indirectly.
- Biological membranes are inextricably integrated into all living processes.

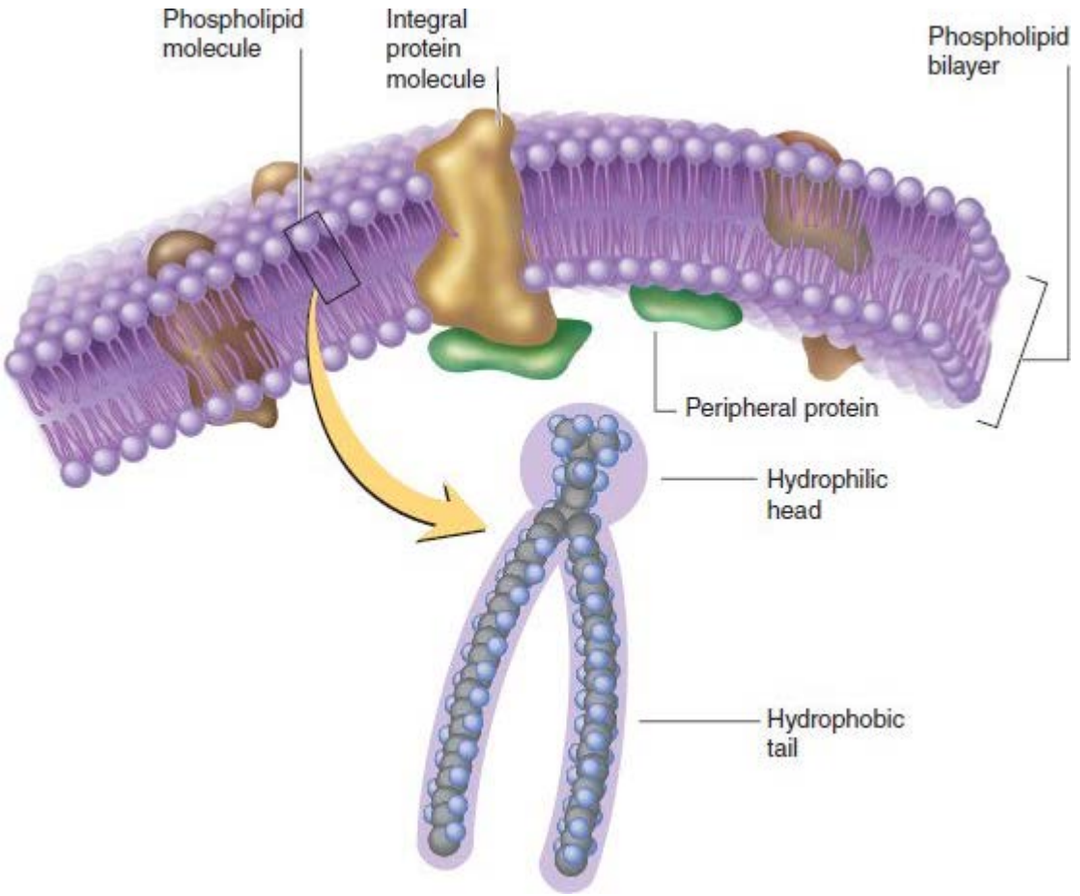


FIGURE 2.3

Membrane Structure

Biological membranes are bilayers composed largely of phospholipid molecules in which numerous proteins are suspended. Some proteins extend completely across the membrane. A space-filling model of a phospholipid is also shown.

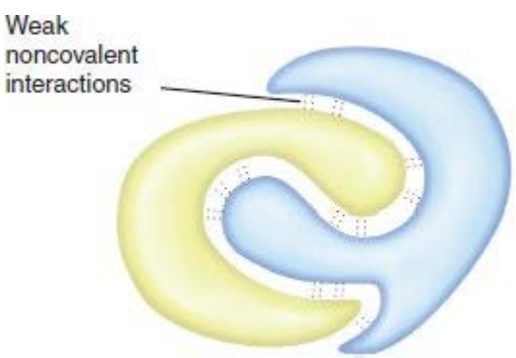


FIGURE 2.4

Self-Assembly

The information that permits the self-assembly of biomolecules consists of the complementary shapes and distributions of charges and hydrophobic groups in the interacting molecules. Large numbers of weak interactions are required for supramolecular structures to form. In this diagrammatic illustration, several weak noncovalent interactions stabilize the binding of two molecules that possess complementary shapes.

Molecular Machines

Researchers now recognize that many of the multisubunit complexes involved in cellular processes function as molecular machines: physical entities with moving parts that perform work, the product of force and distance. Like the mechanical devices used by humans, molecular machines ensure that precisely the correct amount of applied force results in the appropriate amount and direction of movement required for a specific task to be completed. Machines permit the accomplishment of tasks that often would be impossible without them.

KEY CONCEPTS

- In living organisms, the molecules in supramolecular structures assemble spontaneously.
- Biomolecules are able to self-assemble because of the steric information they contain.

Although biological machines are composed of relatively fragile molecules (primarily proteins) that cannot withstand the physical conditions (e.g., heat and friction) associated with human-made machines, the two device classes do share important features. In addition to being composed of moving parts, both require energy-transducing mechanisms; that is, they both convert energy into directed motion. Despite the wide diversity of types of work performed by biological machines, they all share one key feature: energy-driven changes in the three-dimensional shapes of proteins. One or more components of biological machines bind nucleotide molecules such as ATP or GTP (guanosine triphosphate). The binding of nucleotide molecules to these protein subunits, referred to as **motor proteins**, and the release of energy that occurs when the nucleotide is hydrolyzed, result in a precisely targeted change in the subunit's shape (**Figure 2.5**). This wave of change is transmitted to nearby subunits in a process that resembles a series of dominoes falling. Biological machines are relatively efficient because the hydrolysis of nucleotides is essentially irreversible; therefore, the functional changes that occur in each machine occur in one direction only.

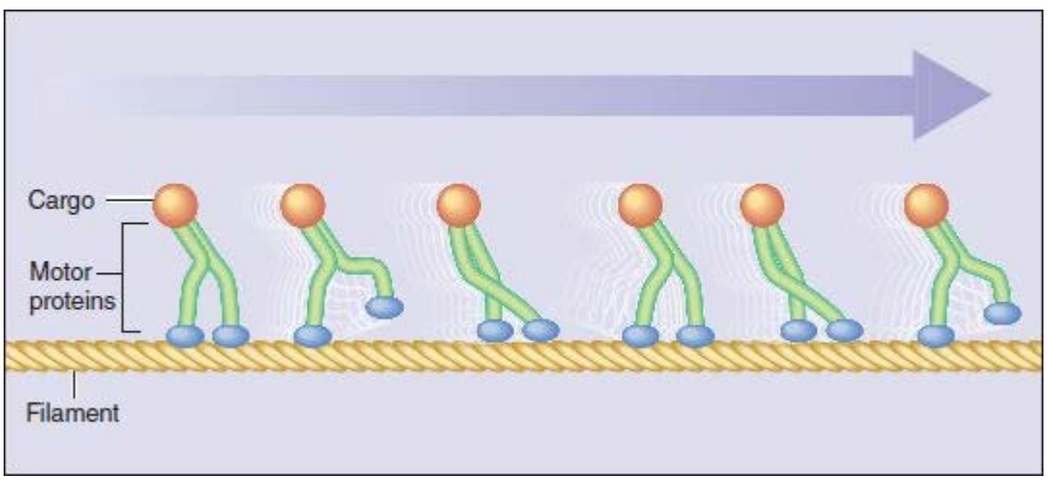


FIGURE 2.5

Biological Machines

Proteins perform work when motor protein subunits bind and hydrolyze nucleotides such as ATP. The energy-induced change in the shape of a motor protein subunit causes an orderly change in the shapes of adjacent subunits. In this diagrammatic illustration, a motor protein complex moves attached cargo (e.g., a vesicle) as it "walks" along a cytoskeletal filament.

KEY CONCEPTS

Many molecular complexes in living organisms function as molecular machines; that is, they are mechanical devices with moving parts that perform work.

Macromolecular Crowding

The interior space within cells is dense and crowded. The concentrations of proteins, the dominant type of cellular macromolecule, may be as high as 200 to 400 mg/mL. The term *crowded* rather than *concentrated* is used because macromolecules of each type usually are present in low numbers. Estimates of the volume occupied by macromolecules, called the *excluded volume*, in individual cell types vary between 20 and 40%. As illustrated in **Figure 2.6**, nonspecific steric repulsion prevents the introduction of additional macromolecules under macromolecular crowding conditions. In contrast, the remaining 70% of the space is available to small molecules. The consequences of macromolecular crowding for living systems are significant. It is now believed to be an important factor in biochemical reaction rates, protein folding, protein–protein binding, chromosome structure, gene expression, and signal transduction.

KEY CONCEPTS

Cells are densely crowded with macromolecules of diverse types. Macromolecular crowding is a significant factor in a wide variety of cellular processes.

Proteostasis

Each type of living cell has its own characteristic set of proteins, referred to as its **proteome**, which changes constantly in response to environmental conditions. Mammalian cells have an average of 10,000 types of protein, most of which are produced in multiple copies, for an estimated total of 1 billion molecules per cell. Bacteria cells such as those of *E. coli* have about 2000 different types, for a total of about 4 million molecules per cell. Following their synthesis on ribosomes, these enormous numbers of proteins must fold into their functional shapes, be transported to their proper destinations, and then be promptly degraded when they become damaged or obsolete. Adding to this complexity, cells must protect themselves from *proteotoxic stress*, a potentially lethal condition in which there is an accumulation of misfolded proteins caused by genetic variations or environmental insults such as oxidative stress (pp. 384–94), elevated temperatures, and exposure to toxins. As a result, all organisms have evolved stringent protein quality-control processes that prevent or correct protein misfolding and aggregation (the formation of usually toxic clumps of misfolded proteins) or, if necessary, destroy damaged proteins or even the cell itself.

Cells in which protein quality control is high are said to be in a state of protein homeostasis, or **proteostasis**. The processes that monitor and restore proteostasis are referred to as the **proteostasis network** (PN). The PN in mammalian cells consists of at least 2000 proteins. PN components include *molecular chaperones* (proteins that assist in protein folding or unfolding: pp. 170–72), simple proteolytic enzymes, and elaborate pathways that degrade selected proteins or

organelles. Examples of degradative processes include the *unfolded protein response* (p. 49), the *ubiquitin–proteasome system* (a mechanism in which the multiprotein proteasome complex destroys proteins that are covalently bound to ubiquitin: p. 572), lysosomal degradation (p. 52), and *autophagy* (a mechanism that destroys unnecessary or dysfunctional cell components: p. 574). Numerous signaling pathways detect unfolded proteins and the stressful conditions that threaten proteostasis. Significant research efforts have been devoted to proteostasis because PN deficiencies are an important feature of numerous human diseases. Examples include type 2 diabetes (p. 614), cardiovascular disease, lysosomal storage diseases (p. 54), and neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's.

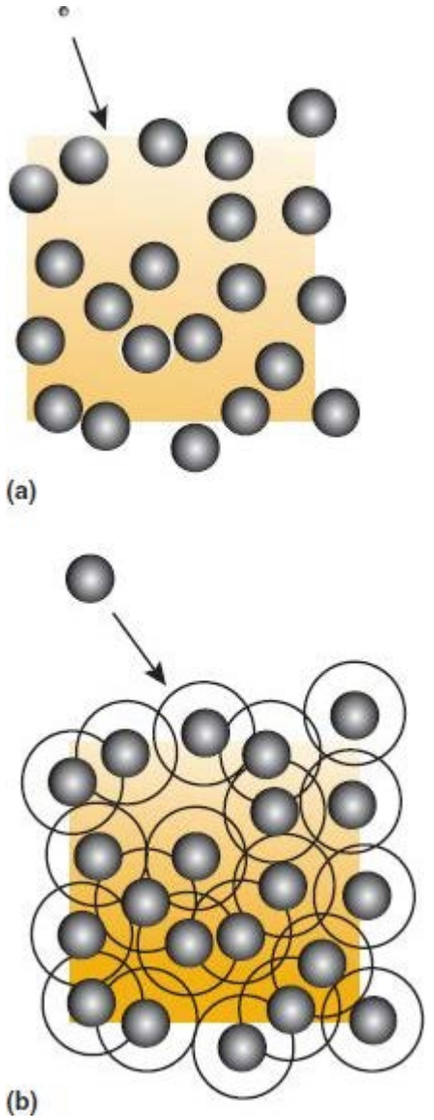


FIGURE 2.6

Volume Exclusion

Macromolecules and small molecules are depicted with large balls and small balls, respectively. Within each square, macromolecules occupy 30% of available space. (a) An introduced small molecule can penetrate into virtually all of the remaining 70% of the space. (b) Steric repulsion between macromolecules (open circles) limits the ability of these molecules to approach each other. Although the macromolecules occupy only 30% of the volume, the introduced macromolecule is excluded.

Signal Transduction

If energy is the force that drives biochemical processes, then information is the power to specify what is done. Self-organizing living organisms are so complicated that they must have not only precise structural specifications for each type of biomolecule, but also specifications for how, when, and where each type is to be synthesized, utilized, and degraded. In other words, living organisms require both energy and information to create order. Survival requires that organisms process information from their environment. For example, bacterial cells track down food molecules, plants adapt to changing light levels, and animals seek to avoid predators. Information, or *signals*, come in the form of molecules (e.g., nutrients), physical stimuli (e.g., light) and mechanical force. Although organisms are bombarded with signals, they can adapt to changing environmental conditions only if they can recognize, interpret, and respond to each type of message. The process that organisms use to receive and interpret information is referred to as **signal transduction**. Although both prokaryotes and eukaryotes process environmental information, most research efforts have been concerned with eukaryotic signal transduction. Consequently, the following discussion focuses on information processing in eukaryotes. Examples of eukaryotic signal molecules include **neurotransmitters** (products of neurons), **hormones** (products of glandular cells), and *cytokines* (products of white blood cells). All information-processing mechanisms can be divided into four phases:

1. **Reception.** A signal molecule, called a **ligand**, binds to and activates a receptor.
2. **Transduction.** Ligand binding triggers a change in the three-dimensional structure of the receptor, which results in the conversion of a primary message or signal to a secondary message, often across a membrane barrier.
3. **Response.** Once initiated, the internal signal causes a **signaling cascade**, a series of reactions that involve covalent modifications (e.g., phosphorylation) of intracellular proteins. Results of this process include changes in enzyme activities and/or gene expression, cytoskeletal rearrangements, cell movement, or cell cycle progression (e.g., cell growth or division).
4. **Termination.** The efficiency and effectiveness of signal mechanisms require that they be terminated in a timely manner. Living organisms use a variety of signal termination methods. For example, signaling molecules may be destroyed or removed (e.g., neurotransmitters such as acetylcholine and serotonin, respectively), activated proteins are inactivated by changes in covalent modification (e.g., removal of phosphate groups), and nonprotein signals are degraded by enzymes.

KEY CONCEPTS

- Living organisms receive, interpret, and respond to environmental information by means of the process of signal transduction.
- Signal transduction can be divided into four phases: reception, transduction, response, and termination.

The protein hormone insulin is a signal molecule. When released from the pancreas in response to high blood glucose levels, insulin binds to its receptor on a target cell. The insulin receptor is a member of a class of receptors called the *tyrosine kinase receptors*. Upon activation, these receptors initiate an intracellular response by catalyzing the transfer of phosphate groups to tyrosine (an amino acid residue that contains an OH group) in specific target proteins. Cellular responses triggered by insulin binding to its receptor include uptake of glucose into target cells and increased fat and glycogen synthesis.

CALCIUM IONS (Ca^{2+}): A UNIVERSAL SIGNALING DEVICE Cells respond to external stimuli by increasing their cytoplasmic Ca^{2+} concentrations, which are normally kept quite low (approximately 100 nM) (1 nM = 1×10^{-9} M) by ATP-driven pump complexes in the plasma membrane and in eukaryotes in the membrane of organelles such as the endoplasmic reticulum (p. 48). Calcium signal decoding is dependent on the amplitude and localization of brief spikes in the concentration of cytoplasmic calcium ions $[Ca^{2+}]_{cyt}$. Each type of stimulus triggers a specific signaling cascade composed of a set of Ca^{2+}–responsive proteins that change in both their shape and their functional properties when bound to the ion. Prevention of nonspecific activation of Ca^{2+}-dependent processes requires precise localization of Ca^{2+} release and then rapid clearance of the ion from the cytoplasm.

In animals, calcium ions are involved in an astonishingly diverse set of signaling processes that include neurotransmitter release from nerve cells, hormone secretion, protein folding (assisted by calcium-dependent molecular chaperone proteins), and contraction of all muscle types. In insulin secretion, for example, the release of insulin from pancreatic β-cells is triggered by calcium ions. The detection of high blood glucose levels sets in motion an intracellular signal transduction process that causes cytoplasmic Ca^{2+} levels to rise near the β-cells plasma membrane. Insulin secretion occurs because the binding of Ca^{2+} to specific calcium-sensitive membrane proteins facilitates the fusion of the membrane of insulin-containing secretory vesicles with the plasma membrane in a process known as exocytosis (p. 51).

SIGNAL TRANSDUCTION AND METABOLISM Signal transduction mechanisms in living organisms are vital. They detect relevant information in cell environments in which there is a profusion of stimuli, integrate this information, and then execute an appropriate response. Such responses involve precise alterations in gene expression and the flow of metabolites in biochemical pathways. Research efforts over several decades have revealed that such signal transduction processes are hierarchical and immensely intricate. This textbook covers the most basic signal transduction mechanisms, yet even these may be viewed as complicated. The most essential features of signal transduction processes and their effects on metabolic regulation (the effects of hormones and transcription factors on biochemical reactions) are introduced in Chapter 8 (Carbohydrate Metabolism). Later chapters devoted to metabolic networks (e.g., lipid and energy metabolism) introduce other facets of signal transduction and metabolic regulation that will allow a more integrated understanding of the biochemistry of cell and organ function. Finally, Chapter 16 provides an overview of a complex metabolic process (the human feeding–fasting cycle) and how it is regulated.

2.2 STRUCTURE OF PROKARYOTIC CELLS

The prokaryotes are an immense and heterogeneous group that are similar in external appearance: cylindrical or rod-like (bacillus), spheroidal (cocci), or helically coiled (spirilla). Prokaryotes are also characterized by their relatively small size (a typical rod-shaped bacterial cell has a diameter of 1 μm and a length of 2 μm), their capacity to move (i.e., whether they have flagella, whip-like appendages that propel them), and their retention of specific dyes. Most are identified on the basis of nutritional requirements, energy sources, chemical composition, and biochemical capacities. Despite their diversity, most prokaryotes possess the following common features: cell walls, a plasma membrane, circular DNA molecules, and no internal membrane-enclosed organelles. The anatomical features of a typical bacterial cell are illustrated in **Figure 2.7**.

Cell Wall

The prokaryotic cell wall is a complex, semirigid structure that maintains the shape of the organism and protects it from mechanical injury. The cell wall's strength is largely caused by the presence of a polymeric network made up of *peptidoglycan*, a covalent complex of short peptide chains linking long carbohydrate chains. The thickness and chemical composition of the cell wall and its adjacent structures determine how avidly a cell wall takes up and/or retains specific dyes.

The cell wall of Gram-negative bacteria—organisms that do not retain the dye crystal violet—is illustrated in **Figure 2.7**. A thin peptidoglycan layer lies between the outer membrane and the plasma membrane and within the periplasmic space. The lipid component of the *outer membrane* is lipopolysaccharide instead of phospholipids. Lipopolysaccharide, composed of a membrane-bound lipid (lipid A) attached to a polysaccharide, acts as an endotoxin. So called because they are released when the cell disintegrates, *endotoxins* are responsible for symptoms such as fever and shock in animals infected by Gram-negative bacteria. The outer membrane is relatively permeable, and small molecules move across it through *porins*, transmembrane protein complexes that contain channels. The *periplasmic space*, the region between the outer membrane and the plasma membrane, is filled with a gelatinous fluid that contains, in addition to peptidoglycan, a variety of proteins. Many of these proteins participate in nutrient digestion, transport, or *chemotaxis* (a process in which cells alter movement in response to certain chemicals such as nutrients in their environment).

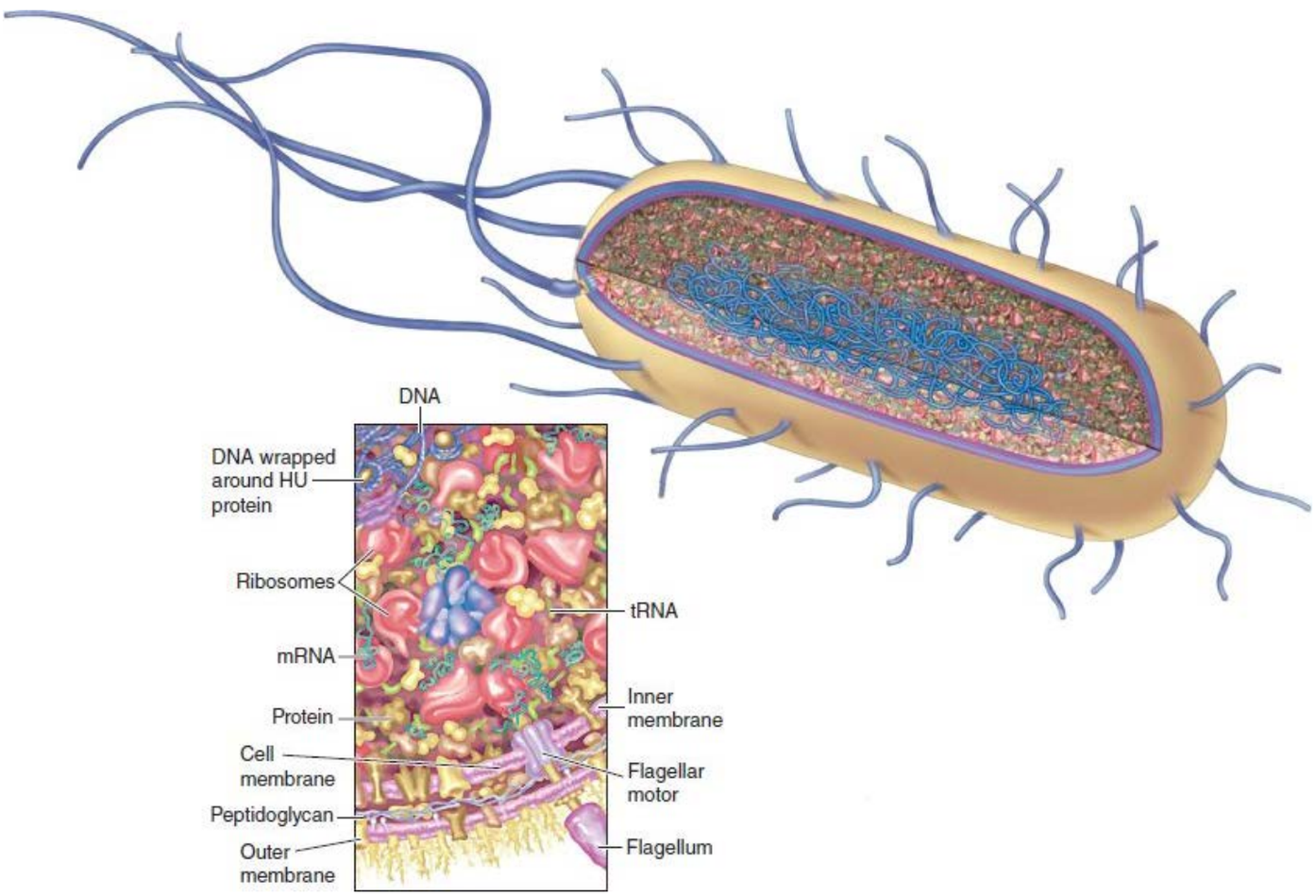


FIGURE 2.7

Structure of a Typical Bacterial Cell

All living cells contain vast numbers of densely packed and interacting molecules, each of which performs specific tasks that, taken together, are required for life. The enlargement indicates relative sizes and shapes of the major biomolecules in a bacterial cell.

Some bacteria secrete substances such as polysaccharides and proteins, collectively known as the *glycocalyx*. Depending on the structure and composition of this material, which accumulates on the outside of the cell, the bacterial glycocalyx may also be referred to as a capsule or a slime layer. Some *pathogenic* (disease-causing) bacterial species possess thick capsules that allow them to avoid detection or damage by host immune systems and to attach to host cells to facilitate colonization. Slime layers, also referred to as *biofilms*, are disorganized accumulations of polysaccharides that form when microorganisms adhere to surfaces and grow. In time, as more cells and secreted material accumulate, biofilms become thicker. Biofilms provide microorganisms with a protective barrier and are a significant feature in a variety of medical conditions (e.g., tooth decay, cystic fibrosis, and tuberculosis). The bacteria in biofilms are very resistant to immune system attack and antibiotic therapy.

Plasma Membrane

Directly inside the cell wall of bacteria is the **plasma membrane** (**Figure 2.8**). Also called the cytoplasmic membrane, the plasma membrane is a phospholipid bilayer that is reinforced with hopanoids, a group of relatively rigid molecules that resemble sterols (e.g., cholesterol) that stiffen membranes in eukaryotes. A diverse group of proteins are embedded in the lipid bilayer.

In addition to acting as a selective permeability barrier, the bacterial plasma membrane possesses receptor proteins that detect nutrients and toxins in their environment. Numerous types of transport proteins involved in nutrient uptake and waste product disposal also occur here. Depending on the species of organism, there may also be proteins involved in energy transduction processes such as **photosynthesis** (the conversion of light energy into chemical energy) and **respiration** (the oxidation of fuel molecules to generate energy).

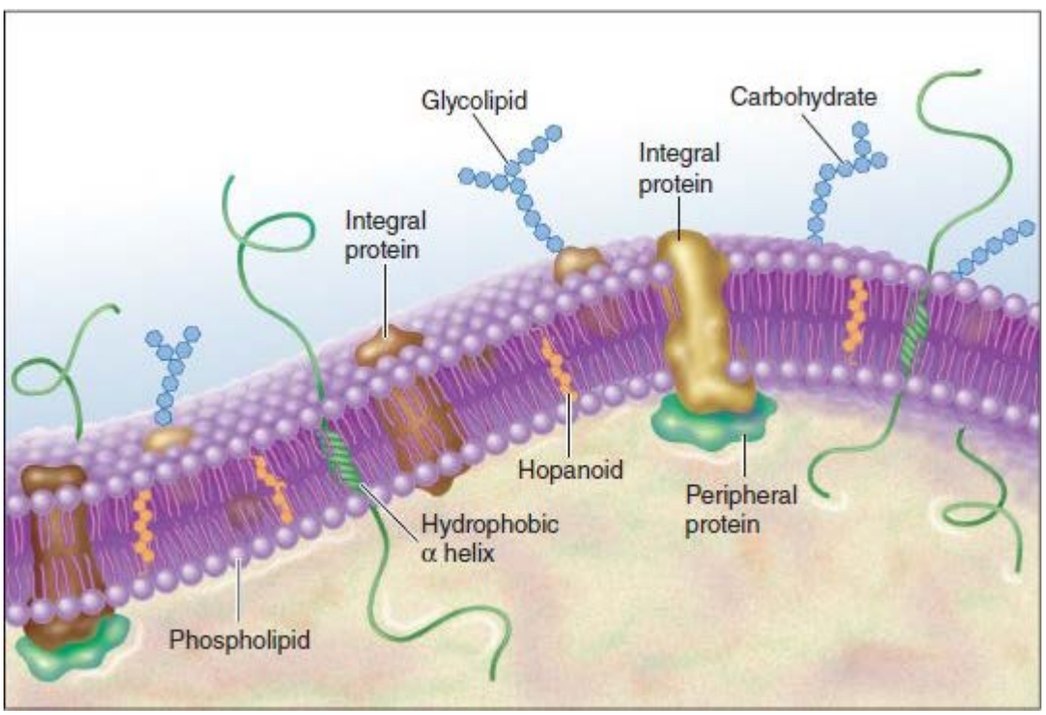


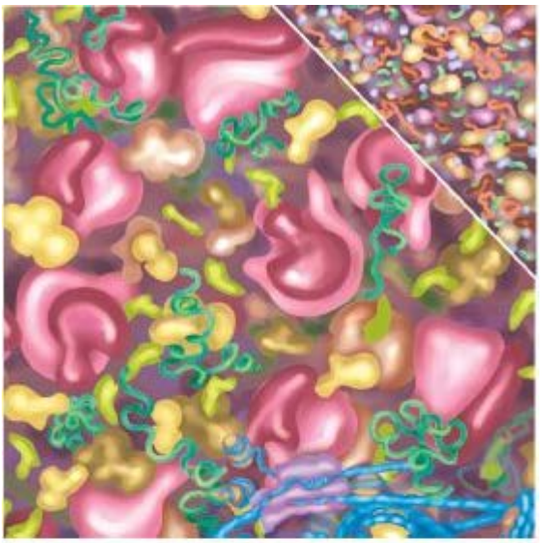
FIGURE 2.8

The Bacterial Plasma Membrane

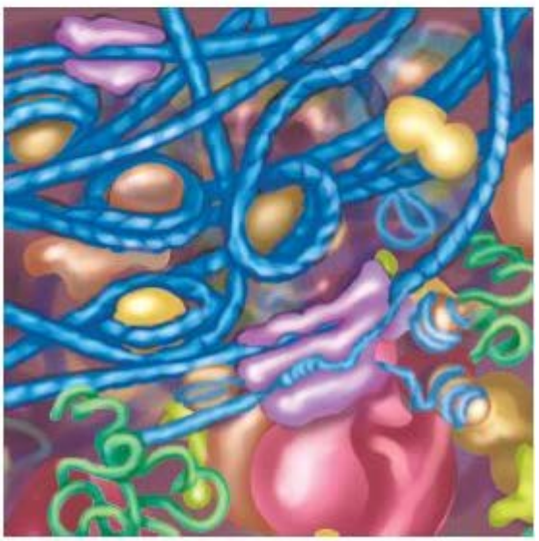
Simplified view of the plasma membrane illustrating several classes of protein and lipid. Many of these proteins and certain lipids are covalently bound to carbohydrate molecules. (Glycolipids contain carbohydrate groups.) Hopanoids are complex lipid molecules that stabilize bacterial membranes.

Cytoplasm

Despite the absence of internal membranes, prokaryotic cells do appear to have functional compartments (**Figure 2.9**). The most obvious of these is the **nucleoid**, a spacious, irregularly shaped, centrally located region that contains a long, circular DNA molecule called a **chromosome**. The bacterial chromosome typically comprises numerous regions of highly coiled and uncoiled structures. Protein complexes involved in DNA synthesis and gene expression regulation are also found within the nucleoid. Many bacteria also contain additional small circular DNA molecules called **plasmids** that can replicate independently of the chromosome. Although they are not required for growth or cell division, plasmids usually provide the cell with a biochemical advantage over cells that lack plasmids. For example, DNA sequences that code for antibiotic resistance are often found on plasmids. In the presence of the antibiotic, resistant cells synthesize a protein that inactivates the antibiotic before it can damage the cell. Such cells continue to grow and reproduce, whereas susceptible cells die.



(a)



(b)

FIGURE 2.9

Bacterial Cytoplasm

(a) Cytoplasm is a complex mixture of proteins, nucleic acids, and an enormous variety of ions and small molecules. For clarity, the small molecules appear only in the upper right corner. (b) Close-up view of the nucleoid. Note that DNA is coiled and folded around protein molecules (brown).

Under low magnification, the cytoplasm of prokaryotes has a uniform, grainy appearance except for inclusion bodies, large granules that contain organic or inorganic substances. Some species use glycogen or poly-β-hydroxybutyric acid as carbon storage polymers. Polyphosphate inclusions are a source of phosphate for nucleic acid and phospholipid synthesis. Prokaryotes that derive energy by oxidizing reduced sulfur compounds form sulfur granules. The iron mineral magnetite (Fe_3O_4) forms inclusions, called magnetosomes, which allow some species of aquatic anaerobic prokaryotes to orient themselves with the earth's magnetic field. The remaining space in the cytoplasm is filled with **ribosomes** (molecular machines composed of RNA and proteins that synthesize polypeptides) and a diverse number of macromolecules and smaller metabolites.

Pili and Flagella

Many bacterial cells have external appendages. *Pili* (singular *pilus*) are fine, hair-like structures that may allow cells to attach to food sources and host tissues. Sex pili are used by some bacteria to transfer genetic information from donor cells to recipients, a process called *conjugation*. In bacteria, the *flagellum* (plural *flagella*) is a flexible corkscrew-shaped protein filament that is used for locomotion. Cells are pushed forward when flagella rotate in a counterclockwise direction, whereas clockwise rotation results in a stop-and-tumble motion, allowing the cell to reorient for a forward run. The filament of the flagellum is anchored into the cell by a protein complex (**Figure 2.7**). Motor proteins in this complex convert chemical energy into rotational motion.

KEY CONCEPTS

- Prokaryotic cells are small and structurally simple. They are bounded by a cell wall and a plasma membrane. They lack a nucleus and other organelles.
- Their DNA molecules, which are circular, are located in an irregularly shaped region called the nucleoid.
- At low magnification, ribosomes and inclusion bodies of several types appear to be present in an otherwise featureless cytoplasm when viewed under low magnification.

QUESTION 2.1

A typical, roughly spheroidal, hepatocyte (liver cell) is a widely studied eukaryotic cell that has a diameter of about 20 μm. Calculate the volume of both a prokaryotic and a eukaryotic cell. To appreciate the magnitude of the size difference between the two cell types, estimate how many bacterial cells would fit inside the liver cell. [*Hint:* Use the expression $V = \pi r^2 h$ for the volume of a cylinder and $V = 4\pi r^3/3$ for the volume of a sphere.]

2.3 STRUCTURE OF EUKARYOTIC CELLS

The structural complexity of eukaryotic cells allows more sophisticated regulation of living processes than is possible in the prokaryotes. The most obvious features of eukaryotic cells are their large sizes (diameters of 10–100 μm) in comparison to prokaryotes. More importantly, the membrane surface area is greatly expanded by the presence of internal membrane-bound organelles. Each organelle within the cell contains a characteristic set of biomolecules and is specialized to perform specific functions. The biochemical processes within an organelle proceed efficiently because of locally high enzyme concentrations and because they can be individually regulated.

Most organelles are components of the **endomembrane system**, an extensive set of interconnecting internal membranes that divide the cell into functional compartments. The endomembrane system consists of the plasma membrane, endoplasmic reticulum, Golgi apparatus, lysosomes, and nucleus. Either through direct physical contact between compartments or by transport vesicles, the endomembrane system processes and transports a vast array of molecules through cells, as well as to and from cell exteriors. **Vesicles** are membranous sacs that bud off from a donor membrane and then fuse with the membrane of a different compartment. Once formed, each vesicle acquires a “coat” of specific proteins that facilitates its transport or targets it to its destination. Other membrane-bound organelles are mitochondria and peroxisomes and the chloroplasts in plant cells.

In addition to membranous organelles, eukaryotic cells possess several components that are devoid of membranes. Included in this group are protein-synthesizing molecular machines called ribosomes and the cytoskeleton. The cytoskeleton is a complex, dynamic, and force-generating network of filaments that give eukaryotic cells shape, structural support, and the capacity for the directed movement of molecules and organelles.

Although most eukaryotic cells possess similar structural features, there is no "typical" eukaryotic cell. Each cell type has its own characteristic structural and functional properties. They are sufficiently similar, however, that a discussion of their basic components is useful. The generalized structures of cells from animals and plants, the major forms of multicellular eukaryotic organisms, are illustrated in **Figures 2.10** and **2.11**.

Plasma Membrane

The plasma membrane isolates the cell from the outside environment. It is composed of a lipid bilayer and an enormous number and variety of integral and peripheral proteins (**Figure 2.12**). Channels and carriers within the plasma membrane regulate the passage of various ions and molecules in and out of the cell. Immense numbers of receptors play key roles in signal transduction. The extracellular face of a eukaryotic cell is heavily "decorated" with carbohydrate; that is, much of the membrane protein and lipid contains covalently attached carbohydrate. This carbohydrate "coat" is referred to as the **glycocalyx** (**Figure 2.13**). The carbohydrate molecules play important roles in cell–cell recognition (a form of intercellular communication) and adhesion, receptor specificity, and self-identity (an immune system requirement). The basic blood group antigens (A, B, AB, or O) are an example of this self-identity function.

In most eukaryotes, the plasma membrane is protected with extracellular and intracellular structures (**Figure 2.12**). Within animal tissues, the specialized cells called fibroblasts synthesize and secrete structural proteins and complex carbohydrates that form the **extracellular matrix (ECM)**, a gelatinous material that binds cells together. In addition to its support and protective functions, the ECM plays roles in the regulation of cell behavior through the binding of some of its components to specific membrane receptors in biochemical and mechanical signaling processes of various types. The inner surface of the eukaryotic plasma membrane is reinforced by a three-dimensional meshwork of proteins called the **membrane skeleton**, which is attached to the membrane by extensive noncovalent bonding to peripheral proteins. In animal cells, this protein network—composed of actin (p. 45), several types of actin-binding proteins, and spectrin (p. 426)—provides mechanical strength to the plasma membrane and determines cell shape. Direct and indirect interactions among membrane skeleton components, plasma membrane integral proteins, and lipid molecules intermittently partition the membrane into compartments (p. 427). The resulting temporary confinement of transmembrane proteins and membrane microdomains is believed to facilitate signal transduction processes.

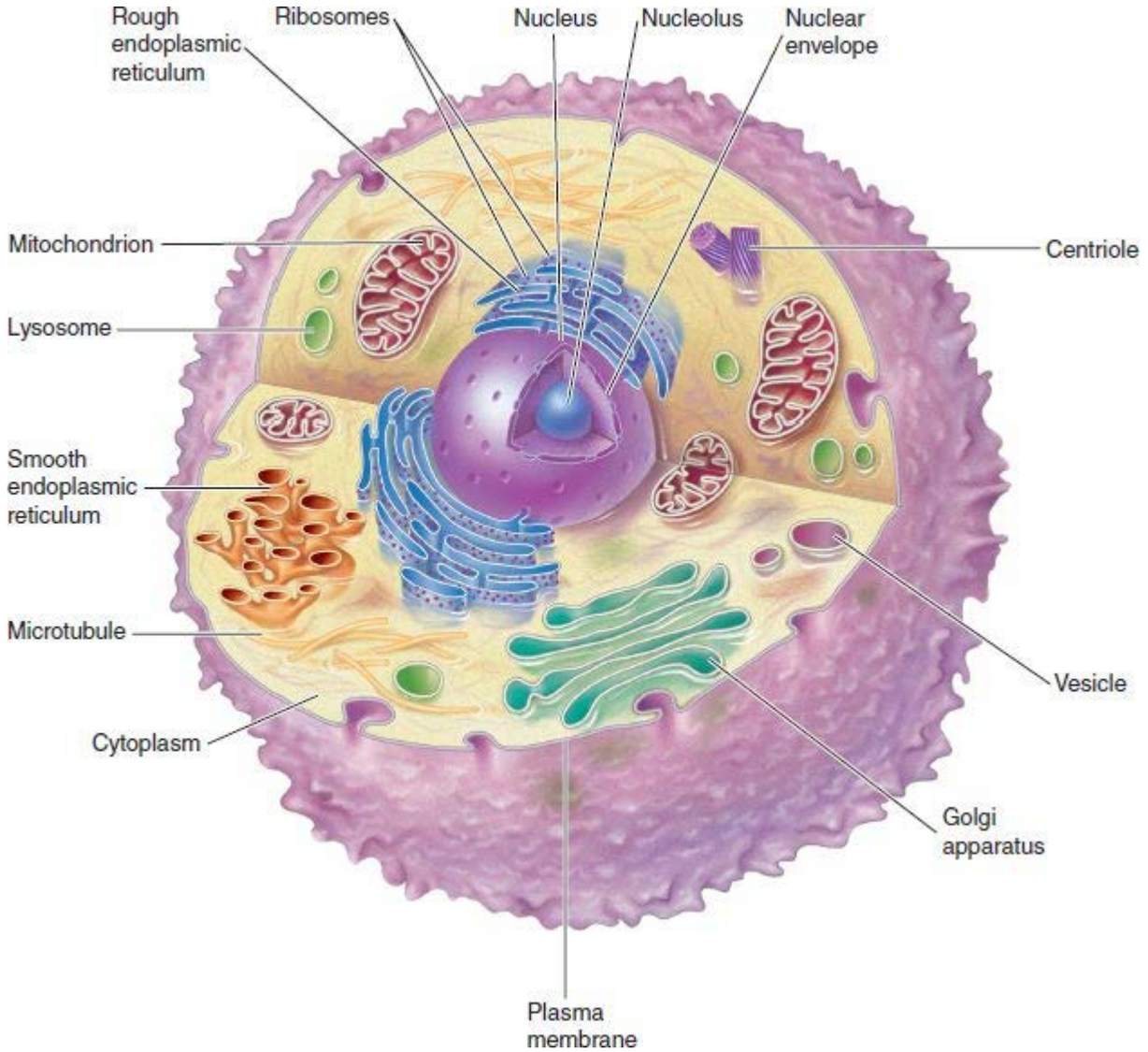


FIGURE 2.10

Animal Cell Structure

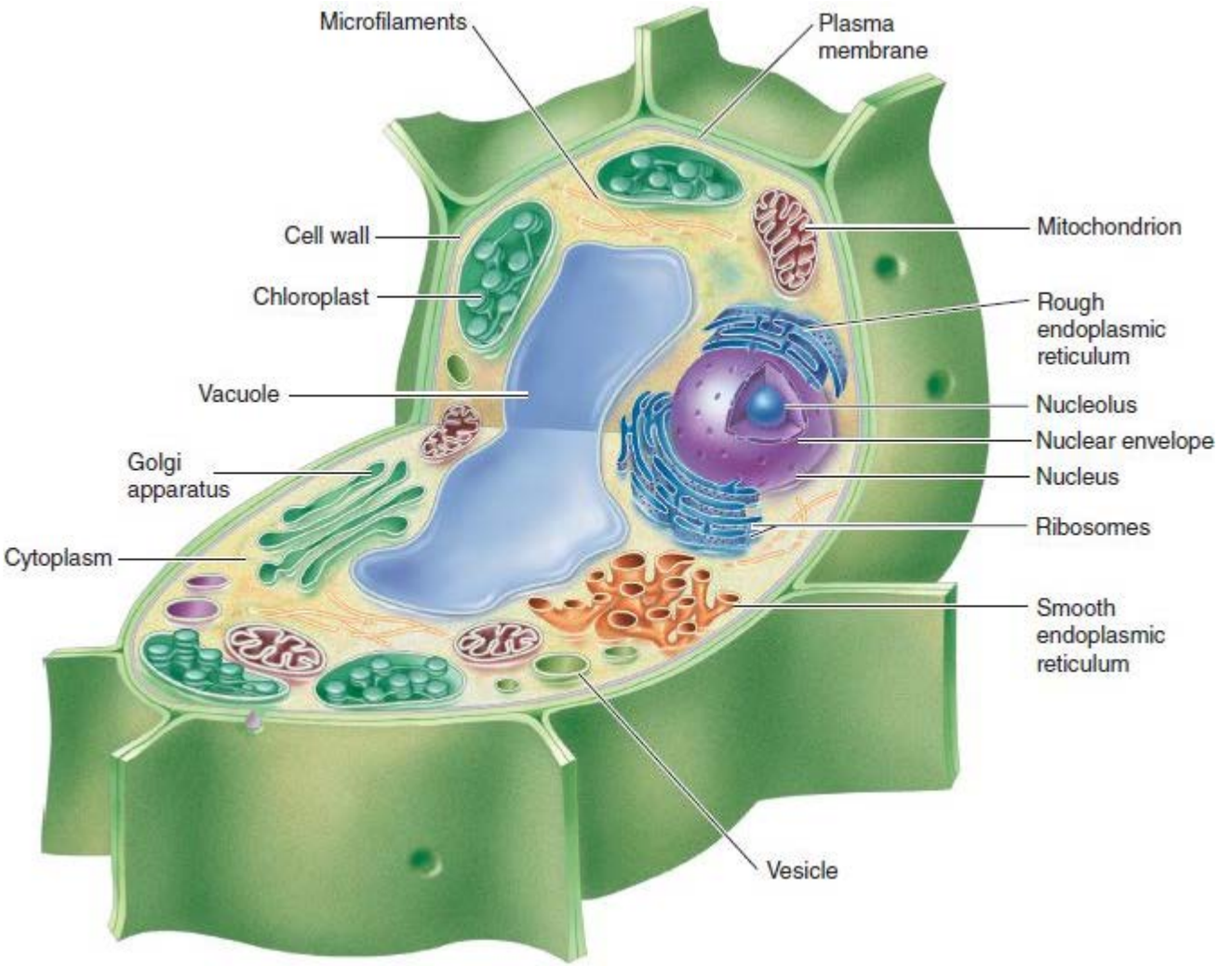


FIGURE 2.11

Plant Cell Structure

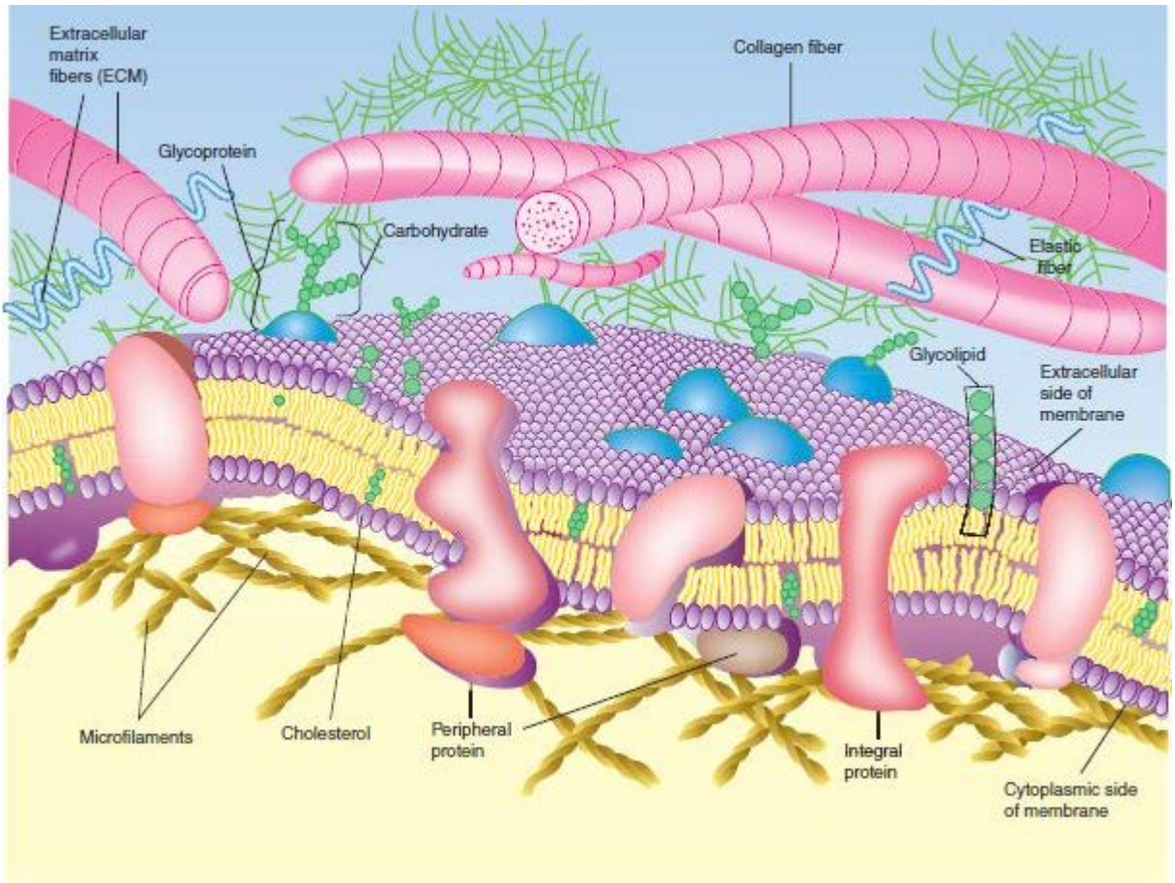


FIGURE 2.12

The Plasma Membrane of an Animal Cell

The plasma membrane (PM) is composed of a lipid bilayer in which a wide variety of integral proteins are embedded. Note that numerous integral proteins and lipid molecules are covalently attached to carbohydrate. Peripheral proteins are attached by noncovalent bonds to the cytoplasmic surface of the PM. Specialized cells of the connective tissue of higher animals called fibroblasts synthesize and secrete proteins into the extracellular matrix (ECM; e.g., elastin and collagen). The inner surface of the PM is reinforced by the membrane skeleton, which is composed of a meshwork of actin microfilaments and other proteins linked to the cell's cytoskeleton.

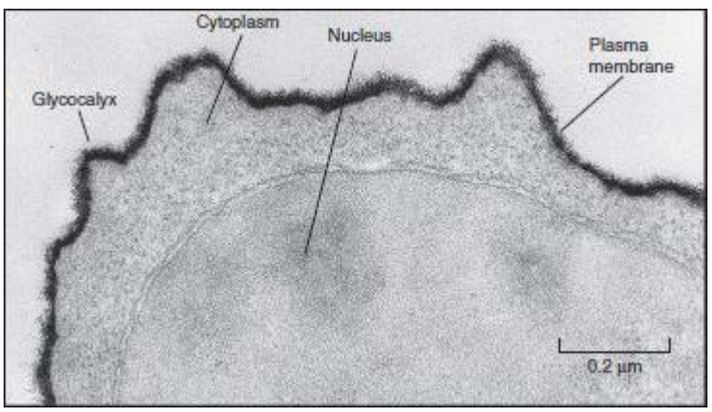


FIGURE 2.13

The Glycocalyx

Electron micrograph of the surface of a lymphocyte stained to reveal the glycocalyx (cell coat).

KEY CONCEPTS

- The plasma membrane provides mechanical strength and shape to the cell, and is actively involved in selecting the molecules that can enter or exit the cell.
- Receptors on the plasma membrane's surface allow the cell to respond to external stimuli.

Endoplasmic Reticulum

The **endoplasmic reticulum (ER)** is a system of interconnected membranous tubules, vesicles, and large flattened sacs that often constitutes more than half of a cell's total membrane. The repeatedly folded, continuous sheets of ER membrane enclose an internal space called the ER *lumen*. This compartment, which is often referred to as the *cisternal space*, is entirely separated from the cytoplasm by the ER membrane. The ER is responsible for several vital processes. Among these processes are the synthesis of several kinds of protein, a variety of membrane lipids and steroid molecules, and the storage of calcium ions.

ER comes in two interconnected forms: **rough ER (RER)** and **smooth ER (SER)** (**Figure 2.14**). The precise functional properties and relative sizes of both ER types vary with cell type and physiological conditions. For example, the percentages of RER in hepatocytes and *pancreatic acinar* (digestive enzyme-producing) cells are approximately 35% and 60%, respectively. SER percentages in these two cell types are 16% and 1%, respectively. RER is so named because of the numerous ribosomes that stud its cytoplasmic surface. Several protein classes are processed by the RER: membrane proteins, and water-soluble proteins destined for retention within the ER, transport to other organelles or export out of the cell. Polypeptides enter the RER during ongoing protein synthesis as they are threaded, or translocated, through the membrane.

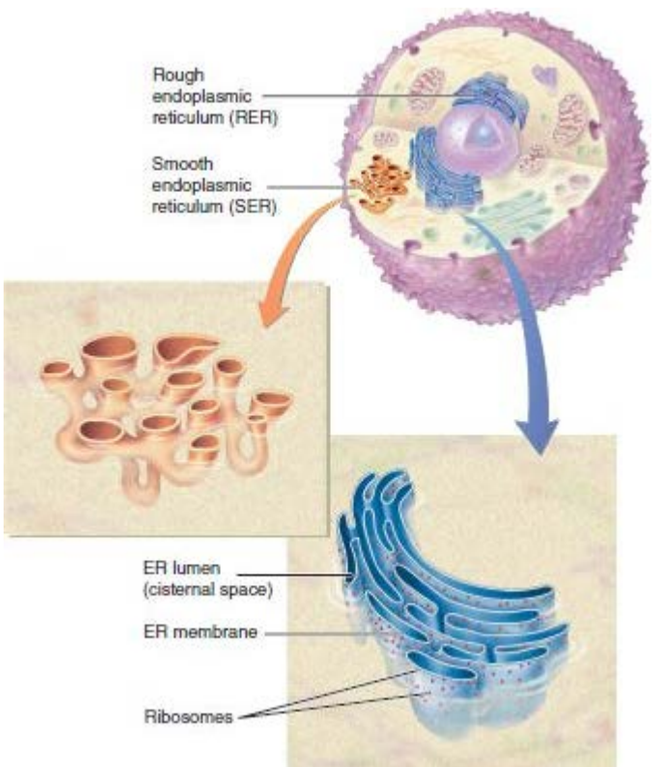


FIGURE 2.14

The Endoplasmic Reticulum

There are two forms of endoplasmic reticulum (ER): RER, the rough endoplasmic reticulum, and SER, the smooth endoplasmic reticulum. Note that in living eukaryotic cells, RER and SER are interconnected.

Transmembrane polypeptides (i.e., those that contain one or more hydrophobic sequence segments) remain embedded in the membrane because the translocation process is halted when hydrophobic segments enter the membrane. As water-soluble polypeptides emerge into the ER lumen, the folding process, facilitated by processing enzymes and *molecular chaperones* (proteins that facilitate protein folding), begins. Glycosylation reactions, the attachment of carbohydrate groups to specific amino acid residues, are a prominent example of ER-processing reactions. The binding of molecular chaperones to short hydrophobic segments in partially folded polypeptides facilitates efficient folding and prevents aggregation.

The failure of polypeptides to fold within the RER, resulting in an accumulation of misfolded molecules, is a potential threat to the entire cell since overall cell function can be disrupted. This phenomenon, called **ER stress**, is caused by environmental factors such as metabolic stress (changes in metabolism triggered by injury, illness, or infection), oxidative stress (from oxygen radicals), and activated inflammatory signaling processes, as well as genetic factors. **ER-associated protein degradation** (ERAD) is a cellular mechanism that targets misfolded polypeptides and transports them into the cytoplasm, where they are degraded by proteasomes (p. 572). If stress is severe, the RER initiates the **unfolded protein response (UPR)** in an attempt to restore proteostasis. Signals sent to the nucleus result in the inhibition of protein synthesis, with the exception of additional molecular chaperones. In addition to proteosomal protein destruction, autophagy (controlled digestion of damaged or unnecessary organelles or other cell components; see p. 574) can be utilized in an attempt to prevent cell death. If protein homeostasis cannot be achieved within a certain time period, *apoptosis* (p. 58), a programmed cell death process can be initiated.

Smooth ER lacks attached ribosomes, and its membranes are continuous with those of RER. The size and functional properties of SER vary considerably in different cell types from sparse to abundant. In most cells, SER is involved in the synthesis of lipid molecules. The SER is especially prominent in hepatocytes and striated muscle cells. Hepatocyte SER performs a wide variety of functions, which include biotransformation and synthesis of the lipid components of very low-density lipoproteins (water-soluble lipid transport complexes that deliver lipids to tissue cells). **Biotransformation reactions** convert an enormous variety of water-insoluble metabolites and xenobiotics (foreign and potentially toxic molecules) into more soluble products that can then be excreted. The SER in striated muscle is so highly specialized in both structure and function that it has a different name, the *sarcoplasmic reticulum* (SR). The SR membrane extends throughout the muscle cell and is in close proximity to all myofibrils, the organized arrays of contractile proteins. SR is a reservoir for calcium, the signal that triggers muscle contraction.

Newly synthesized protein and lipid molecules exit the ER in coated vesicles that bud off from exit sites in an ER subdomain referred to as the *transitional ER* (tER). The vesicular coat of *COPII* (*co*at *p*rotein complex II) and its adapter proteins ensure that the vesicle is directed to the correct target membrane. After exiting the tER, the vesicles are transported to the *ER-Golgi intermediate compartment* (ERGIC), a structure of membranous tubules and vesicles that facilitates the sorting of cargo molecules from resident ER proteins. Newly formed COPII-coated vesicles deliver the molecular cargo to the Golgi complex for further processing. The resident ER molecules, identified by retrieval signals in their structures, are recycled by returning to the ER via vesicles that have *COPI* (*co*at *p*rotein I) coats.

KEY CONCEPTS

- The RER is primarily involved in protein synthesis. The external surface of the RER membrane is studded with ribosomes.
- SER lacks attached ribosomes and is involved in lipid synthesis, biotransformation, and Ca^{2+} storage.

Golgi Apparatus

The **Golgi apparatus** (also known as the **Golgi complex**) is formed from relatively large, flattened, saclike membranous vesicles that resemble a stack of plates. The Golgi apparatus is involved in the processing, packaging, and distribution of cell products (e.g., glycoproteins) to internal and external compartments (**Figure 2.15**). The primary role of the Golgi apparatus is the glycosylation (covalent attachment of sugar molecules) to proteins and lipids. Sulfation and phosphorylation reactions also occur.

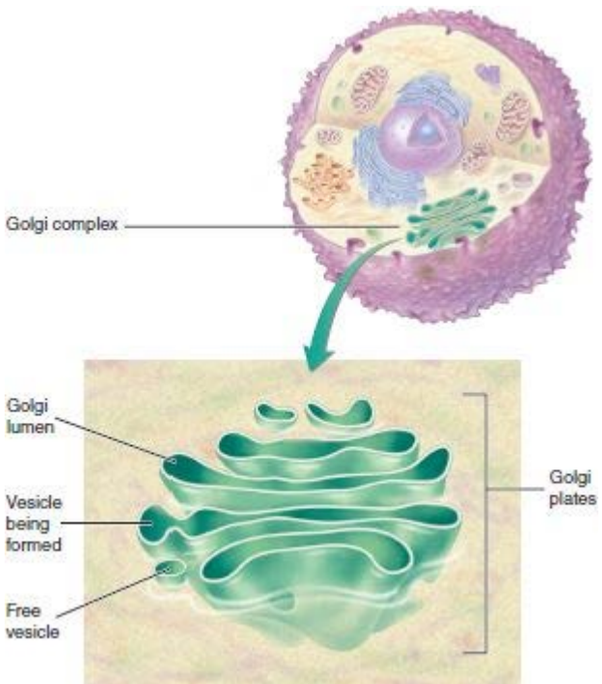


FIGURE 2.15

The Golgi Apparatus

The Golgi apparatus is essentially a factory that synthesizes and/or processes a diverse group of proteins and lipids. These biomolecules are then sorted for transport to their final destination.

The Golgi apparatus has two faces. The plate (or *cisterna*) positioned closest to the ER is on the forming (cis) face, whereas the one on the maturing (trans) face is typically close to the portion of the cell's plasma membrane that is engaged in secretion. Several small membranous vesicles coated with COPII and containing newly synthesized protein and lipid bud off from the ER. Moving quickly along microtubules, the vesicles, in a process referred to as **anterograde transport**, then lose their coats and fuse together to form the *vesicular tubule cluster* (VTC). The VTC then proceeds to the cis face of the Golgi apparatus. COPI-coated transport vesicles containing resident ER proteins bud off from the edges of VTCs and return these molecules to the ER in a process called **retrograde transport**. According to the *cisternal maturation model*, the

VTCs mature into the *cis*-Golgi network (CGN), which in turn forms the cis cisterna. As the cis cisterna moves forward, eventually becoming the trans cisterna, the newly synthesized proteins and lipids enzymes undergo an ordered series of covalent modifications (the addition of sugar molecules and sulfate or phosphate groups) that prepare them for their functions. Golgi proteins are constantly recycled. At each cisterna, COPI-coated vesicles are formed that transfer these molecules from older to younger cisternae.

A complex network of *trans*-Golgi membrane and vesicles, called the *trans-Golgi network* (TGN), sorts the processed molecules and packages them into vesicles coated with a protein called clathrin (p. 52). Clathrin adapter proteins link clathrin to membrane-bound receptors and target clathrin-coated vesicles to destinations such as endosomes (p. 52), lysosomes, and the plasma membrane. Clathrin-coated vesicles are also used to transport vesicles from the plasma membrane to destinations such as endosomes and the TGN.

In the *secretory process*, secretory vesicles containing molecules such as digestive enzymes, hormones, or neurotransmitters are delivered to the plasma membrane, where the cargo molecules are released from the cell. Often referred to as **exocytosis** (**Figure 2.16**), this process involves the fusion of the vesicle membrane with the plasma membrane. *Constitutive exocytosis* (unregulated secretion) occurs continually in all cells. Examples include the secretion of the structural protein collagen by fibroblasts and serum albumin by liver cells. In regulated exocytosis, secretion is a Ca^{2+}-triggered process that occurs only in response to an external signal. For example, when an action potential of a motor neuron reaches the presynaptic terminal, it causes calcium channels to open. Calcium ions then trigger the fusion of neurotransmitter vesicle membrane with the nerve cell membrane, thus releasing their contents into the neuromuscular junction. The binding of sufficient numbers of acetylcholine molecules to acetylcholine receptors on the surface of the postsynaptic muscle cell results in muscle contraction.

KEY CONCEPTS

Formed from relatively large, flattened, saclike membranous vesicles, the Golgi apparatus prepares proteins and lipids for their cellular functions and packages and secretes certain cell products.

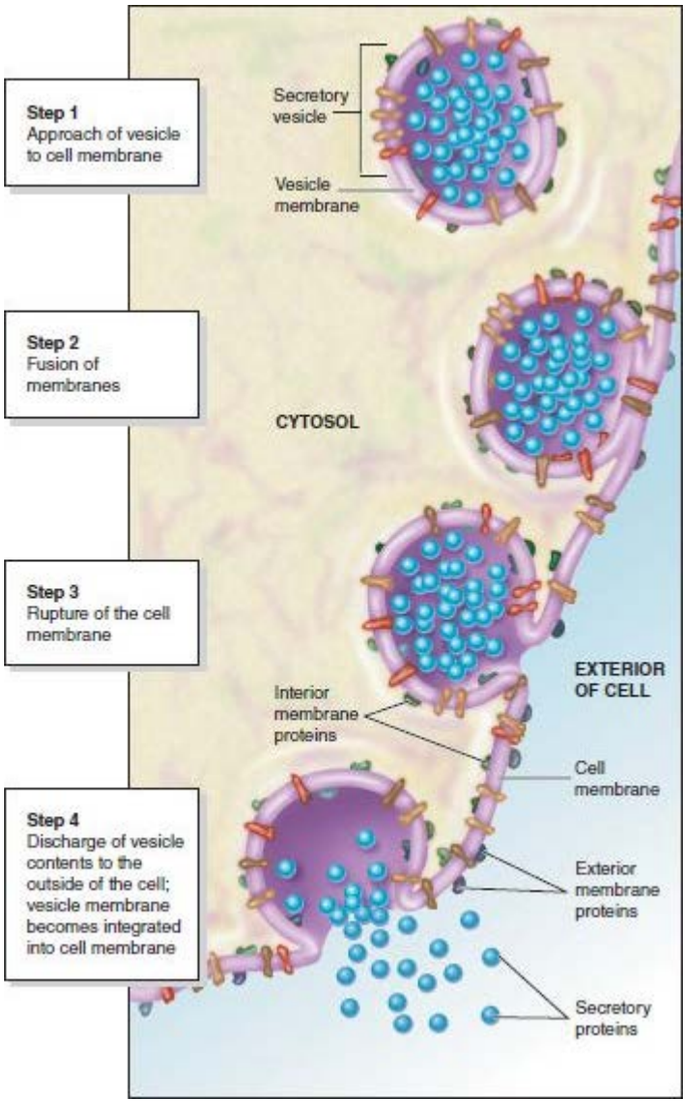


FIGURE 2.16

Exocytosis

Proteins destined to be secreted by a cell are produced in the ER and processed by the Golgi apparatus, where they are packaged into vesicles that migrate to the plasma membrane and merge with it.

Vesicular Organelles and Lysosomes: The Endocytic Pathway

Endocytosis (**Figure 2.17**) is a cellular process in which plasma membrane protein receptors bound to specific substances such as lipoproteins are taken into cells, most notably involving the pinching off of regions of the plasma membrane. The newly made vesicles then enter the *endocytic pathway* when they fuse with a membrane-bound organelle called an early endosome. Usually located near the cell's periphery, *early endosomes* serve as the nexus or focal point of the endocytic pathway because it is here that the fate of internalized molecules is determined. An elaborate mechanism involving regulatory proteins ensures that internalized molecules are appropriately recycled back to the plasma membrane, delivered to the TGN for transport to locations throughout the cell, or degraded within organelles called lysosomes. **Lysosomes** are vesicles that contain granules consisting of digestive enzymes called *acid hydrolases*, which catalyze the attack of a water molecule on ester and amide linkages under acidic conditions. In addition to their role in endocytosis, lysosomes also contribute to the autophagic degradation of debris within cells (p. 574).

Early endosomes are tubular-vesicular networks that mature to form *late endosomes*, which are also called *multivesicular bodies* because they contain numerous closely packed vesicles. The maturation process is achieved in part by an increase in hydrogen ion concentration (i.e., a reduction in internal pH) through the activity of V-ATPase (an ATP-dependent proton pump) and the arrival of TGN vesicles containing lysosomal acid hydrolases and membrane proteins. Late endosomes are converted to fully functional lysosomes when the internal pH is less than 5, a circumstance that activates the acid hydrolases. Late endosomes may also fuse with existing lysosomes.

There are several forms of endocytosis: the best-researched example is clathrin-dependent endocytosis, which is described next. Caveolar endocytosis, a type of clathrin-independent process, is briefly described in an online reading.

Clathrin-dependent endocytosis, also referred to as *receptor-mediated endocytosis*, is a versatile and widely used mechanism in which clathrin-coated vesicles containing cargo, bound to membrane receptors, are taken into cells. Examples of the use of clathrin-dependent endocytosis include nutrient uptake [e.g., low-density lipoproteins (LDLs), sources of lipids such as cholesterol (p. 417); and transferrin, an iron-binding protein], intercellular signal transduction, and membrane recycling. The process begins with the binding of a specific ligand to its cognate receptor on the external surface of the plasma membrane. Adapter proteins then bind to the cytoplasmic side of the receptor–ligand complex, after which clathrin is recruited. **Clathrin** is a soluble protein complex called a *triskelion* (three heavy chains and three light chains) because of its shape (**Figure 2.18**). As the clathrin triskelia bind to the adapter proteins, a basketlike latticework forms that forces the membrane into the shape of a bud. The clathrin-coated vesicle is then excised from the plasma membrane by dynamin, a GTP-requiring protein that encircles and constricts the vesicle's neck until a fully formed coated vesicle is released from the plasma membrane. Vesicle fusion with the early endosome is preceded by removal of the clathrin coat. When vesicles, such as those containing LDLs, have fused with early endosomes, the reduction in pH releases the cargo from their receptors. The LDL receptors are recycled back to the plasma membrane, and LDL molecules (lipids and proteins) are degraded within lysosomes.

Once believed to be a simple process, endocytosis is now recognized as being fully integrated into cellular signaling and regulation. Endocytosis in combination with exocytosis (p. 51), referred to as the **endocytic cycle**, plays a central role in cellular information processing. Until recently, the endocytic cycle has been regarded as a means of controlling a cell's response to a signal molecule by regulating the number of its cognate receptors in the plasma membrane. In this view endocytosis is a mechanism for downregulating receptors, thereby desensitizing cells to signaling molecules. Recent evidence indicates that endocytosis pathways contribute in other ways to signal transduction. For example, signaling from some receptors, such as insulin receptor and thyroid-stimulating hormone receptor, has been shown to continue after they enter endosomes. In addition, endosomes may serve as signaling platforms because they contain protein or lipid components not present in the plasma membrane, and endosomal pathways offer opportunities for signal diversification.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay Caveolar Endocytosis.

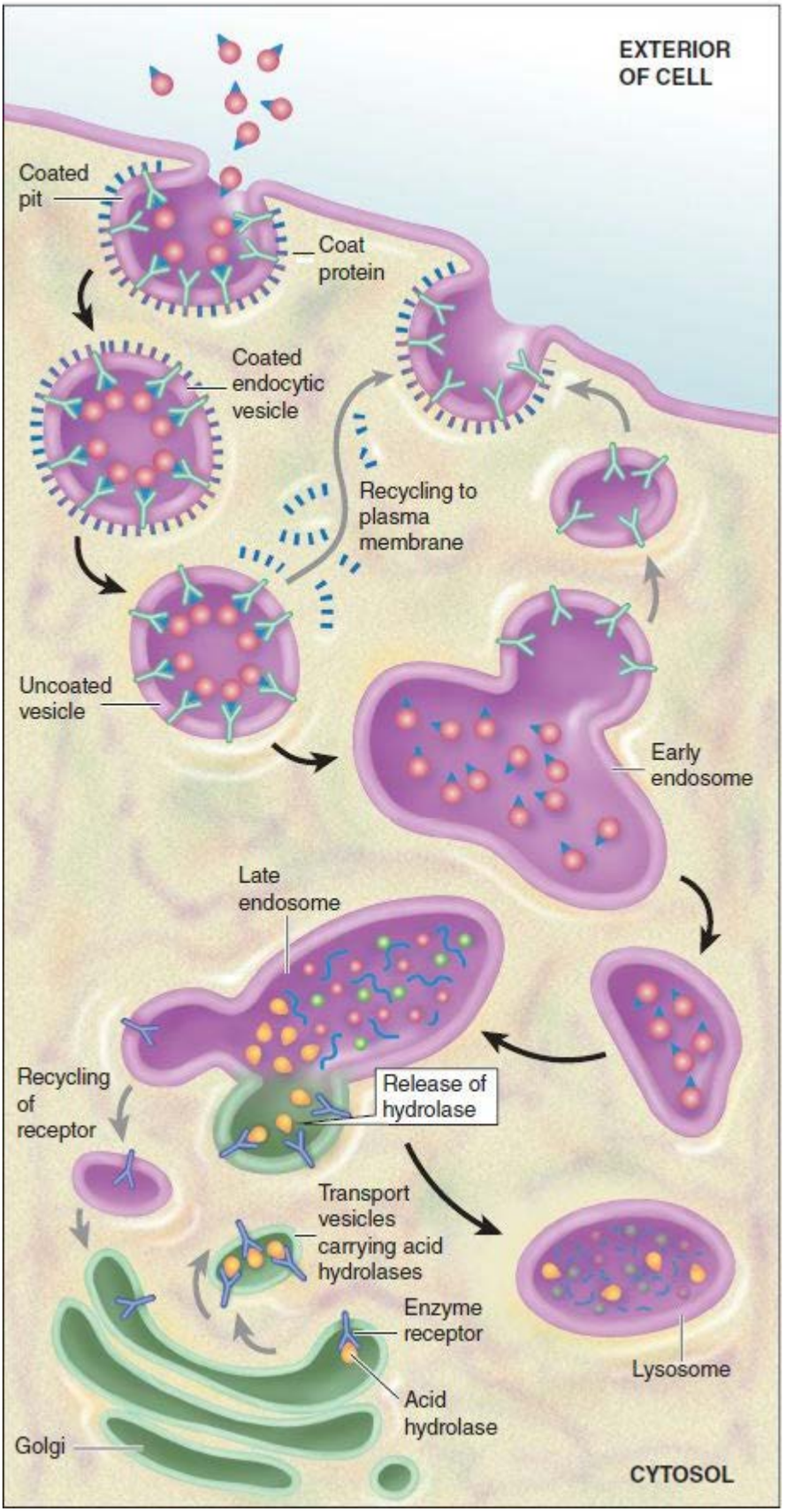


FIGURE 2.17

Receptor-Mediated Endocytosis

Extracellular substances may enter the cell during endocytosis, a process in which receptor molecules in the plasma membrane bind to the specific molecules or molecular complexes called ligands. Specialized regions of plasma membrane called coated pits (composed of clathrin triskelia, not shown) progressively invaginate to

form closed vesicles. After the coat proteins are removed, the vesicle fuses with an early endosome, the precursor of lysosomes. The coat proteins are then recycled to the plasma membrane. During endosomal maturation, the proton concentration rises and the ligands are released from their receptors, which are subsequently also recycled back to the plasma membrane. As endosomal maturation continues, lysosomal hydrolases are delivered from the Golgi apparatus. Lysosomal formation is complete when all the hydrolases have been transferred to the late endosome and the Golgi-derived vesicular membrane has been recycled back to the Golgi apparatus.

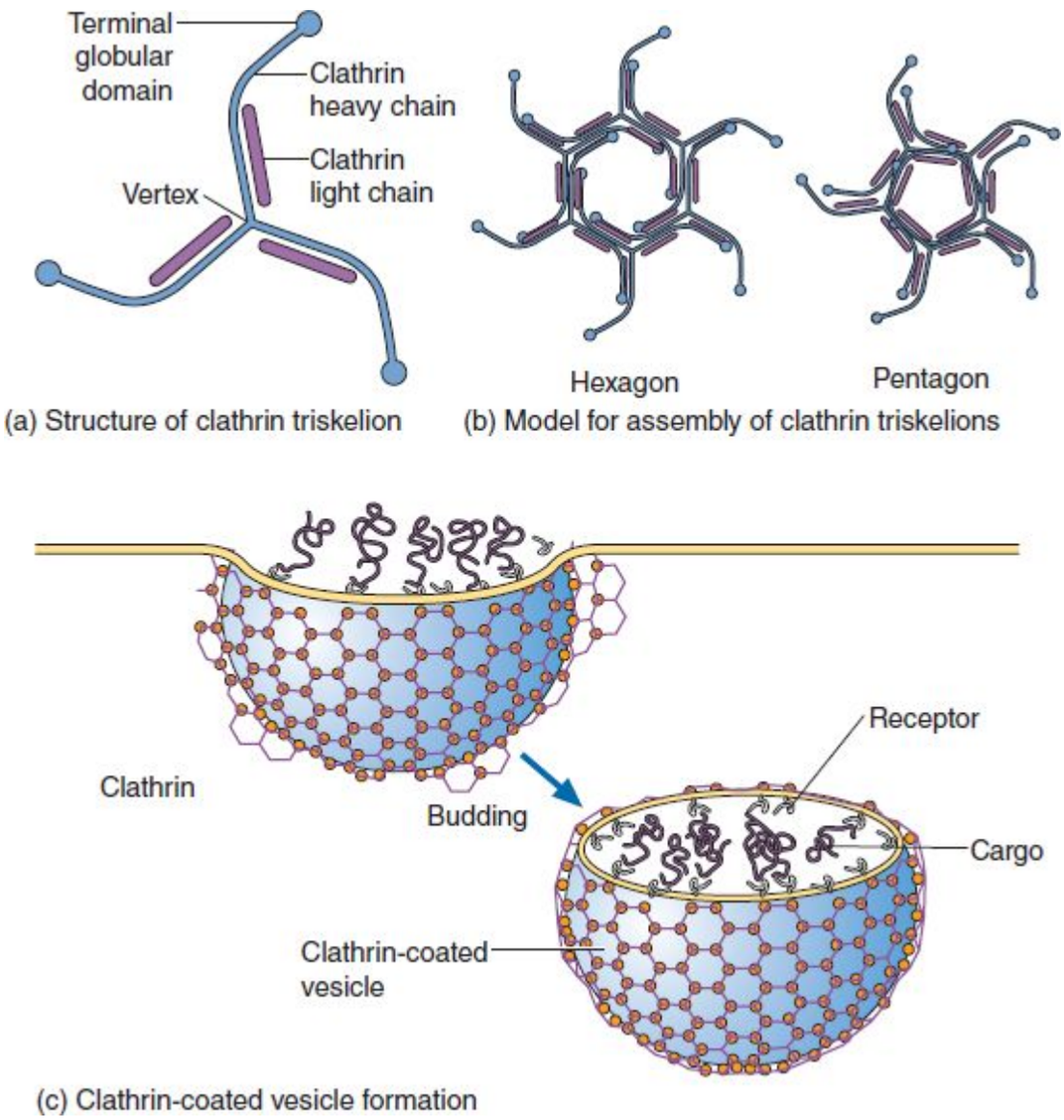


FIGURE 2.18

Clathrin-Dependent Endocytosis

(a) Each clathrin triskelion is formed by three heavy chains and three light chains. (b) Triskelia combine to form the hexagons and pentagons observed in the latticework of clathrin-coated vesicles. (c) Clathrin-coated vesicle formation is initiated by GTP-binding proteins (not shown) that recruit adaptor proteins that serve as binding sites for clathrin. The association of clathrin triskelia to form hexagons (in this figure) causes the membrane distortion required in the vesicle-forming process.

QUESTION 2.2

In many genetic disorders, a lysosomal enzyme required to degrade a specific molecule is missing or defective. These maladies, often referred to as *lysosomal storage diseases*, include Tay–Sachs disease (p. 415). Afflicted individuals inherit from each parent a defective gene that codes for an enzyme that degrades a complex lipid molecule. Symptoms include severe mental retardation and death before the age of five years. What is the nature of the process that is destroying the patient's cells? [*Hint:* Synthesis of the lipid molecule continues at a normal rate.]

Lysosomal Storage Diseases

Nucleus

The **nucleus** (**Figure 2.19**) is the most prominent organelle in eukaryotic cells. It contains most of the cell's DNA. Low-resolution micrographs reveal that nuclear structure consists of a seemingly amorphous nucleoplasm surrounded by membrane, the nuclear envelope. **Nucleoplasm** contains a network of **chromatin fibers**, which during the mitotic phase of the cell cycle are condensed to form the chromosomes that will be distributed to daughter cells. (In humans, for example, 2 m of DNA are compressed into a nucleus with a diameter of 10 μm.) Chromatin is highly structured, consisting of DNA and DNA packaging proteins known as the *histones*. Chromatin has been traditionally classified according to how densely it is packaged. *Euchromatin*, the lightly packaged form, is usually gene-rich and is easily accessed by transcription factors and transcription enzyme complexes. The other form, called *heterochromatin*, is tightly packed and inaccessible for transcription. *Constitutive heterochromatin*, containing the sequences for chromosomal structures such as centromeres and telomeres and remnants of ancient viruses, is highly condensed and permanently inactive. *Facultative heterochromatin* is capable of changing in response to specific signaling processes from a condensed inactive state into actively transcribed euchromatin. Each type of differentiated cell has its own specific set of facultative heterochromatin. Chromatin distribution within the nucleus is not random. Chromosomes occupy discrete locations called *chromosome territories*. In general, chromosome segments that are gene-poor or that contain heterochromatin typically occur in the periphery of the nucleus. Actively transcribed, gene-dense euchromatin is located more centrally. Chromosome location varies among different cell types.

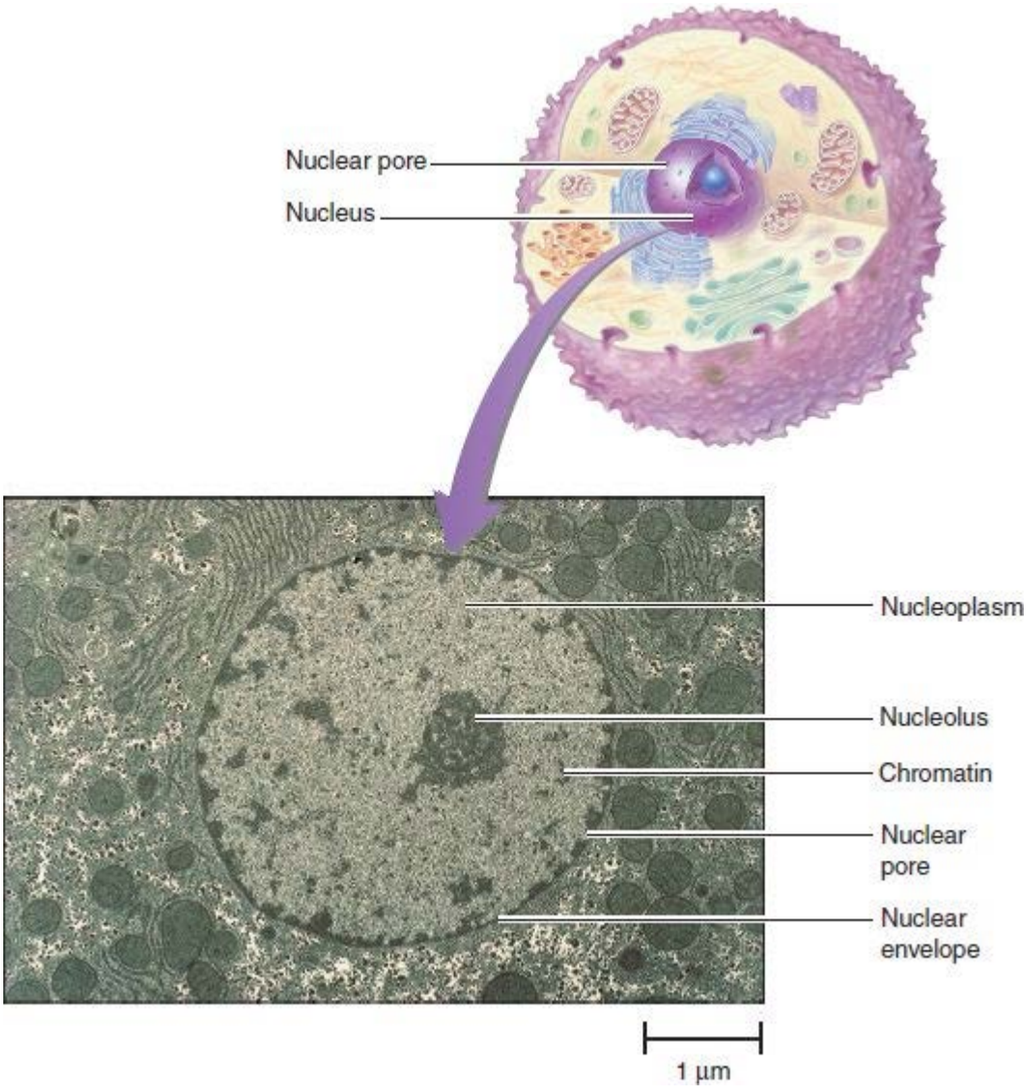


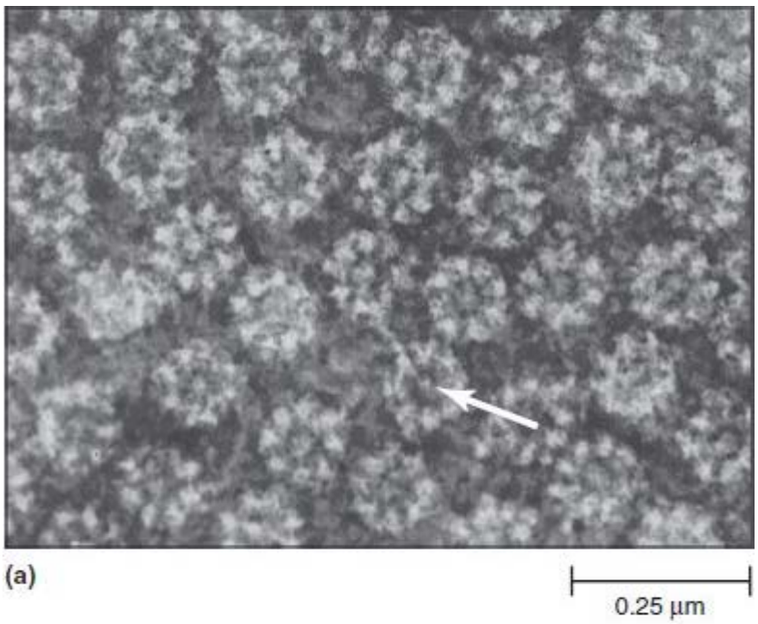
FIGURE 2.19

The Eukaryotic Nucleus

The nucleus is an organelle surrounded by a double membrane, the nuclear envelope. The nuclear envelope, a barrier that prevents the free passage of molecules between the nuclear compartment and the cytoplasm, plays a vital role in gene expression regulation.

The **nuclear envelope** (NE) acts as a barrier that prevents the free passage of molecules between the nucleus and the cytoplasm. As a result, processes such as DNA replication and transcription are more easily regulated. The NE is composed of two concentric membranes. The **outer nuclear membrane** (ONM) is continuous with the RER, and ribosomes are attached to its cytoplasmic surface. Among the many RER proteins on the cytoplasmic side of the ONM, several (e.g., nesprins) bind to the filaments of the cytoskeleton. Unlike the outer membrane, the **inner nuclear membrane** (INM) contains integral proteins that are unique to the nucleus. In addition to stabilizing NE structure, the functions of these proteins include chromatin binding, chromatin remodeling protein recruitment, and various enzyme activities. The space between the two membranes, the **perinuclear space** (diameter 20–50 nM), is continuous with the lumen of the rough ER. The inner and outer membranes fuse at structures called *nuclear pores*, which are elaborate macromolecular structures that regulate molecular traffic between the cytoplasm and the nucleus. Nuclear pores, referred to as **nuclear pore complexes** (NPCs) (**Figure 2.20**), vary in

number in vertebrates from 2000 to 4000 per nucleus. Each NPC is a 60 to 100 MDa structure (diameter = 100 nm). In vertebrates, NPCs consist of several hundred proteins called the *nucleoporins* (a family of 30 proteins). The function of NPCs was once thought to be limited to nucleocytoplasmic molecular transport. Recent research, however, has revealed that nucleoporins also have roles in chromatin organization and in DNA replication and repair.



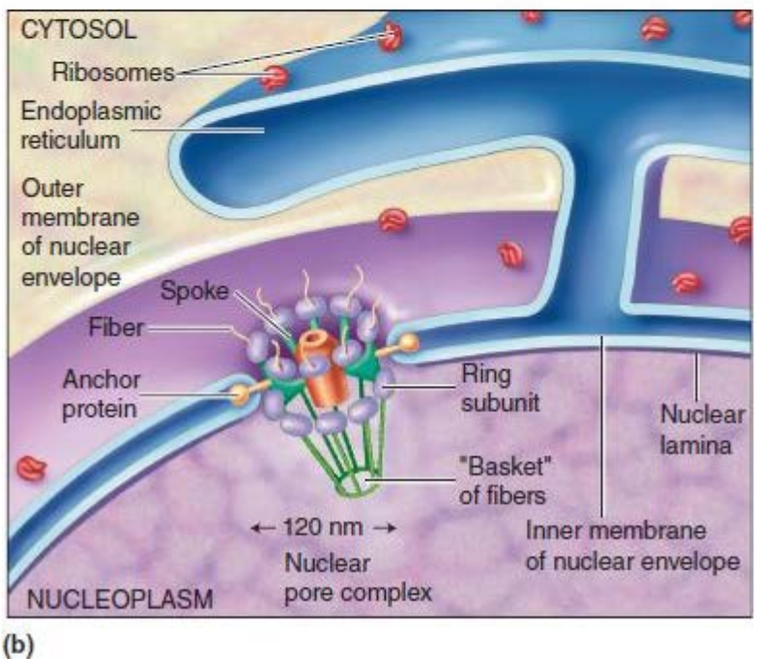


FIGURE 2.20

The Nuclear Pore Complex

(a) The nuclear envelope is studded with thousands of nuclear pore complex structures, one of which is indicated by the arrow. (b) The basic structure of NPCs is a donut-shaped scaffold seated on top of a basketlike structure; the center is a pore-like opening. Unstructured polypeptides (not shown) in the center ensure the selective transport of cargo molecules bound to nuclear transport proteins.

The membrane-embedded ring-shaped core of the NPC is attached to a basket-shape structure.

Filaments that extend from the cytoplasmic and nucleoplasmic side of the NPC function as docking sites for large molecules that will subsequently be transported through the pore. A meshwork formed from flexible nucleoporins that line the central pore restricts transport through the NPC only to those macromolecules (e.g., RNAs and large proteins) bound to either import or export chaperone proteins. Small substances such as ions and small proteins (40 kDa or less) diffuse through the NPC, which has a functional diameter of about 9 nm. Driven by the hydrolysis of the nucleotide GTP, traffic through the NPC is brisk and efficient.

Proteins entering the nucleus must have a nuclear localization amino acid sequence signal, which is recognized by a cargo transport protein called an *importin*. In a process driven by the energy released by GTP hydrolysis, the protein–importin complex is then transported through the nuclear pore. The nuclear export process, also driven by GTP hydrolysis, is similar to the import process. Cargo leaving the nucleus binds to proteins with nuclear export signal sequences. The newly formed molecular complexes then bind to *exportins*, the nuclear transport proteins that mediate cargo transport through the NPC and into the cytoplasm. About 1000 macromolecules pass through each NPC per second.

The **nuclear lamina** (**Figure 2.21**) is a thin, dense protein meshwork attached to the inner surface of the inner nuclear membrane. Once thought to only provide shape and mechanical stability to the nuclear envelope, the nuclear lamina is now believed to have roles in numerous nuclear processes, including DNA replication, transcription, and chromatin organization. The nuclear lamina consists largely of lamin filaments and lamin-associated proteins. As a result of linkages between lamin filaments, lamin-associated proteins, and the cell's cytoskeleton (pp. 62–66), external mechanical force strong enough to deform the cytoskeleton can alter the shape of the NE and possibly alter chromatin organization within the nucleus.

The lamins are intermediate filament proteins (see p. 65) that are classified as type A (lamins A and C) and type B (lamins B1 and B2). Type A and type B lamins each polymerize to form separate types of filament. Examples of lamin-associated proteins include the integral INM proteins emerin, LBR, and SUN-domain proteins. Emerin binds directly to lamin A filaments and to BAF, a DNA-bridging protein with roles in chromatin organization and gene expression. LBR (lamin B receptor) binds to both B lamin filaments and heterochromatin. It is noteworthy that the nuclear lamina and its associated heterochromatin do not extend to NPCs. INM SUN-domain proteins bind with one or more of the nesprins in the ONM to form LINC (*li*nks *n*ucleoskeleton and *c*ytoskeleton) complexes. Since the nesprins bind directly or indirectly to cytoskeletal filaments, LINC connects the nucleoplasm to the cytoskeleton.

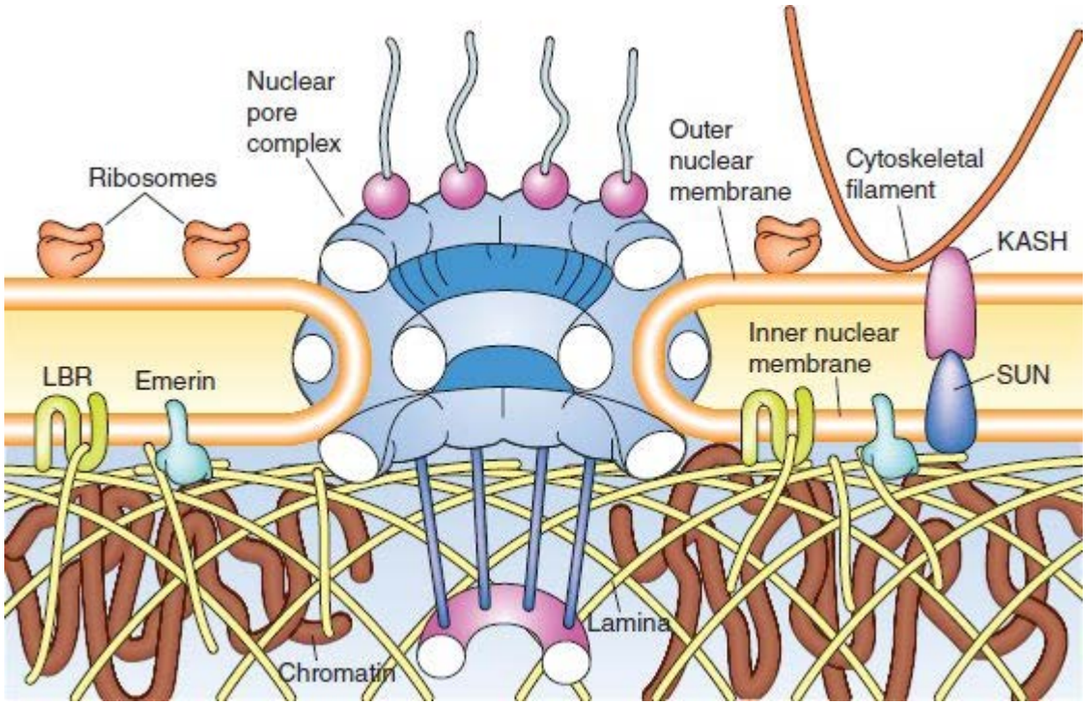


FIGURE 2.21

The Nuclear Lamina

A thin, dense network of lamin filaments is linked to the inner nuclear membrane via interactions with lamin-associated proteins such as emerin, LBR (lamin B receptor), and SUN-domain proteins. Note that emerin and LBR also bind to chromatin, although emerin does so indirectly via the protein BAF (not shown). SUN domain proteins, which are integral INM proteins, bind to the KASH domains of the ONM nesprins to form the LINC complex. The LINC complex thus connects the nuclear lamina to the cell's cytoskeleton.

Regions of the nucleus called *nuclear bodies* have been found to contain certain types of chromatin sequences and nuclear proteins. Examples include nucleoli, speckles, and Cajal bodies. The **nucleolus** is the largest of the nuclear bodies. Its best understood functions are the transcription of *rRNA* genes, rRNA processing reactions, and ribosomal subunit synthesis. After their transport into the cytoplasm, the ribosomal subunits bind to mRNAs to form ribosomes, the macromolecular complexes that synthesize proteins. The most recognizable feature of the nucleolus is the *nucleolar organizer region*, which is formed by the association of several chromosomal segments, each with multiple copies of *rRNA* genes. *Speckles*, as many as 50 per nucleus, are storage sites for ncRNA-protein complexes involved in converting pre-mRNAs into mature mRNAs. There are between 1 and 10 *Cajal bodies* (0.2 μM–2 μM = diameter) per nucleus. Named after the Spanish histologist Santiago Ramón y Cajal (1852–1934), Cajal bodies are sites of processing reactions of histone mRNAs and several ncRNAs. There are also discrete sites within the nucleoplasm, called *transcription factories*, in which active genes come together to be transcribed to yield mRNAs that code for polypeptides or numerous types of other RNAs by multiple copies of transcription enzyme complexes.

The **nuclear matrix** (nucleoskeleton) is a scaffold within the nucleoplasm composed of a large number of proteins in which loops of chromatin are organized. It is believed to be analogous to the cell's cytoskeleton because various forms of cytoskeletal proteins (e.g., actins, actin binding proteins, and myosins; see p. 62) have been located within the nucleoplasm. Although numerous nuclear processes have been shown to require specific nucleoskeletal proteins, their structural properties remain unresolved.

KEY CONCEPTS

- The nucleus contains the cell's genetic information and the machinery for converting that information into a code for protein synthesis.
- The nucleolus plays an important role in the synthesis of ribosomal RNA.

Mitochondria

Mitochondria (singular mitochondrion) are organelles that have long been recognized as the site of **aerobic metabolism**, the mechanism by which the chemical bond energy of food molecules is captured and used to drive the oxygen-dependent synthesis of adenosine triphosphate (ATP), the cell's energy storage molecule. They are often observed positioned near sites of high-energy demand. For example, most striated muscle cell mitochondria are positioned along the entire length of myofibrils.

Mitochondria play vital roles as central integrators of other metabolic processes. Prominent examples include the metabolism of amino acids and lipids, the synthesis of iron–sulfur clusters (p. 364) used in redox reactions, heme required to make hemoglobin, calcium homeostasis, and the disposal of ammonia (NH_4^+) in the urea cycle (p. 376). In recent years, mitochondria have also been recognized as key regulators of **intrinsic apoptosis**, one form of a genetically programmed series of events triggered by cell stresses (e.g., DNA damage, hypoxia, or nutrient deprivation) that leads to cell death (**Figure 2.22**). Mitochondria have traditionally been described as sausage-shaped structures with lengths ranging from 1 to 10 μm. This view has changed considerably as researchers have discovered that mitochondria have no fixed sizes. Mitochondria cannot be generated de novo; that is, new mitochondria are the result of *biogenesis*, the growth and division of preexisting mitochondria. They are dynamic organelles that are continuously dividing (*fission*), branching, and merging (*fusion*).

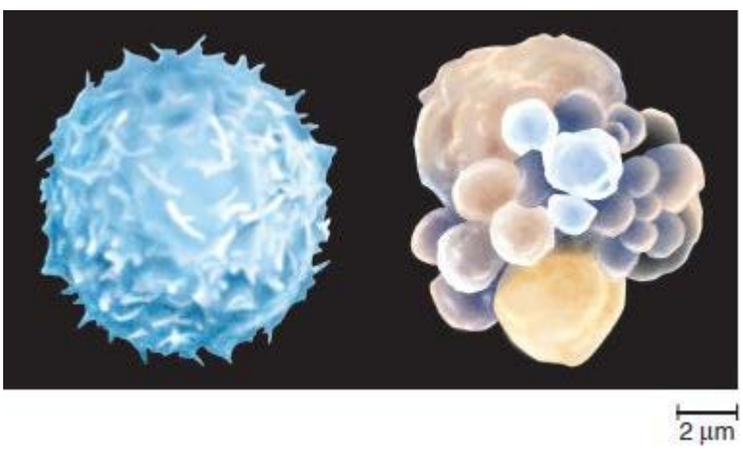


FIGURE 2.22

Apoptosis

White blood cells before (left) and during (right) apoptosis. In response to cell stress, mitochondria release a protein called cytochrome c into the cytoplasm, which facilitates the activation of enzymes that proceed to degrade cell components. The apoptotic cell shown is forming blebs that will eventually fragment into apoptotic bodies. Ultimately, phagocytes (immune system cells that digest cell debris) will ingest the apoptotic bodies.

In healthy cells, there is a constant remodeling of mitochondrial networks by continuous cycles of fission and fusion in response to the ever-changing metabolic status of cells. **Mitochondrial fission** (**Figure 2.23**) allows biogenesis when cell energy requirements are high or when damaged or inactive portions of mitochondria are segregated prior to destruction. **Mitochondrial fusion** generates extended mitochondrial networks and facilitates the rescue of mitochondria with minor damage by allowing the mixing of their contents with healthy mitochondria. An autophagic process called mitophagy removes heavily damaged mitochondria. Disruption of the delicate balance between fission and fusion has been observed in metabolic diseases such as diabetes and obesity.

Each **mitochondrion** is bounded by two membranes (**Figure 2.24**). The relatively porous smooth **outer membrane** is permeable to most molecules with masses less than 10,000 Da. The **inner membrane** is permeable to O_2, CO_2, and H_2O and impermeable to ions and a variety of organic molecules. It projects inward into folds that are called *cristae* (singular: crista). Embedded in this membrane are protein complexes and other molecules that comprise the *mitochondrial respiratory chain* (MRC). In a process referred to as *oxidative phosphorylation* (OXPHOS), the energy released by the oxidation of nutrient molecules (e.g., glucose, fatty acids, and amino acids) is transformed into the chemical bond energy of ATP. Also present are proteins that are responsible for the transport of specific molecules and ions.

Together, the inner and outer membranes create two separate compartments: (1) the *intermembrane space* and (2) the *matrix*. The intermembrane space contains several enzymes involved in nucleotide metabolism, whereas the gel-like matrix consists of high concentrations of enzymes and ions and myriad small, organic molecules. The matrix also contains 2 to 10 circular DNA molecules.

Mitochondrial DNA (mt DNA) resembles bacterial DNA in that both molecules are “naked” (i.e., not packaged with histones) and are located within a nucleoid. Of the 3000 mitochondrial genes, only 37 are encoded in the mitochondrial genome: 2 rRNAs, 22 tRNAs, and 13 MRC protein components. The remaining genes are located on nuclear chromosomes.

The number of mitochondria that cells possess varies across cell types. For example, human oocytes and hepatocytes may have as many as 200,000 and 2000 mitochondria, respectively. Most cell types have several hundred mitochondria. Erythrocytes (red blood cells) have none. Notably, the configuration of mitochondria changes with the physiological status of the cell. For example, the internal appearance of liver mitochondria has been observed to change dramatically during active respiration, that is, from a low-energy (expanded mitochondrial matrix) conformation to a high-energy (contracted mitochondrial matrix) conformation.

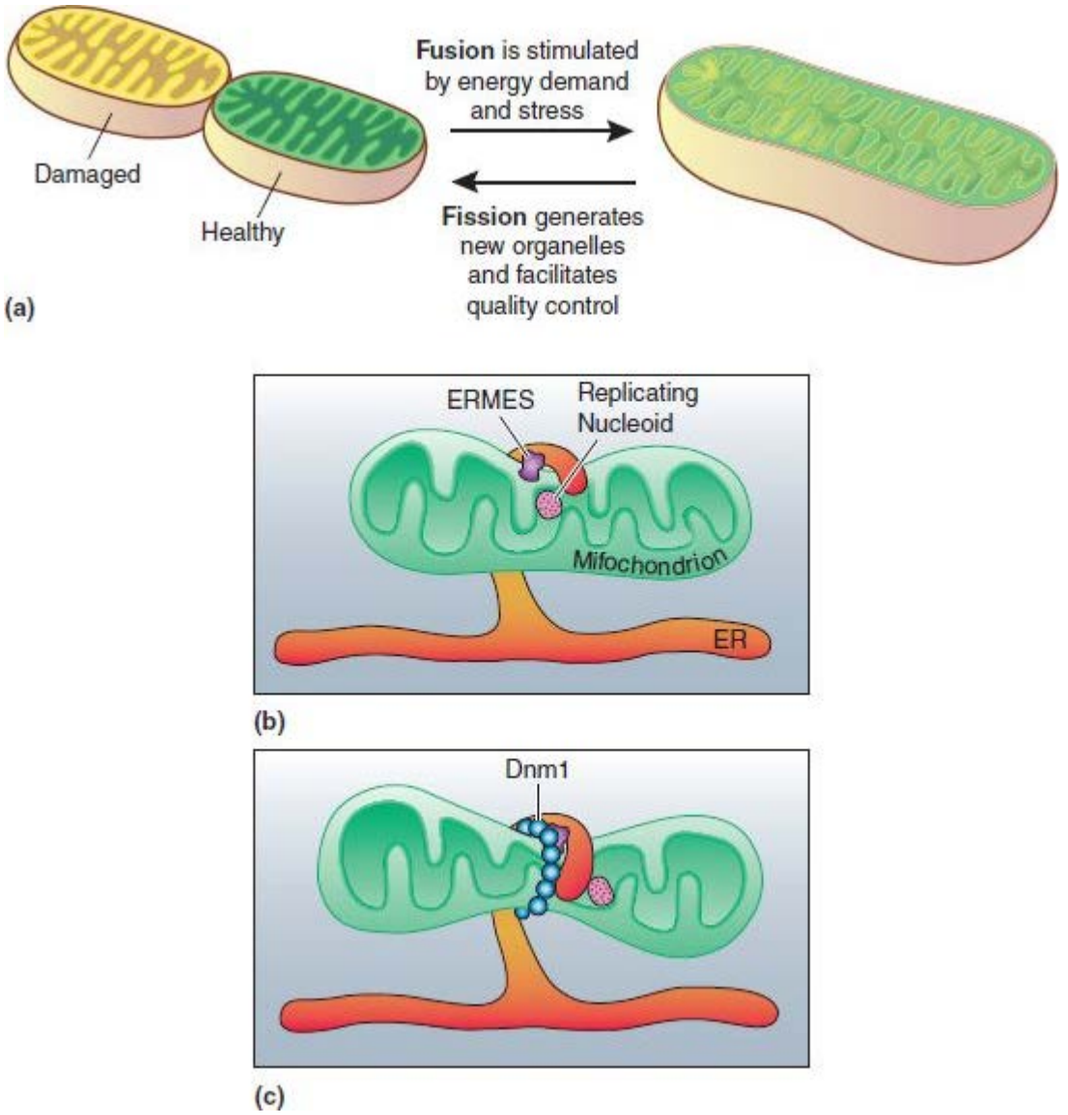


FIGURE 2.23

Mitochondrial Fission and Fusion

(a) Both fission and fusion help mitochondria remain functional when cells experience metabolic, oxidative, and other stresses. Fission creates new mitochondria and serves as a mechanism for removing damaged mitochondria. The damaged mitochondrial product of fission is then destroyed by a form of autophagy called mitophagy. Mildly damaged mitochondria can be rescued when they fuse with healthy mitochondria (i.e., components from a healthy mitochondrion compensate for those in the damaged organelle). (b) In yeast cells mitochondrial fission involves an interaction with an ER tubule, which is mediated in part by a protein tether called ERMES (ER-mitochondria encounter structure). (c) Fission is subsequently facilitated by the GTP-hydrolyzing protein dynamin-1 (Dnm1) (Drp1 in mammals), which then constricts and severs the membrane stalk between the two mitochondria.

Mitochondria form stable contact sites with regions of the ER called *mitochondria-associated membranes* (MAMs). These contacts, which consist of closely apposed membrane segments tethered by proteins, have several functions that together regulate mitochondrial dynamics:

1. **Calcium signaling.** Calcium-dependent signaling cascades initiated by a hormone, growth factor, or other type of stimulus always involve the rapid transfer of calcium into mitochondria via MAMs. The subsequent activation of Ca^{2+}-responsive proteins in every phase of ATP synthesis ensures that sufficient energy is available for the processes that are initiated by the signaling event and quenching of the calcium signal by ATP-pump

complexes in the ER and plasma membrane. Mitochondria have low-affinity calcium channels that also contribute to calcium signal quenching.

2. **Lipid exchange.** Although cellular phospholipid biosynthesis occurs predominantly in the SER, several reactions require mitochondrial enzymes. Bidirectional lipid transfer between mitochondria and the SER via MAMs is required for maintenance of the unique lipid composition of both the inner and the outer mitochondrial membranes.
3. **Mitochondrial fission regulation.** The wrapping of an ER tubule around a mitochondrion (refer to **Figure 2.23c**) is an early step in the fission process. It is believed that proteins at ER-mitochondria contact sites initiate fission by marking the division location on the mitochondrion and recruiting the fission proteins that cause constriction and division.

KEY CONCEPTS

- Aerobic respiration, the process that generates most of the energy required in eukaryotes, takes place in mitochondria.
- Embedded in the inner membrane of a mitochondrion are respiratory assemblies, where ATP is synthesized.

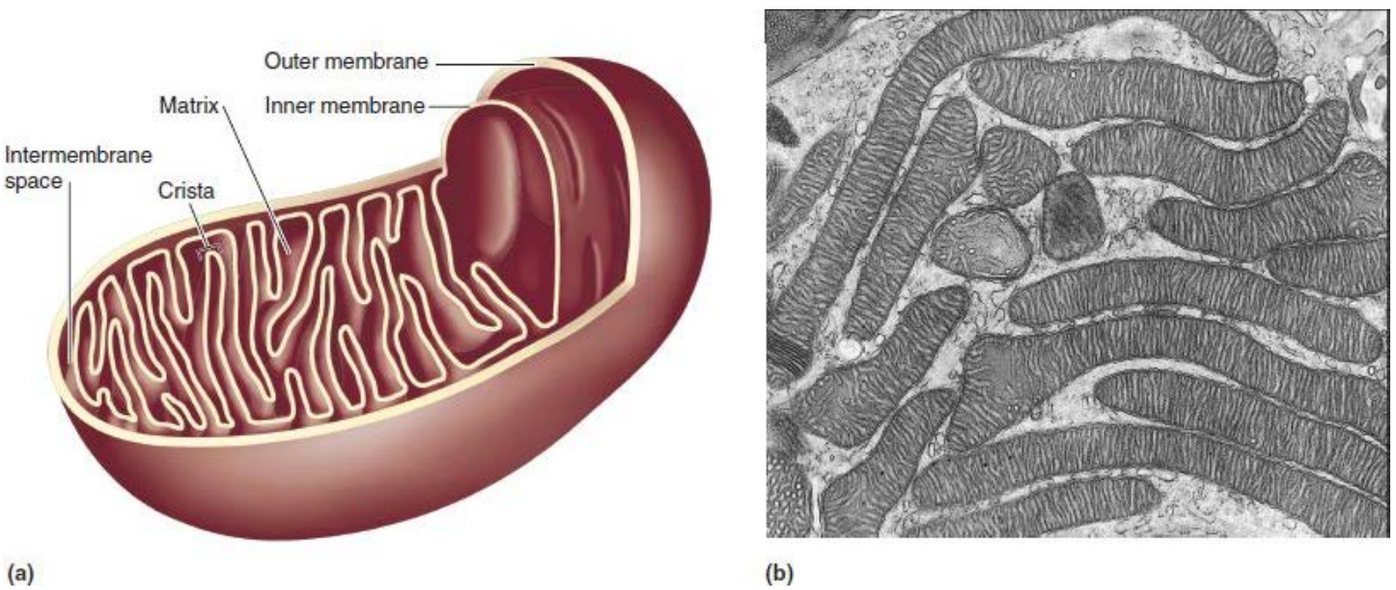


FIGURE 2.24

The Mitochondrion

(a) Membrane and crista. The internal structure depicted in this diagram is referred to as the baffle model because of the bellow-like shape of the crista. Electron tomography studies (a microscopic technique in which electron beams are used to create three-dimensional reconstructions of specimens) have revealed a more complicated anatomy. Complex arrays of fusing and dividing inner membrane tubules have been observed in the mitochondria of some tissues. The functional significance of these structural features is unknown. (b) Mitochondria from adrenal cortex, the outer layer of cells of the adrenal glands located above the kidneys.

QUESTION 2.3

It has been estimated that mitochondria occupy 20% of the volume in the human body. For a 70-

kg adult, the average number of mitochondria has been estimated to be 1×10 (10,000 trillion). Using this information, provide a rough estimate of the average mass of a mitochondrion.

Peroxisomes

Peroxisomes are small, spherical organelles (0.1 μm–0.5 μm = diameter) consisting of a single membrane that surrounds a crystalline protein matrix containing about 50 enzymes. Found in all human cells except for erythrocytes (red blood cells), peroxisomes most notably occur in large numbers in the liver and kidney. Peroxisomal enzymes are involved in a variety of anabolic and catabolic pathways, including synthesis of certain membrane phospholipids and other lipids, purine and pyrimidine bases (pp. 554–55), and bile acids (p. 480). Peroxisomes are also involved in the degradation of long-chain fatty acids, branched chain fatty acids (p. 459), polyamines (p. 650), and purine bases. As their name suggests, peroxisomes are most noted for their involvement in the generation and breakdown of toxic molecules known as peroxides. Hydrogen peroxide (H_2O_2) is generated when molecular oxygen (O_2) is used to remove hydrogen atoms from specific organic molecules.

$$RH_2 + O_2 \rightarrow R + H_2O_2$$

For example, H_2O_2 is produced in the pathway that converts purine bases into the nitrogenous waste molecule uric acid (p. 590). Peroxisomes use H_2O_2 to oxidize toxic molecules such as formaldehyde or alcohol. If not used in such reactions, this highly reactive molecule is detoxified by the enzyme catalase.

$$2H_2O_2 \rightarrow 2H_2O + O_2$$

Peroxisome biogenesis occurs by two distinct pathways. In the de novo process, preperoxisomal vesicles bud off from a specialized region of the ER. These vesicles then fuse together to form mature peroxisomes. As many as 32 peroxins, proteins required for peroxisome assembly, have been identified. Preexisting peroxisomes can grow larger with the transfer of membrane from the ER and then divide to produce new peroxisomes.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay Organelles and Human Disease.

Chloroplasts

Chloroplasts are specialized chromoplasts that convert light energy into chemical energy. (A chromoplast is a cellular organelle in plants that accumulates the pigments responsible for the colors of leaves, flower petals, and fruits.) In this process, called **photosynthesis**, which will be described in Chapter 13, light energy is used to drive the synthesis of carbohydrate from CO_2. The structure of chloroplasts (**Figure 2.25**) is similar in several respects to that of mitochondria. For example, the outer membrane is highly permeable, whereas the relatively impermeable inner

membrane contains special carrier proteins that control molecular traffic into and out of the organelle. In addition, chloroplasts undergo fission.

An intricately folded internal membrane system, called the **thylakoid membrane**, is responsible for the metabolic function of chloroplasts. Chlorophyll molecules, which capture light energy during photosynthesis, are bound to thylakoid membrane proteins. Certain portions of thylakoid membrane form tightly stacked structures called **grana** (singular: granum); the entire membrane encloses a compartment known as the *thylakoid lumen*. Surrounding the thylakoid membrane is the **stroma**, a dense enzyme-filled substance, analogous to the mitochondrial matrix. In addition to enzymes, the stroma contains DNA, RNA, and ribosomes. Membrane segments that connect adjacent grana are referred to as *stroma lamellae* (singular: lamella).

Cytoskeleton

The **cytoskeleton** is a network of fibers, filaments, and associated proteins (**Figure 2.26**). Its principal components include microtubules, microfilaments, and intermediate fibers.

Microtubules, the largest constituent of the cytoskeleton (outer diameter 25 nm, inner diameter 12 nm), are girder-like hollow cylinders composed of protofilaments. Each protofilament is formed by the reversible polymerization of the protein tubulin. *Tubulin* is a dimer that consists of two polypeptides: α-tubulin and β-tubulin, which is a GTP-binding protein. Microtubules are polar; that is, their ends are different. At the plus (+) end, polymerization can occur rapidly. The minus (–) end grows more slowly. As the microtubule grows at the plus end, it extends toward the cell's periphery. Microtubule dynamics are regulated by microtubule-associated proteins (MAPs), a series of molecules that control microtubule stability by promoting or preventing the assembly process. Other functions of MAPs include guiding microtubules toward specific cellular locations and cross-linking that creates microtubule bundles. The ATP-dependent motor proteins kinesin and dynein move along microtubules. In general, kinesin moves cargo such as vesicles or organelles toward the plus end, and dynein moves toward the minus end. Although found in many cellular regions, microtubules are most prominent in long, thin structures that require support (e.g., the extended axons and dendrites of nerve cells). They are also found in the *mitotic spindle* (the structure formed in dividing cells that is responsible for the equal dispersal of chromosomes into daughter cells) and the slender, hair-like organelles of locomotion known as cilia and flagella (**Figure 2.27**).

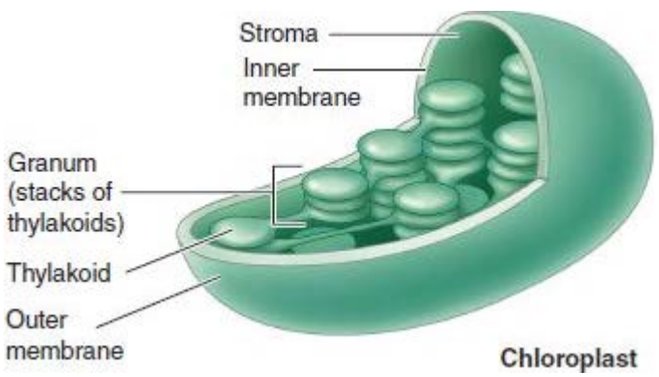


FIGURE 2.25

The Chloroplast

Chloroplasts convert light energy into the chemical bond energy of organic biomolecules.

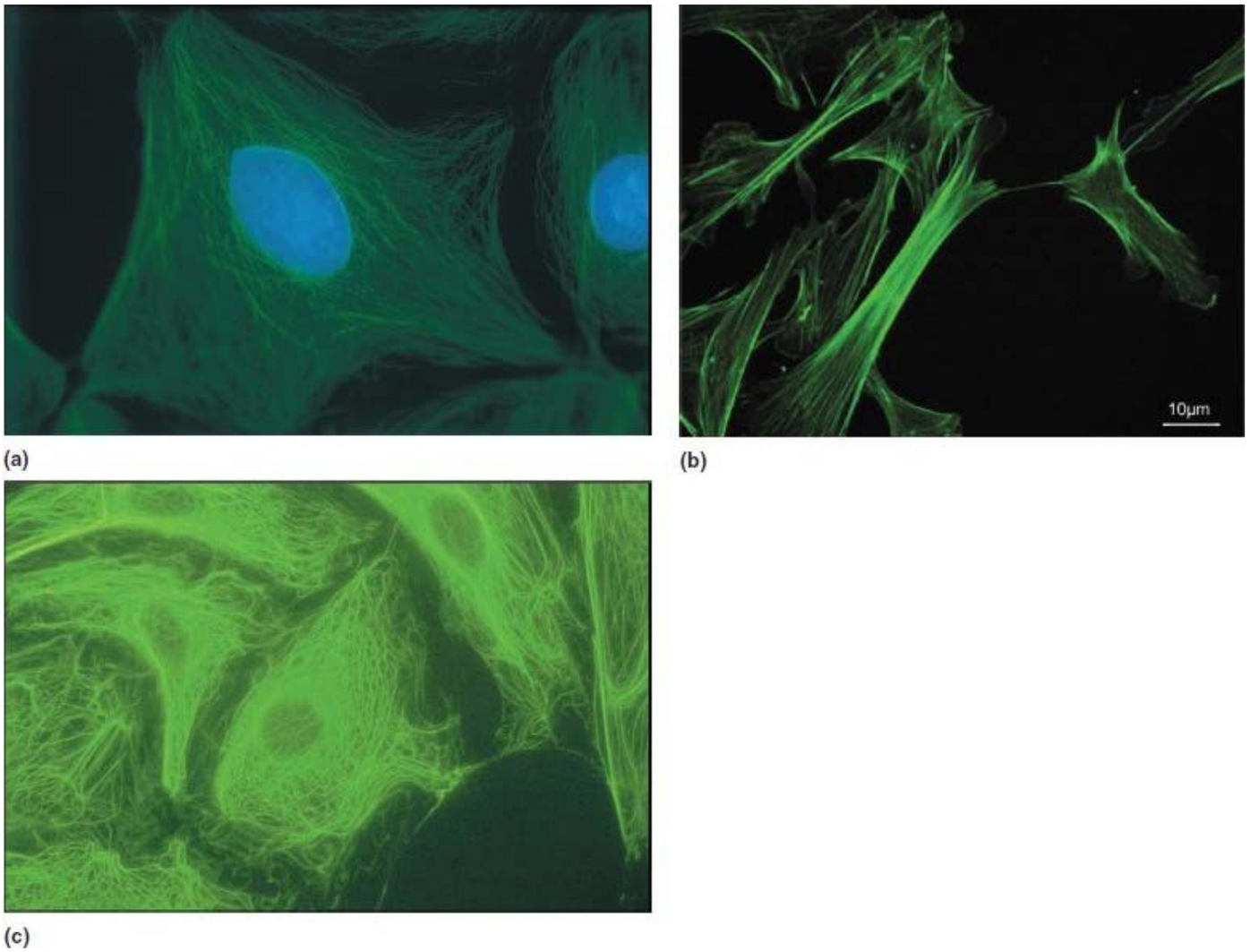


FIGURE 2.26

The Cytoskeleton

The major components of the cytoskeleton are (a) microtubules, (b) microfilaments, and (c) intermediate filaments. The intracellular distribution of each type of cytoskeletal component is visualized by staining with fluorescent dyes.

Cilia and flagella, whip-like appendages encased in plasma membrane, are highly specialized for their roles in propulsion. The most prominent examples include the motile cilia on the surface of tracheal cells that move debris-laden mucus away from the lungs and the flagellum of sperm cells that seek out egg cells. The microtubules in the inner core of flagella and cilia, referred to as the *axoneme*, form a ring of nine fused pairs with a centrally located unfused pair (a 9 + 2 pattern). The undulating motion of cilia and flagella is the result of the outer microtubule pairs sliding relative to each other. Bending occurs as the ATP-driven structural changes in the dynein molecules (called “arms”) cause them to alternately attach to and “walk along” the adjacent microtubule and then detach. The microtubules also transport cargo (e.g., newly synthesized axonemal proteins) within cilia and flagella. In a process called *intraflagellar transport* (*IFT*), kinesins move particles containing molecules required for ciliary or flagellar assembly and maintenance along the outer pair of microtubules toward the cell periphery. Dyneins move substances (e.g., kinesins that have discharged their cargo) in the opposite direction. A nonmotile version of cilia, referred to as *primary cilia*, is an important structural feature of most vertebrate

cells. Their impact on human health is described in the Biochemistry in Perspective essay Primary Cilia and Human Disease on p. 67.

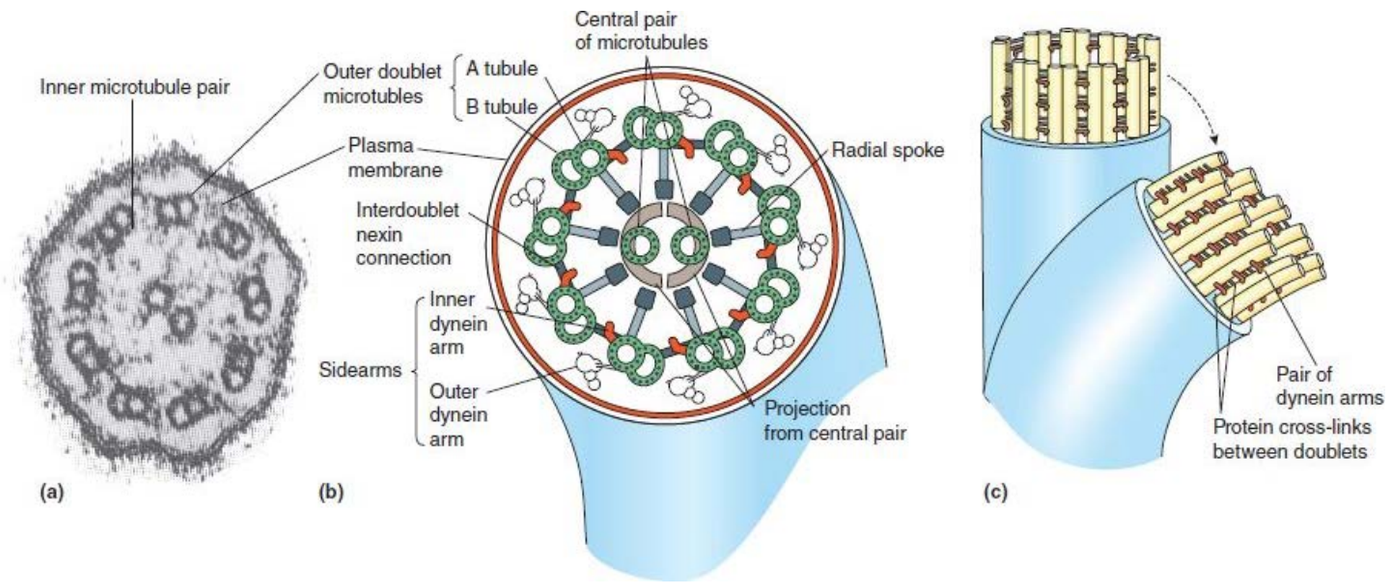


FIGURE 2.27

The Axoneme of Eukaryotic Cilia and Flagella

The axoneme is microtubule-based cytoskeletal structure. In whip-like motile cilia and flagella, the axoneme has a classic 9 × 2 + 2 microtubule pattern (i.e., nine outer pairs [doublet microtubules] and an inner microtubule pair). (a) Transmission electron micrograph of a cross section of a motile eukaryotic flagellum. (b) This diagram of an axoneme in cross section illustrates the classic axoneme pattern. The dynein arms are ATP-hydrolyzing motor proteins that "walk along" the microtubules to produce axoneme bending. The nexin linkages connect the outer doublet microtubules. Radial spokes are believed to regulate axoneme motion. (c) This diagram illustrates axoneme bending, the result of the sliding of the microtubule doublets relative to each other.

Microfilaments are small fibers (5–7 nm in diameter) composed of polymers of the globular protein actin (G-actin). The filamentous or polymeric form (F-actin) exists as a coil of two actin polymers with a plus end and a minus end. Polymerization, driven by ATP hydrolysis, occurs more rapidly at the plus end. The individual filaments, being highly flexible, are usually cross-linked into bundles of different sizes. A large variety of actin-binding proteins regulate the structure and functional properties of microfilaments. They cross-link, stabilize, sever (cut into fragments), or cap (block polymerization) microfilaments. Microfilaments can exert force simply by polymerizing or depolymerizing. Together with the myosins, a large family of ATP-dependent motor proteins, microfilaments generate contractile forces that create tension. Important roles of microfilaments include involvement in cytoplasmic streaming (a process that is most easily observed in plant cells in which cytoplasmic currents rapidly displace organelles such as chloroplasts), ameboid movement (a type of locomotion created by the formation of temporary cytoplasmic protrusions), and muscle contraction.

Intermediate filaments (8–12 nm in diameter) are a large group of flexible, strong, and relatively stable polymers. They provide cells with significant mechanical support. A network of intermediate filaments (IFs) extends from a ring-like meshwork around the nucleus to attachment points on the plasma membrane. There are six classes of IF proteins, which differ in their amino acid sequences. Well-known examples are the keratins found in skin and hair cells and the lamins that reinforce the nuclear envelope. Despite this diversity, each IF type consists of a rod-like

domain flanked by globular head and tail domains. IF polypeptides assemble into dimers (two polypeptides), tetramers (four polypeptides), and higher-order structures. IFs are especially prominent in cells that are subjected to mechanical stress.

The cytoskeleton, a dynamic mechanical system, is an integral feature of most cell activities. The unique functional properties of the cytoskeleton are made possible by a balance of mechanical forces between compression-resistant microtubules and tension generated by contractile microfilaments. IFs connect microtubules and microfilaments to each other and to the nucleus and plasma membrane. As a result of this functional "cytoarchitecture," opposing forces are continuously equilibrated throughout all cytoskeletal elements (**Figure 2.28**). Living cells are, therefore, in a constant state of dynamic instability. Cytoskeletal reorganization, triggered by a vast array of chemical and physical signals, is a principal feature of most cellular processes. For example, mechanical forces such as stretching or pressure transmitted through the extracellular matrix of cells activates a process that sequentially realigns the bionanowires of the cytoskeleton and the nuclear lamina, often resulting in changes in cellular biochemistry and gene expression.

Among the most important functions made possible by the properties of the cytoskeleton are the following.

1. **Cell shape.** Eukaryotic cells come in a vast variety of shapes including the blob-like amoeba, columnar epithelial cells, and neurons with complex branching architecture. Changes in cell shape result from responses to external signals. Amoebas, for example, rapidly change shape as they move closer to a source of nutrient molecules.
2. **Large- and small-scale cell movement**. Large-scale cellular movements are made possible by a dynamic cytoskeleton that can rapidly assemble and disassemble its structural elements according to the cell's immediate needs. Organelles are moved around within cells by being attached to cytoskeletal structures. For example, after cell division, the extension of the endoplasmic reticulum membrane from the newly formed nuclear membrane out toward the cell's periphery and the reformation of the Golgi complex are accomplished by attachment to microtubules. Movement occurs as motor proteins linked to microtubules and to the membrane cargo undergo ATP hydrolysis–dependent conformational changes.
3. **Solid state biochemistry.** Many of the biochemical reactions previously believed to occur within the liquid phase of the cytoplasm instead proceed in large measure on a cytoskeletal platform. Biochemical pathways are both more efficient and more easily controlled when enzymes assemble into complexes on a solid surface. Prominent examples are the reactions of glycolysis, an ATP-generating pathway in carbohydrate metabolism. The binding of glycolytic enzymes to cytoskeletal filaments has been observed to vastly increase reaction rates. Drugs that disrupt cytoskeletal structure cause glycolytic enzyme detachment and a rapid decrease in cytoplasmic ATP production.

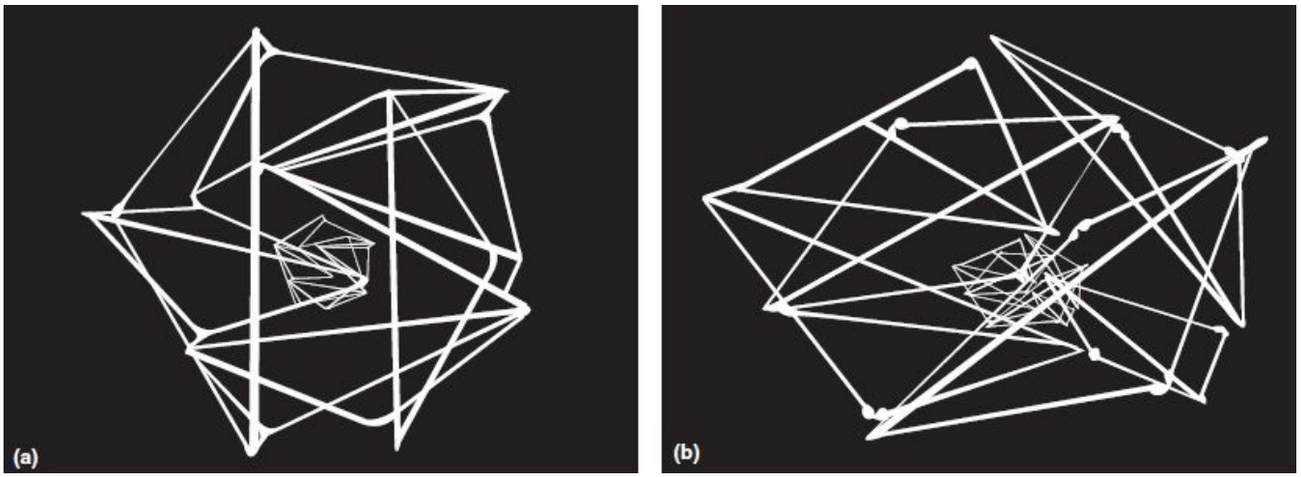


FIGURE 2.28

Cytoskeletal Reorganization Model

In this tensegrity model, both (a) and (b) are stable structures that are held together by balanced mechanical stresses, namely, tensional strings and rigid struts. The "cell" here is composed of aluminum struts ("intermediate filaments") and thin elastic strings ("microfilaments"); the "nucleus," a geodesic sphere, is constructed of wooden sticks and white elastic thread. When an external force is applied to structure (a), it rearranges to form structure (b).

4. **Signal transduction.** Cells are information-processing systems that respond to a wide range of external chemical and physical stimuli. Examples include the binding of hormones and growth factors to cell-surface receptors, action potentials in nerve and muscle cell membranes, and mechanical forces such as tensional force and hydrostatic pressure. Cells possess a constellation of signal transduction mechanisms (voltage- and stretch-sensitive ion channels, signal complexes, biochemical pathways, and gene expression devices) that can be thought of as resembling the *integrated circuits* (microchips) in computers: information-processing devices composed of transistors and capacitors, connected by wires, and driven by electricity. It is the filaments of the cytoskeleton and nucleoskeleton that facilitate and support signal transduction processes.

KEY CONCEPTS

The cytoskeleton, a highly structured and adaptable network of proteinaceous filaments, is responsible for maintenance of cell shape, large- and small-scale cell movement, solid state biochemistry, and signal transduction.

QUESTION 2.4

Cancer is a group of diseases characterized by unregulated cell division. Taxol, a drug used to treat ovarian cancer, attaches to and stabilizes microtubules. Briefly, what is the basis of Taxol's anticancer action?

Primary Cilia and Human Disease

What effects do nonmotile cilia have on human health? Most differentiated vertebrate cells possess a single nonmotile cilium called the **primary cilium**. In contrast to motile cilia, primary cilia lack the central microtubule pair within the axoneme (a 9 + 0 pattern) and the dynein arms and radial spokes that are required for motility. The primary cilium functions, instead, as a sensory organelle; that is, it acts as a cellular antenna. A large number of receptor molecules and other proteins embedded in the ciliary membrane facilitate the sensing of environmental cues, such as mechanical pressure, signal molecules, and light. For example, the primary cilium of kidney tubule cells protrudes into the tubular lumen, where it senses the flow of urine. Mechanical bending of the cilium caused by urine flow results in the inward flow of Ca^{2+}. One consequence of inward calcium flow is cell division suppression. Other examples of primary cilia functions include wound healing (fibroblasts migrate toward a wound when primary cilium plasma membrane receptors bind to platelet-derived growth factor [PDGF], p. 603); olfaction (the primary cilia of olfactory sensory neurons detect odorant molecules); and sight (the outer segment of rod cells in the retina is essentially a highly modified primary cilium with an enlarged tip packed with visual pigments).

The confined space within a primary cilium enables the tight integration of several signaling systems. Examples include the hedgehog and Wnt signaling pathways, both of which have important roles in animal development. Efficient transport by intraflagellar transport (IFT) of ciliary components and intermediate signal molecules back and forth between the cilium and the cytoplasm also facilitates signal transduction. Signal molecules transported by IFT ultimately result in gene expression changes within the nucleus. A number of human diseases, referred to as *ciliopathies*, are attributed to defects in primary cilia.

Human Ciliopathies

Considering that primary cilia are present on most cells in the human body, the wide spectrum of primary cilia-related human diseases is not surprising. Some ciliopathies appear to affect one or a small number of cell types or organs, whereas others affect many of the body's systems and can vary greatly in their symptoms. Retinitis pigmentosa (RP) is a group of more than 30 different progressive genetic eye disorders that lead to blindness. One form of RP is caused by a defective version of RP1, a gene that codes for a MAP in the outer segment of rod cells. In polycystic kidney disease, loss of function in the kidneys and several other organs is caused by cyst formation that is linked to defects in either of two genes that code for the primary cilium proteins polycystin 1 (PC1) and polycystin 2 (PC2). Together PC1 and PC2 (a cation channel) act as a mechanoreceptor that monitors fluid flow in the kidney tubule cells. When this function is disrupted by mutation of either protein, cell division, which is controlled in part by primary cilium function, is stimulated. The resulting increased cell division leads to the formation of thousands of cysts (fluid-filled sacs) and eventually causes kidney failure.

Bardet–Biedl syndrome (BBS) is an example of a pleiotropic disease (a condition in which a genetic defect results in numerous and seemingly unrelated symptoms). In BBS, which can be caused by mutations in any of 12 genes, retinal degeneration and kidney and liver cysts occur in

addition to an array of clinical symptoms that include several of the following: obesity, hearing loss, olfactory deficits, diabetes, mental retardation, polydactyly (extra fingers or toes on either or both hands or feet), and situs inversus (left-to-right reversal of the internal organs). The genes linked to BBS form the BBSome, a protein complex that plays an important role in IFT in primary cilia.

SUMMARY Nonmotile primary cilia have vital roles in the health of vertebrate cells. Defects in primary cilia result in numerous human diseases.

Biochemistry IN THE LAB

Cell Technology

During the past 50 years, our understanding of the functioning of living organisms has undergone a revolution. Much of our current knowledge of biochemical processes is a direct result of technological innovations. Four of the most important cellular techniques used in biochemical research are briefly described: cell fractionation, electron microscopy, autoradiography, and live cell imaging.

Cell Fractionation

Cell fractionation techniques (**Figure 2A**) allow the study of organelles in a relatively intact form outside of cells. For example, functioning mitochondria can be used to study cellular energy generation. In these techniques, cells are gently disrupted and separated into several organelle-containing fractions.

Cells may be disrupted by several methods, but homogenization is the most commonly used. In this process, a cell suspension is placed either in a glass tube fitted with a specially designed glass pestle or into an electric blender. The resulting homogenate is separated into several fractions by means of **differential centrifugation**. In this procedure, a refrigerated instrument, the *ultracentrifuge*, generates enormous centrifugal forces that separate cell components on the basis of size, surface area, and relative density. (Forces as large as 500,000 times the force of gravity, or 500,000 *g*, can be generated in unbreakable test tubes placed in the rotor of an ultracentrifuge.) The homogenate is first spun in the ultracentrifuge at low speed (700–1000 *g*) for 10 to 20 minutes. The heavier particles, such as the nuclei, form a sediment, or *pellet*. Lighter particles, such as mitochondria and lysosomes, remain suspended in the *supernatant*, the liquid above the pellet. The supernatant is then transferred to another centrifuge tube and spun at a higher speed (15,000–20,000 *g*) for 10 to 20 minutes. The resulting pellet contains mitochondria, lysosomes, and peroxisomes. The supernatant, which contains **microsomes** (small closed vesicles formed from ER during homogenization), is transferred to another tube and spun at 100,000 *g* for 1 to 2 hours. Microsomes are deposited in the pellet, and the supernatant contains ribosomes, various cellular membranes, and granules such as glycogen, a carbohydrate polymer. After this latest supernatant

has been recentrifuged at 200,000 *g* for 2 to 3 hours, ribosomes and large macromolecules are recovered from the pellet.

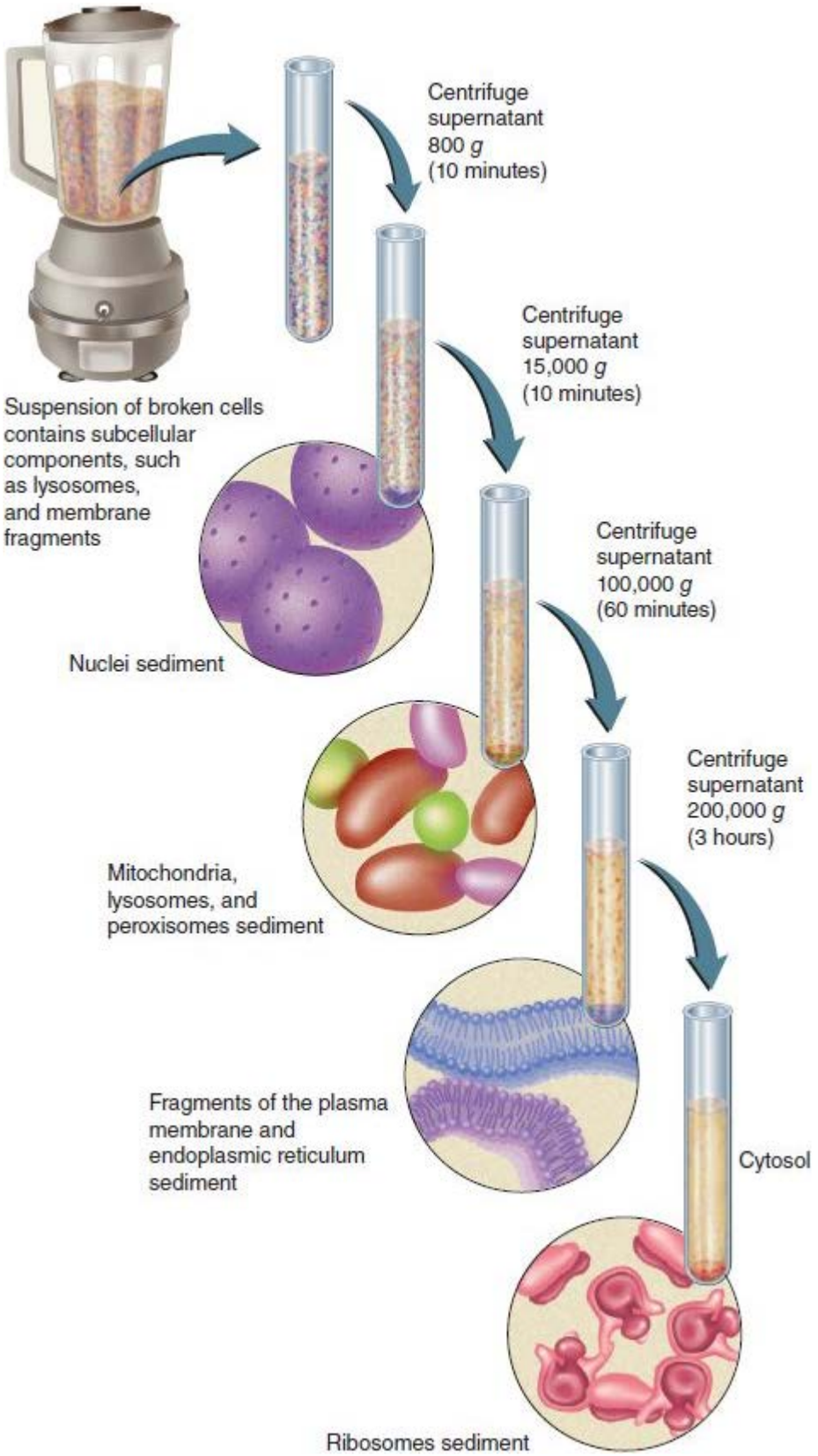


FIGURE 2A

Cell Fractionation

After homogenization of cells in a blender, cell components are separated in a series of centrifugations at

increasing speeds. As each centrifugation ends, the supernatant is removed, placed into a new centrifuge tube, and then subjected to greater centrifugal force. The collected pellet can be resuspended in liquid and examined by microscopy or biochemical tests.

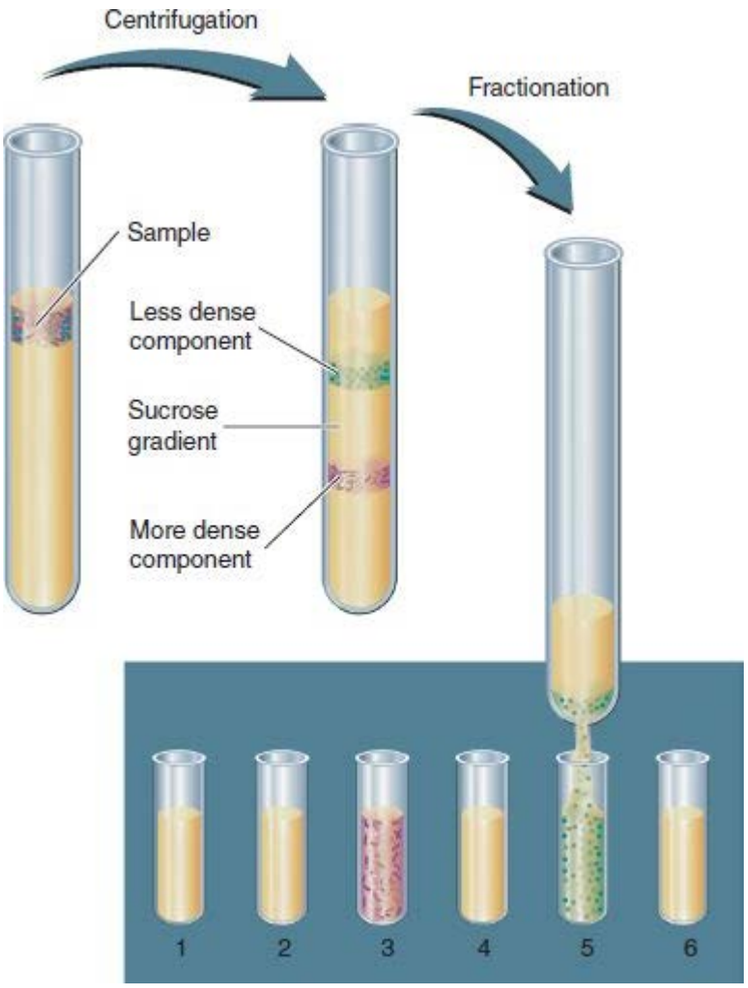


FIGURE 2B

Density-Gradient Centrifugation

The sample is gently layered onto the top of a preformed gradient of an inert substance such as sucrose. As centrifugal force is applied, particles in the sample migrate through the gradient bands according to their densities. After centrifugation, the bottom of the tube is punctured and the individual bands are collected in separate tubes.

Often, the organelle fractions obtained with this technique are not sufficiently pure for research purposes. One method often employed to further purify cell fractions is **density- gradient centrifugation** (**Figure 2B**). In this procedure, the fraction of interest is layered on top of a solution that consists of a dense substance such as sucrose. (In the centrifuge tube containing the solution, the sucrose concentration increases from the top to the bottom to form a gradient.) During centrifugation at high speed for several hours, particles move downward in the gradient until they reach a level that has a density equal to their own. The plastic centrifuge tube is then punctured, and the cell components are collected in drops from the bottom. The purity of the individual fractions can be assessed by visual inspection (electron microscopy). However, assays for **marker enzymes** (enzymes that are known to be present in especially high concentration in specific

organelles) are more commonly used. For example, glucose-6-phosphatase, the enzyme responsible for converting glucose-6-phosphate to glucose in the liver, is a marker for liver microsomes. DNA polymerase, the enzyme involved in DNA synthesis, is a marker for nuclei.

Electron Microscopy

The electron microscope (EM) permits a view of cell ultrastructure that is not possible with the more commonly available light microscope. Direct magnifications as high as 1,000,000× have been obtained with the EM. Electron micrographs may be enlarged photographically to 10,000,000×. The light microscope, in contrast, magnifies an image to about 1,000×. The greater resolving power of EM (0.5 nm) compared with that of the light microscope (0.2 μm) is a result of the wavelength sizes of their illumination sources. EM uses a stream of electrons, which have much shorter wavelengths than those of visible light. As a result, more detailed images can be obtained. In general, shorter wavelengths allow greater resolution.

There are two types of EM. In transmission electron microscopy (TEM) electrons pass through thin specimens. Images are formed because of variations in the absorption of electrons by the specimen. Scanning electron microscopy (SEM) is used to form three-dimensional images by detecting electrons emitted from specimen surfaces coated with a thin layer of a heavy metal. Although only surface features can be examined with the SEM, this form of microscopy provides very useful information about cell structure and function.

Autoradiography

Autoradiography is used to study the intracellular location and behavior of cellular components. It has been an invaluable tool in biochemistry. Radioactively labeled molecules have been used in the investigation of nucleic acid and protein synthesis, gene expression, signal transduction, and metabolic pathways. Commonly used radioisotopes include ^{3}H (tritium), ^{32}P, and ^{35}S. The tritiated nucleotide thymidine, for example, is used to study DNA synthesis because thymidine is incorporated only into DNA molecules. After exposure to the radioactive precursor, the cells are processed for light or electron microscopy. The resulting slides are then dipped in photographic emulsion. After storage in the dark, the emulsion is developed by standard photographic techniques. The location of radioactively labeled molecules is indicated by the developed pattern of silver grains.

Live Cell Imaging

The dynamic activities of living cells are best observed by live cell imaging using light microscopy. Two examples are phase contrast microscopy and fluorescence microscopy.

Phase contrast microscopy takes advantage of variations in the refraction of light as it passes through substances with different densities. Phase contrast microscopes are fitted with an annulus aperture (which limits the angle of the incoming light rays) and a phase plate (which contains a phase ring that shifts light wavelengths (λ) along the horizontal axis). Although living cells are translucent, a phase shift of 90° relative to the background light results in sufficient contrast that cell structures can be observed. Phase contrast microscopy is used in investigations such as chemotaxis when low-resolution level is sufficient. (Chemotaxis is the movement of an organism such as an amoeba toward or away from a chemical stimulus.)

Cell biologists and biochemists use *fluorescence microscopy* to investigate cell function with *fluorophores* (molecules that absorb photons of light and then reemit photons of lower energy). (Fluorescence is described on p. 502.) Fluorophores can be small molecules such as DAPI, TRITC, and FITC or fluorescent proteins such as green fluorescent protein (GFP) linked to a protein of interest. Fluorophores are used in the time-lapse investigation of a wide variety of cell functions, including signal transduction mechanisms.

A fluorescence microscope is fitted with a light source, an excitation filter (which transmits only the wavelengths that excite a specific fluorophore), a dichroic mirror (a beamsplitter composed of coated glass that reflects the excitation light and transmits the emitted fluorescence), an emission filter (which transmits peak emission light waves), and a camera.

Several fluorophores can be used simultaneously. In **Figure 2C,** cell nuclei are stained blue with DAPI, a fluorescent molecule that binds to DNA; actin filaments appear red because they are labeled with TRITC bound to the actin filament–specific binding protein phalloidin; and microtubules are green because they are bound to antibody proteins linked to FITC.

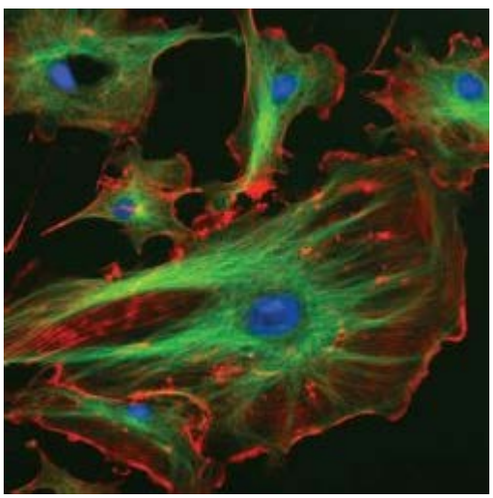


FIGURE 2C

Fluorescence Micrograph of an Arterial Endothelial Cell

Nuclei are stained blue because the fluorophore DAPI binds to DNA; actin filaments appear red because they are labeled with TRITC; and microtubules are green because microtubule-specific antibodies are linked to FITC.

Note that the use of fluorescence microscopy in the observation of living cells is limited to short periods of time by the toxic nature of fluorescent stains.

Chapter Summary

1. Cells are the structural units of all living organisms. Within each living cell are hundreds of millions of densely packed biomolecules. The unique chemical and physical properties of water are a crucial determining factor in the behavior of all other biomolecules. Biological membranes are thin, flexible, and relatively stable sheet-like structures that enclose cells and organelles. They are formed from biomolecules such as phospholipids and proteins that together form a selective physical barrier.
2. Self-assembly of supramolecular structures occurs within living cells because of the steric information encoded into the intricate shapes of biomolecules that allows numerous weak, noncovalent interactions

between complementary surfaces. Many of the multisubunit complexes involved in cellular processes are now known to function as molecular machines; that is, they are mechanical devices composed of moving parts that convert energy into directed motion. Macromolecular crowding, created by the density of proteins within the cell, is an important factor in the wide variety of cellular phenomena. Signal transduction mechanisms allow cells to process internal and external information. Proteostasis (protein homeostasis) exists in cells as the result of a dynamic equilibrium between protein synthesis and folding and protein degradation.

3. All currently existing organisms contain either prokaryotic or eukaryotic cells. Prokaryotes are simpler in structure than eukaryotes. They also have vast biochemical diversity across species lines because almost any organic molecule can be used as a food source by some species of prokaryote. Unlike the prokaryotes, the eukaryotes carry out their metabolic functions in membrane-bound compartments called organelles.
4. DNA molecules in prokaryotic cells are located in an irregularly shaped region called the nucleoid. Many bacteria contain additional small circular DNA molecules called plasmids. Plasmids may carry genes for special function proteins that provide protection, metabolic specialization, or reproductive advantages to the organism.
5. The plasma membrane of both prokaryotes and eukaryotes performs several vital functions. The most important of these functions is controlled molecular transport, which is facilitated by carrier and channel proteins.
6. The ER is a system of interconnected membranous tubules, vesicles, and large flattened sacs found in eukaryotic cells. There are two forms of ER. The RER, which is primarily involved in protein synthesis, is so named because of the numerous ribosomes that stud its cytoplasmic surface. The second form lacks attached ribosomes and is called SER. Functions of the SER include lipid synthesis and biotransformation.
7. Formed from relatively large, flattened, sac-like membranous vesicles that resemble a stack of plates, the Golgi apparatus is involved in the modification, packaging, and release of cell products into the vesicular compartment for delivery to target locations in the cell.
8. The cell contains a system of vesicular organelles involved in processing both endogenous and exogenous materials around, into, and out of the cell and performing specialized biochemical functions.
9. The nucleus of any eukaryote contains DNA, the cell's genetic information. Ribosomal RNA is synthesized in the nucleolus, found within the nucleus. Separating DNA replication and transcription processes from the cytoplasm is the nuclear envelope; it is composed of two membranes that fuse at structures called the nuclear pores.
10. Aerobic respiration, a process by which cells use oxygen to generate energy, takes place in mitochondria. Each mitochondrion has two membranes. The smooth outer membrane is permeable to most molecules with masses less than 10,000 Da. The inner membrane, which is impermeable to ions and a variety of organic molecules, projects inward into folds that are called cristae. Embedded in this membrane are mitochondrial respiratory chains, molecular complexes that are responsible for the synthesis of ATP.
11. Peroxisomes are small spherical membranous organelles that contain a variety of oxidative enzymes. These organelles are most noted for their involvement in the generation and breakdown of peroxides.
12. Chromoplasts accumulate the pigments that are responsible for the color of leaves, flower petals, and fruits. Chloroplasts are a type of chromoplast that is specialized to convert light energy into chemical energy.
13. The cytoskeleton, a supportive network of fibers and filaments, is involved in the maintenance of cell shape, large-and small-scale cellular movement, solid-state biochemistry, and signal transduction.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on water to help you prepare for exams.

Chapter 2 Review Quiz

Suggested Readings

Clemente JC, et al. 2012. The impact of the gut microbiota on human health: an integrative view. Cell 148(6):1258–70.

Conion MA, Bird AR. 2014. The impact of diet and lifestyle on gut microbiota and human health. Nutrients 7:17–44.

Goodsell DS. 2009. The machinery of life, 2nd ed. New York (NY): Springer Verlag.

Hamilton G. 2015. The mitochondrial mystery. Nature 525:444–6.

Ingber DE. 1998. The architecture of life. Sci Am 278(1):48–57.

Ingber DE, Wang N, Stamenovic D. 2014. Tenegrity, cellular biophysics, and the mechanics of living systems. Rep Prog Phys 77(4):046603.

Lane N. 2005. Power, sex, suicide: mitochondria and the meaning of life. New York (NY): Oxford.

Madhuparna R, et al. 2015. Mitochondrial division and fusion in metabolism. Curr Opin Cell Biol 33:111–18.

Meyerovich K, et al. 2016. Endoplasmic reticulum stress and the unfolded protein response in pancreatic islet inflammation. J Mol Endocrin 57:R1–R17.

Pedersen LB, et al. 2016. Endocytic control of cellular signaling at the primary cilium. Trends Biochem Sci 41(9):784–97.

Powers ET, Balch WE. 2013. Diversity in the origins of proteostasis networks—a driver for protein function in evolution. Nat Rev Mol Cell Biol 14(4):237–48.

Satir P, Pederson LB, Christensen ST. 2010. The primary cilia at a glance. J Cell Sci 123:499–503.

Suez J, et al. 2014. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. Nature 514:181–6.

Villasenor R, et al. 2016. Signal processing by the endosomal system. Curr Opin Cell Biol 39:53–60.

Key Words

aerobic metabolism, *58*

apoptosis, *58*

biotransformation reaction, *49*

carrier protein, *37*

cell fractionation, *68*

channel protein, *70*

chloroplast, *62*

chromatin fiber, *55*

chromosome, *44*

chromosome territory, *55*

clathrin, *52*
clathrin-dependent endocytosis, *52*
cytoskeleton, *62*
density-gradient centrifugation, *69*
differential centrifugation, *68*
endocytic cycle, *52*
endocytosis, *52*
endomembrane system, *45*
endoplasmic reticulum (ER), *48*
ER-associated protein degradation, *49*
ER stress, *49*
exocytosis, *51*
extracellular matrix (ECM), *45*
glycocalyx, *45*
Golgi apparatus (Golgi complex), *49*
granum, *62*
hormone, *40*
hydrophilic, *36*
hydrophobic, *37*
inner mitochondrial membrane, *59*
inner nuclear membrane, *56*
integral protein, *37*
intermediate filament, *65*
ligand, *40*
lysosome, *52*
marker enzyme, *69*
membrane skeleton, *45*
microbiota, *34*
microfilament, *64*
microsome, *68*
microtubule, *62*
mitochondrion, *59*
motor protein, *38*
neurotransmitter, *40*
nuclear envelope, *56*
nuclear lamina, *57*
nuclear matrix, *58*
nuclear pore complexes, *56*
nucleoid, *44*
nucleolus, *58*
nucleoplasm, *55*
nucleus, *55*

organelle, *36*
outer mitochondrial membrane, *59*
outer nuclear membrane, *56*
perinuclear space, *56*
peripheral protein, *37*
peroxisome, *61*
photosynthesis, *43*
plasma membrane, *43*
plasmid, *44*
primary cilium, *67*
proteome, *39*
proteostasis, *39*
proteostasis network, *39*
receptor, *37*
respiration, *43*
rough ER (RER), *48*
signaling cascade, *40*
signal transduction, *40*
smooth ER (SER), *48*
stroma, *62*
thylakoid membrane, *62*
unfolded protein response, *49*
vesicles, *45*

Review Questions

SECTION 2.1

Comprehension Questions

1. Define the following terms:
    a. hydrophillic
    b. hydrophobic
    c. lipid bilayer
    d. integral protein
    e. peripheral protein
2. Define the following terms:
    a. channel protein
    b. carrier protein
    c. receptor
    d. prokaryote
    e. eukaryote

3. Define the following terms:
    a. polar head group
    b. hydrophobic tail
    c. ligand
    d. motor protein
    e. GTP
4. Define the following terms:
    a. ribosome
    b. macromolecular crowding
    c. excluded volume
    d. signal transduction
    e. hormone
5. Define the following terms:
    a. cytokine
    b. lipopolysaccharide
    c. proteostasis
    d. proteostasis network
    e. signal cascade
6. Define the following terms:
    a. proteotoxic stress
    b. proteosome
    c. proteome
    d. ubiquitin-proteosome system
    e. autophagy

Fill in the Blanks

7. Techniques that the body uses to protect itself from the microbe members of the human superorganism are impenetrable tissue barriers and _____ system cells.
8. _____ compounds exclude water.
9. The two types of membrane proteins are peripheral and _____.
10. Living organisms require both information and _____ to create order.
11. The four phases of signal transduction are _____, _____, _____, and _____.

Short-Answer Questions

12. Describe the functional properties of motor protein subunits in molecular machines.
13. Describe the properties of molecules that self-assemble into a supermolecular machine.
14. Approximately 70% of the body's immune system cells are located in or near the wall of the lower intestinal tract (intestines). Can you suggest a reason for this phenomenon?
15. What factors can promote protein misfolding in cells?
16. Explain why the term *crowded* rather than *concentrated* is used to describe the densely packed molecules in the interior of living cells.
17. Describe the four phases of signal transduction in living organisms.

18. Why is it difficult for the colon to reestablish a beneficial flora after several antibiotic treatments?
19. Describe the causes of proteolytic stress in a living cell.
20. Give three examples of human diseases in which there are proteostasis deficiencies.
21. Compare the functions of neurotransmitters, hormones, and cytokines.
22. Provide several examples of the consequences of signaling cascades in living cells.
23. Provide examples of biochemical phenomena that are directly affected by macromolecular crowding.

Critical-Thinking Questions

24. Describe the connection between antibiotic use and inflammatory bowel disease.
25. Describe how cytoplasmic calcium ions are involved in insulin secretion.

SECTION 2.2

Comprehension Questions

26. Define the following terms:
    a. plasma membrane
    b. nucleoid
    c. conjugation
    d. bacterial capsule
    e. biofilm
27. Define the following terms:
    a. endotoxin
    b. periplasmic space
    c. slime layer
    d. lipopolysaccharide
    e. pilus
28. Define the following terms:
    a. flagellum
    b. photosynthesis
    c. respiration
    d. capsule
    e. polyphosphate inclusions

Fill in the Blanks

29. There are two types of prokaryotes: _____ and _____.
30. _____ is the biochemical conversion of light energy into chemical energy.
31. The three types of external bacterial appearances are _____, _____, and _____.
32. The _____ is the region between the outer membrane and inner membrane of Gram-negative bacteria.

33. _____ is a polymeric network in the prokaryotic cell wall in which short peptide chains are linked to carbohydrate chains.
34. Examples of inclusion bodies in bacterial cytoplasm are _____ and _____ .

Short-Answer Questions

35. Describe why bacterial biofilms can be a threat to human health.
36. Draw a diagram of a bacterial cell. Label and explain the function of each of the following components: nucleoid, plasmid, cell wall, pili, and flagella.
37. Why does soap kill bacteria?
38. Describe the functional properties of bacterial pili and flagella.
39. Antibiotic resistance can be transferred from one bacterial cell to another. Describe how this process can occur.
40. Why is lipopolysaccharide considered to be an endotoxin?
41. Describe the possible functions of proteins in the bacterial plasma membrane.
42. What property of biofilms contributes to their impact on human medical conditions?
43. What is the major structural difference between prokaryotic and eukaryotic cells?

Critical-Thinking Questions

44. Several pathogenic bacteria (e.g., *Bacillus anthracis*, the cause of anthrax) produce an outermost mucoid layer called a capsule. Capsules can be composed of polysaccharide or protein. What effect do you think this "coat" would have on a bacterium's interactions with a host animal's immune system?
45. Mycoplasmas are unusual bacteria that lack cell walls. With a diameter of 0.3 μm, they are believed to be the smallest known free-living organisms. Some species are pathogenic to humans. For example, *Mycoplasma pneumoniae* causes a serious form of pneumonia. Assuming that mycoplasmas are spherical, calculate the volume of an individual cell. Compare the volume of a mycoplasma with that of *E. coli*.
46. The dimensions of prokaryotic ribosomes are approximately 14 nm by 20 nm. If ribosomes occupy 20% of the volume of a bacterial cell, calculate how many ribosomes are in a typical cell as *E. coli*. Assume that the shape of a ribosome is approximately that of a cylinder.

SECTION 2.3

Comprehension Questions

47. Define the following terms:

    a. glycocalyx
    b. cell cortex
    c. endomembrane system
    d. vesicle
    e. endoplasmic reticulum

48. Define the following terms:

    a. ER lumen
    b. rough ER

c. smooth ER
d. clathrin
e. Golgi apparatus

49. Define the following terms:

a. biotransformation reaction
b. unfolded protein response
c. exocytosis
d. endocytosis
e. nucleoplasm

50. Define the following terms:

a. chromatin
b. nuclear matrix
c. nucleolus
d. nuclear envelope
e. nuclear pore complex

51. Define the following terms:

a. acid hydrolase
b. mitochondrion
c. lysosome
d. autophagy
e. peroxisome

52. Define the following terms:

a. photosynthesis
b. thylakoid membrane
c. MAP
d. IFT
e. polycystic kidney disease

53. Define the following terms:

a. F-actin
b. G-actin
c. intermediate filament
d. ciliopathy
e. anterograde transport

Fill in the Blanks

54. Mitochondrial fission involves an interaction with a _____, which is mediated by a protein tether called _____.
55. Mitochondria form stable contact sites with regions of the ER called _____.
56. _____ are chromoplasts that convert light energy into chemical energy.
57. _____ is a protein meshwork attached to the inner surface of the inner nuclear membrane.
58. The inner and outer nuclear membranes fuse at structures called _____.

Short-Answer Questions

59. What functions does the cytoskeleton perform in eukaryotic cells?
60. Describe the functions of the rough and smooth endoplasmic reticulum.
61. Describe the functions of the Golgi apparatus.
62. List three environmental signals detected by primary cilia.
63. Describe the functions of peroxisomes. How is this type of organelle formed?
64. What are the functions of COPII in eukaryotic cells?
65. Distinguish among the terms ER stress, unfolded protein response, and ER-associated protein degradation.
66. Suggest a reason why eukaryotic cells are so much larger than prokaryotic cells.
67. Describe the structural and functional properties of the nuclear membranes.
68. What is proteostasis? How is it important in the life of cells?
69. What are the components of the endomembrane system? How are these components functionally connected?
70. Outline the role of the cytoskeleton in intracellular signal transduction.

Critical-Thinking Questions

71. The *E. coli* cell is 2 μm long and 1 μm in diameter, whereas a typical eukaryotic cell is 20 μm in diameter. Assuming that the *E. coli* cell is a perfect cylinder and the eukaryotic cell is a perfect sphere, calculate the surface-to-volume ratio for each cell type [cylinder volume, $V = pr^2h$; cylinder area $A = 2pr^2 + 2prh$; sphere volume, $V = 4/3(pr^3)$; sphere area, $A = 4pr^2$]. What do these numbers tell you about the evolutionary changes that would have to occur to generate an efficient eukaryotic cell, considering that many biochemical processes depend on transport across membranes?
72. Familial hypercholesterolemia (FH) is an inherited disease characterized by high blood levels of cholesterol, xanthomas (lipid-laden nodules that develop under the skin near tendons), and early-onset atherosclerosis (the formation of yellowish plaques within arteries). In the milder form of this disease, patients have half the plasma membrane low-density lipoprotein (LDL) receptors needed for cells to bind to and internalize LDL (a plasma lipoprotein particle that transports cholesterol and other lipids to tissues). These individuals have their first heart attacks in young adulthood. In the severe form of FH, in which affected individuals have no functional LDL receptors, heart attacks begin at about age 8, with death occurring a few years later. Based on what you have learned in this chapter, briefly describe the cellular processes that are defective in FH.
73. Cyst formation causes a catastrophic loss of function in polycystic kidney disease. Genetic research has linked this disease to defects in genes that code for primary cilium proteins. Describe in general terms how malfunctioning primary cilia cause the formation of kidney cysts.
74. In addition to providing support, the cytoskeleton immobilizes enzymes and moves organelles to suitable positions in the cytoplasm. What advantage do these circumstances have over allowing the cell contents to freely diffuse in the cytoplasm?
75. Primary cilia have evolved as primary sensory organelles for vertebrate cells. What structural features of these cilia make them ideal for this purpose?

MCAT Study Questions

76. Which of the following is a component of the cytoskeleton?

    a. collagen
    b. dynein
    c. elastin
    d. chromatin

77. A marker enzyme used to identify liver microsomes (ER fragments) after centrifugation is

    a. AMPK
    b. phosphodiesterase
    c. peroxidase
    d. glucose-6-phosphatase

78. Which of the following statements about the nucleolus is **not** true?

    a. The nucleolus is a membrane-bound structure.
    b. rRNA genes are transcribed in the nucleolus.
    c. Ribosomal subunits are synthesized in the nucleolus.
    d. rRNA molecules are chemically modified in the nucleolus.

79. Which of the following statements about insulin is **not** true?

    a. The insulin receptor has tyrosine kinase activity.
    b. Insulin increases fat synthesis in adipose cells.
    c. Insulin increases glycogen synthesis
    d. Insulin increases gluconeogenesis reactions.

80. Proteins in the inner mitochondrial membrane are involved in which of the following processes?

    a. nucleotide synthesis
    b. OXPHOS
    c. autophagy
    d. NADH synthesis

CHAPTER 3 Water: The Matrix of Life

The Water Planet Unique among the planets in the solar system, the Earth is an oceanic world. Water's properties make life on Earth possible.

OUTLINE

WATER, WATER, EVERYWHERE

3.1 MOLECULAR STRUCTURE OF WATER

3.2 NONCOVALENT BONDING

Ionic Interactions

Hydrogen Bonds
Van der Waals Forces

3.3 THERMAL PROPERTIES OF WATER

3.4 SOLVENT PROPERTIES OF WATER
Hydrophilic Molecules, Cell Water Structuring, and Sol-Gel Transitions
Hydrophobic Molecules and the Hydrophobic Effect
Amphipathic Molecules
Osmotic Pressure

3.5 IONIZATION OF WATER
Acids, Bases, and pH
Buffers
Physiological Buffers

Biochemistry in Perspective
Cell Volume Regulation and Metabolism

AVAILABLE ONLINE

Biochemistry in Perspective
Water, Abiotic Stress, and Compatible Solutes

Biochemistry in the Lab
Dialysis

Water, Water, Everywhere

> Water, water, everywhere,
> And all the boards did shrink;
> Water, water, everywhere,
> Nor any drop to drink.
>
> Samuel Taylor Coleridge (1772–1834), *The Rime of the Ancient Mariner*

Coleridge's mariner laments the fate (dying of thirst) he and his shipmates face as their becalmed wooden sailing ship is surrounded by an endless vista of undrinkable ocean water. The plight of these fictional characters is also our own. We are threatened by our dependence on clean, unsalted water to maintain our health when usable, accessible water resources are vanishingly small. Our bodies are, on average, approximately 60% water because it is the principal component of every body fluid (e.g., blood, saliva, lymph, and gastric juices). Among water's numerous roles are nutrient absorption and transport, waste product excretion, body temperature regulation, and joint lubrication. Our requirements for water are so stringent that even mild dehydration causes tiredness, headaches, and loss of concentration. However, of an estimated total world water supply of 366×10^{18} gallons, approximately 97% (355×10^{18} gallons) is ocean water. Only 3% of the world's water is fresh (drinkable), and most of that is locked away in glaciers and polar icecaps. Together, fresh water in rivers, lakes, underground aquifers, and the atmosphere accounts for less than 1% of the world's water!

In the 10,000 years since the Neolithic agricultural revolution, humans have usually settled near convenient sources of water (dependable rainfall and/or large rivers) so that agriculture could flourish. The first civilizations in the West were established between the Tigris and Euphrates rivers in

Mesopotamia and the Nile River in Egypt. In Asia, the Yellow River basin has been described as the "cradle of Chinese civilization." As human populations in towns and then cities increased, survival depended on water resource management. In ancient Rome by the third century CE, 11 aqueducts provided sufficient water for an estimated 1 million people. Inadequate water results in catastrophe. The Old Kingdom in Egypt in the third millennium BCE built the great pyramids, but it collapsed in the midst of a 300-year drought.

The Mayan civilization, a series of city-states in the Yucatan peninsula, is an exceptionally well-researched example of the effect of water availability on human survival because of climate data (temperature, volcanic eruptions, forest fires, and precipitation) obtained from Greenland ice cores. Between 250 and 950 CE, the Mayans developed a culture with a written language and astronomical, mathematical, and architectural achievements that rivaled those of other contemporary civilizations (**Figure 3.1**). Despite these magnificent accomplishments, beginning around 750 CE and lasting until about 950 CE, Mayan cities began to decline and were eventually abandoned. Despite sophisticated rain collection and storage facilities and irrigation systems, the Mayan population (at its highest about 13 million) was severely stressed by a serious dry period that lasted more than 250 years. The drought, at its most severe in the ninth century, became especially dire when it was punctuated with 3- to 9-year periods of little or no rainfall. The collapse of the Mayan civilization, the result of severe malnutrition and disease, chronic warfare, and social chaos, was the inevitable consequence of inadequate food production caused by insufficient water. However, the Mayan collapse is not just a cautionary tale of geography and climate. It also illustrates that human-caused environmental destruction can have catastrophic consequences.

The Yucatan peninsula is a seasonal desert. Before the Mayans arrived, this land was a dense rainforest growing on top of a porous limestone bedrock and was dependent almost entirely on summer rains. Analysis of pollen and mineral concentrations of modern lake sediments reveals that as the Mayan population grew, the land was gradually deforested for agricultural purposes. Deforestation disrupts the water cycle, the process by which water continuously moves on, above, and below Earth's surface. *Evapotranspiration*, in which water is transpired by plants and evaporated from soil, is an essential feature of the water cycle. The water vapor created by these processes is carried into the atmosphere by air currents and then returns to the Earth as rain. The immediate effect of deforestation in the Yucatan peninsula in the ninth century was erosion as rainwater washed away soil previously held in place by tree roots. In many areas of the Yucatan peninsula, this unimpeded water percolated through porous limestone down into water tables, many of which were too deep for wells. For the Mayans it was a tragedy that deforestation, with its resultant erosion, water runoff, and decreased rainfall (caused by water cycle disruption), occurred in the midst of a regional drought. Together these circumstances contributed to the failure of the Mayans to sustain their civilization.

FIGURE 3.1

A Mayan Pyramid

The Mayans developed an elaborate and sophisticated culture with remarkable achievements in architecture, mathematics, and astronomy. The Mayan civilization eventually collapsed because its population grew so large that it surpassed water and food resources.

Overview

EARTH IS UNIQUE AMONG THE PLANETS IN OUR SOLAR SYSTEM, PRIMARILY BECAUSE OF ITS VAST OCEANS OF WATER. DESPITE DECADES OF RESEARCH, the source of much of this water is still controversial. Whether it originated from Earth's collisions with ice-rich asteroids and comets, the water vapor outgassing by volcanoes, or high-temperature interactions between atmospheric hydrocarbons and the silicate and iron oxides in Earth's mantle over millions of years, water has profoundly affected our planet. Whether falling as rain or flowing in rivers, water has eroded the hardest rocks and transformed the mountains and continents. Many scientists today believe that life arose in a primordial pudding of clay and water. Shallow clay pools can promote the synthesis of macromolecules and accumulate the building blocks of life. In another scenario, life arose in close proximity to hydrothermal vents, openings in the seafloor out

of which flows heated mineral-rich water. Whatever its origin, it is not an accident that life arose in association with water because this substance has several unusual properties that suit it to be the matrix of life. Among these are its thermal properties and unusual solvent characteristics. Water's properties are directly related to its molecular structure.

Why is water so vital for life? Water's chemical stability, its remarkable solvent properties, and its role as a biochemical reactant have long been recognized. What has not been widely appreciated is the critical role that *hydration* (the noncovalent interaction of water molecules with solutes) plays in the architecture, stability, and functional dynamics of macromolecules such as proteins and nucleic acids. Water is now known to be an indispensable component of biological processes as diverse as protein folding and biomolecular recognition in signal transduction mechanisms, the self-assembly of supramolecular structures such as ribosomes, and gene expression. Understanding how essential water is in living processes requires a review of its molecular structure and the physical and chemical properties that are the consequences of that structure.

3.1 MOLECULAR STRUCTURE OF WATER

The water molecule (H_2O) is composed of two atoms of hydrogen and one of oxygen. Water has a tetrahedral geometry because its oxygen atom is sp^3 hybridized, and at the center of the tetrahedron is the oxygen atom. Two of the corners are occupied by hydrogen atoms, each of which is linked to the oxygen atom by a single covalent bond (**Figure 3.2**). The other two corners are occupied by the unshared electron pairs of the oxygen, which repulse each other because of their negatively charged clouds. It is this repulsion that is responsible for pushing the two hydrogens closer together, resulting in water's overall bent geometry.

Oxygen is more electronegative than hydrogen; in other words, oxygen has a greater capacity to attract electrons when bonded to hydrogen. Consequently, the larger oxygen atom bears a partial negative charge (δ^-), and each of the two hydrogen atoms bears a partial positive charge (δ^+) (**Figure 3.3**). The electron distribution in oxygen–hydrogen bonds is displaced toward the oxygen, and, therefore, the bond is **polar**. If water molecules were linear, like those of carbon dioxide (O=C=O), the bond polarities would balance each other and water would be nonpolar. However, water molecules are bent (the bond angle is 104.5°, slightly less than the symmetrical tetrahedral angle of 109°) because the lone pair electrons occupy more space than the bonding electron pairs of the O—H bonds (**Figure 3.4**).

Molecules in which charge is separated are called **dipoles**. Water is a dipole because there is an uneven distribution of electrons between the oxygen atom and the hydrogens. When molecular dipoles are subjected to an electric field, they orient themselves in the direction opposite to that of the field (**Figure 3.5**).

The electron-deficient hydrogens of one water molecule are attracted to the unshared pairs of electrons of another water molecule because of the large difference in electronegativity of hydrogen and oxygen. (Hydrogens attached to nitrogen and fluorine also behave the same way.) In this interaction, called a **hydrogen bond** (**Figure 3.6**), the hydrogen is unequally shared by the two electronegative centers: oxygen nuclei in the case of a pair of water molecules. Hydrogen bonds are considerably weaker than the covalent hydrogen–oxygen bond in the water molecule itself because of their longer bond lengths.

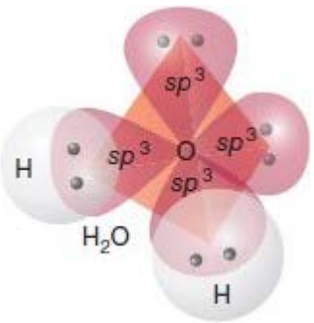


FIGURE 3.2

Tetrahedral Structure of Water

In water, two of the four sp^3 orbitals of oxygen are occupied by two lone pairs of electrons. Each of the other two half-filled sp^3 orbitals is filled by the addition of an electron from hydrogen.

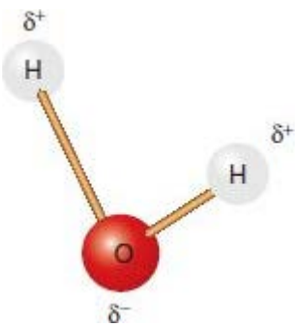


FIGURE 3.3

Charges on a Water Molecule

The two hydrogen atoms in each molecule carry partial positive charges. The oxygen atom carries a partial negative charge.

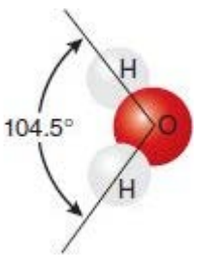


FIGURE 3.4

Space-Filling Model of a Water Molecule

The water molecule has a bent geometry because the distribution of charge within the molecule is asymmetric. Water is therefore polar.

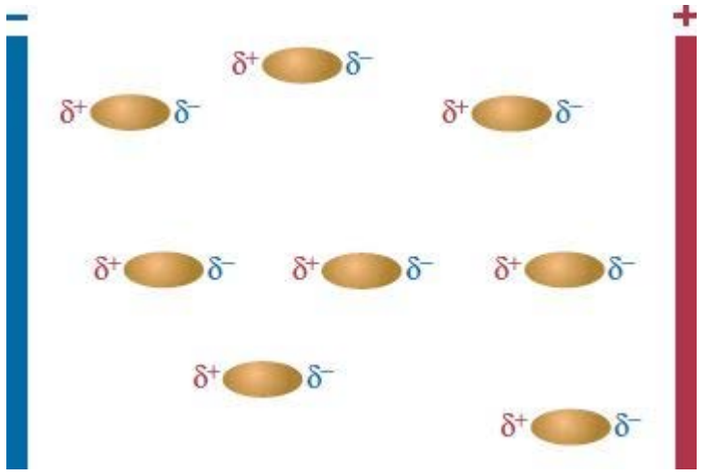


FIGURE 3.5

Molecular Dipoles in an Electric Field

When polar molecules are placed between charged plates, they line up in opposition to the field.

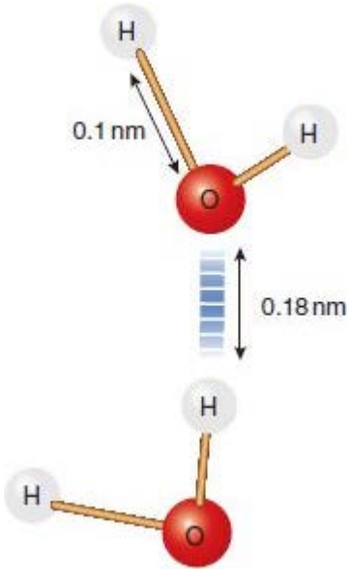


FIGURE 3.6

The Hydrogen Bond

A hydrogen bond results when the electronegative oxygen atoms of two water molecules compete for the same electron-deficient hydrogen atom. The hydrogen bond is represented by short, parallel lines designating the weak covalent character and directionality of the bond.

The hydrogen bond has both electrostatic (ionic) and covalent character. **Electrostatic interactions** occur between any two opposite partial charges (polar molecules) or full charges (ions or charged molecules). **Covalent bonds** involve electron sharing with orbital overlap or mixing. Covalent character confers directionality to the bond or interaction, unlike the uniformly spherical force field around an ion.

3.2 NONCOVALENT BONDING

Noncovalent interactions are usually electrostatic; that is, they occur between the positive nucleus of one atom and the negative electron clouds of another nearby atom. Unlike the stronger covalent

bonds, individual noncovalent interactions are relatively weak and are therefore easily disrupted (**Table 3.1**). Nevertheless, they play a vital role in determining the physical and chemical properties of water and the structure and function of biomolecules because the cumulative effect of many weak interactions can be considerable. Large numbers of noncovalent interactions stabilize macromolecules and supramolecular structures, whereas the capacity of these bonds to rapidly form and break endows biomolecules with the flexibility required for the rapid flow of information that occurs in dynamic living processes. In living organisms, the most important noncovalent interactions include ionic interactions, hydrogen bonds, and van der Waals interactions.

TABLE 3.1 Bond Strengths of Bonds Typically Found in Living Organisms

	Bond Strength*	
Bond Type	**kcal/mol†**	**kJ/mol**
Covalent	>50	>210
Noncovalent		
Ionic interactions	1–20	4–80
Van der Waals forces	<1–2.7	<4–11.3
Mixed: hydrogen bonds	3–7	12–29

* The actual strength varies considerably with the identity of the interacting species.

† 1 cal = 4.184 J.

Ionic Interactions

The ionic interactions that occur between charged atoms or groups are nondirected, meaning they are felt uniformly in space around the center of charge. Oppositely charged ions such as sodium (Na^+) and chloride (Cl^-) are attracted to each other. In contrast, ions with like charges, such as Na^+ and K^+ (potassium), repel each other. In proteins, certain amino acid side chains contain ionizable groups. For example, at physiological pH, the side chain of the amino acid glutamic acid ionizes as $—CH_2CH_2COO^-$ and the side chain group of the amino acid lysine ($—CH_2CH_2CH_2CH_2—NH_2$) ionizes as $—CH_2CH_2CH_2CH_2NH_3^+$. The attraction of positively and negatively charged amino acid side chains forms **salt bridges** ($—COO^- \ ^+H_3N—$), and the repulsive forces created when similarly charged species come into close proximity are an important feature in many biological processes, such as protein folding, enzyme catalysis, and molecular recognition. It should be noted that stable salt bridges rarely form between biomolecules in the presence of water; this is because the hydration of ions is preferred, and the attraction between the biomolecules decreases significantly. Most salt bridges in biomolecules occur in relatively water-free depressions or at biomolecular interfaces where water is excluded.

Hydrogen Bonds

Covalent bonds between hydrogen and oxygen or nitrogen are sufficiently polar that the hydrogen

nucleus is weakly attracted to the lone pair electrons of an oxygen or nitrogen on a neighboring molecule. In the water molecule, each of oxygen's unshared electron pairs can form a weak electrostatic attraction to a hydrogen atom in an interaction, referred to as a hydrogen bond, with nearby water molecules (**Figure 3.7**). Because this interaction is partially covalent, the force of attraction has directionality. Maximum attraction occurs when the two O—H bonds of the participating water molecules are colinear. The resulting intermolecular "bonds" act as a bridge between water molecules. Neither of the hydrogen bonds to oxygen or nitrogen is especially strong (about 20 kJ/mol) in comparison to covalent bonds (e.g., 393 kJ/mol for N—H bonds and 460 kJ/mol for O—H bonds). However, when large numbers of intermolecular hydrogen bonds can be formed (e.g., in the liquid and solid states of water), the molecules effectively become large, dynamic, three-dimensional aggregates. In water, the substantial amounts of energy that are required to break up this aggregate explain the high values for its boiling and melting points, heat of vaporization, and heat capacity. Other properties of water, such as surface tension and viscosity, are also largely a result of its capacity to form large numbers of hydrogen bonds.

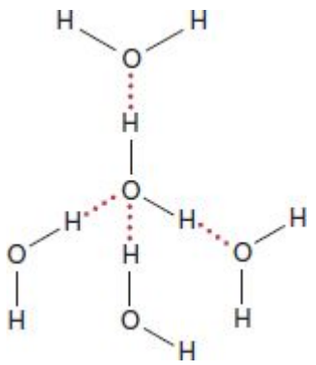


FIGURE 3.7

Tetrahedral Aggregate of Water Molecules

In water, each molecule can form hydrogen bonds with four other water molecules.

Van der Waals Forces

Van der Waals forces are relatively weak electrostatic interactions that arise when biomolecules containing neutral permanent dipoles approach each other or an inducible dipole (such as a π cloud). The more polar and colinear the groups involved, the stronger the van der Waals force. Even in pure hydrocarbons (those with no polar bonds, as in hydrophobic regions of proteins and hydrocarbon tails of lipids), close approach will induce charge delocalization (electrons shift and the charges spread out) that results in cohesiveness. The attraction between molecules is greatest at a distance called the van der Waals radius. If molecules approach more closely, a repulsive force develops, caused by the interpenetration of the outer electron shells of the molecules. In biological systems, the sum of repulsive and attractive forces creates the stable, functional structure of large biomolecules and biomolecular complexes.

There are three types of van der Waals force:

1. **Dipole–dipole interactions**. These forces, which occur between molecules containing electronegative atoms, cause molecules to orient themselves so that the positive end of one polar group is directed toward the negative end of another (**Figure 3.8a**). Hydrogen bonds are an especially strong type of dipole–dipole interaction.
2. **Dipole-induced dipole interactions**. A permanent dipole induces a transient dipole in a nearby molecule by distorting its electron distribution (**Figure 3.8b**). For example, a

carbonyl-containing molecule is weakly attracted to an aromatic ring because of the ability of the permanent dipole of the carbonyl group to delocalize (shift) the electrons of the π electron cloud of the aromatic ring. Dipole-induced dipole interactions are weaker than dipole–dipole interactions.

3. **Induced dipole-induced dipole interactions**. The motion of electrons in nearby nonpolar molecules results in transient charge imbalance in adjacent molecules (**Figure 3.8c**). A transient dipole in one molecule polarizes the electrons in a neighboring molecule. This attractive interaction, often called **London dispersion forces**, is extremely weak. The stacking of the base rings in a DNA molecule, a classic example of this type of interaction, is made possible because of the ability of the loosely held π electrons to distribute unequally above and below closely spaced parallel rings. Although individually weak, these interactions extending over the length of the DNA molecule provide significant stability.

KEY CONCEPTS

- Noncovalent bonds (i.e., ionic interactions and van der Waals forces) are important in determining the physical and chemical properties of living systems.
- Hydrogen bonds, with both dipole–dipole and covalent character, play a critical role in the properties of water and its place in the structure and function of cells.

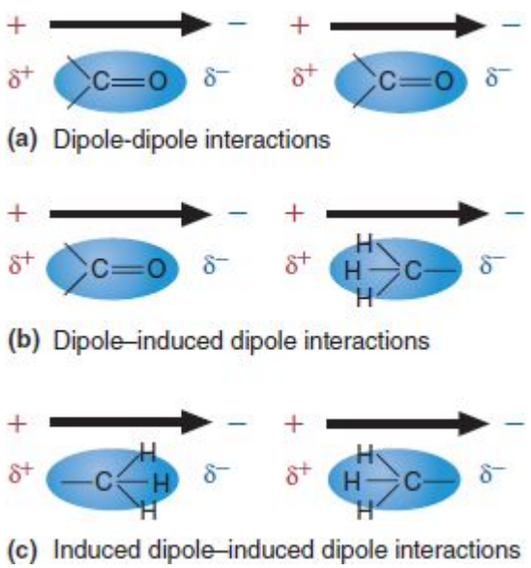


FIGURE 3.8

Dipolar Interactions

The three types of electrostatic interaction involving dipoles are (a) dipole–dipole interactions, (b) dipole-induced dipole interactions, and (c) induced dipole-induced dipole interactions. The relative ease with which electrons respond to an electric field determines the magnitude of van der Waals forces. Dipole–dipole interactions are the strongest; induced dipole–induced dipole interactions are the weakest.

3.3 THERMAL PROPERTIES OF WATER

Perhaps the oddest property of water is that it is a liquid at room temperature. Compared with related molecules of similar molecular weight, water's melting and boiling points are exceptionally high (**Table 3.2**). If water followed the pattern of compounds such as hydrogen

sulfide, it would melt at −100°C and boil at −91°C. Under these conditions, most of Earth's water would be steam, making life unlikely. However, water actually melts at 0°C and boils at +100°C. Consequently, it is a liquid over most of the wide range of temperatures typically found on Earth's surface. Hydrogen bonding is responsible for this anomalous behavior.

Each water molecule can form hydrogen bonds with four other water molecules that, in turn, can form hydrogen bonds with other water molecules. The maximum number of hydrogen bonds form when water has frozen into ice (**Figure 3.9**). Energy is required to break these bonds. When ice is warmed to its melting point, approximately 15% of the hydrogen bonds break. The energy required to melt ice (*heat of fusion*) is substantially higher than expected (**Table 3.3**). Liquid water consists of ice-like clusters of molecules whose hydrogen bonds are continuously breaking and forming. As the temperature rises, the movement and vibrations of the water molecules accelerate, and additional hydrogen bonds are broken. At the boiling point, the water molecules break free from one another and vaporize.

Water is an effective modulator of climatic temperature because of its high *heat of vaporization* (the energy required to vaporize one mole of a liquid at a pressure of one atmosphere) and high *heat capacity* (the energy that must be added or removed to change the temperature of a substance by one degree Celsius). Water also plays an important role in the thermal regulation of living organisms. Its high heat capacity, coupled with the high water content found in most organisms (between 50% and 95%, depending on the species), helps maintain an organism's internal temperature. The evaporation of water serves as a cooling mechanism. An adult human may eliminate as much as 1200 g of water daily in expired air, sweat, and urine. The associated heat loss may amount to approximately 20% of the total heat generated by metabolic processes.

TABLE 3.2 Melting and Boiling Points of Water and Three Other Group VI Hydrogen-Containing Compounds

Name	Formula	Molecular Weight (Da)*	Melting Point (°C)	Boiling Point (°C)
Water	H_2O	18	0	100
Hydrogen sulfide	H_2S	34	−85.5	−60.7
Hydrogen selenide	H_2Se	81	−50.4	−41.5
Hydrogen telluride	H_2Te	129.6	−49	−2

* I dalton (Da) = 1 atomic mass unit (amu).

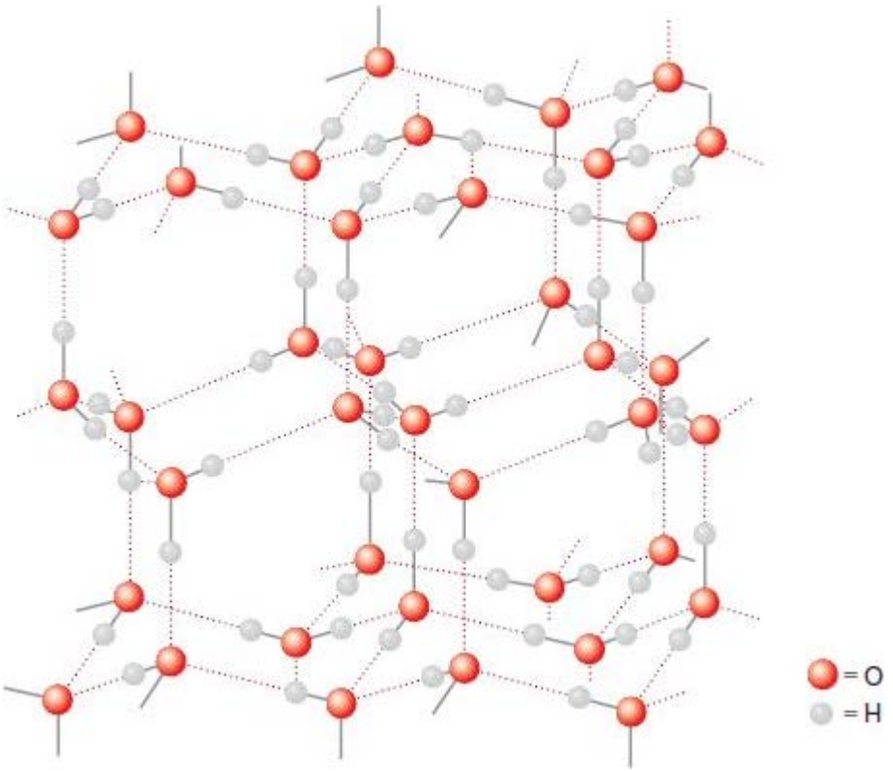


FIGURE 3.9

Hydrogen Bonding between Water Molecules in Ice

Hydrogen bonding in ice produces a very open structure. Ice is less dense than water in its liquid state.

TABLE 3.3 Heat of Fusion of Water and Two Other Group VI Hydrogen-Containing Compounds

Name	Formula	Molecular Weight (Da)	Heat of Fusion* cal/g	J/g
Water	H_2O	18	80	335
Hydrogen sulfide	H_2S	34	16.7	69.9
Hydrogen selenide	H_2Se	81	7.4	31

* The heat of fusion is the amount of heat required to change 1 g of a solid into a liquid at its melting point; 1 cal = 4.184 J.

KEY CONCEPTS

- Hydrogen bonding is responsible for water's unusually high freezing and boiling points.
- Because water has a high heat capacity, it can absorb and release heat slowly. Water plays an important role in regulating body temperature in living organisms.

QUESTION 3.1

Water (H_2O), ammonia (NH_3), and methane (CH_4) have approximately the same molecular weight: 18, 17, and 16 g/mol, respectively. Although all these molecules are structurally in the tetrahedral family, they differ significantly in physical properties. For example, the heat of fusion decreases somewhat from water (6.01 kJ/mol) to ammonia (5.66 kJ/mol) and significantly from water to methane (0.94 kJ/mol). Draw the structure of these molecules and explain the difference in properties based on what you know about hydrogen bonding in the solid state. If it were realistic to generate NH_3 ice (melting point −97.8°C), would you expect it to be more or less dense than liquid ammonia?

WORKED PROBLEM 3.1

Water's capacity to absorb large amounts of energy (heat capacity) with only minimal increases in its temperature is an important factor in the success of life on Earth. Heat capacity (energy absorbed or liberated by a substance when its temperature changes) is given by $q = g \quad C \quad \Delta T$, where

q = energy in joules
g = mass in grams
C = heat capacity
ΔT = change in temperature

Calculate how much energy is required to raise the temperature of 10 g of water 15°C where C_{H_2O} = 4.178 J/g °C. Then calculate the amount of energy needed to raise the temperature of 10 g of sand (SiO_2) (C_{sand} = 0.74 J/g °C) by 15°C.

SOLUTION

The energy absorbed by 10 g of water is determined by substituting the values given for water into the formula:

$$q = (10\ g)\ (4.178\ J/g \cdot °C)\ (15°C) = 627\ J$$

The energy absorbed by sand is calculated as

$$q = (10\ g)\ (0.74\ J/g \cdot °C)\ (15°C) = 111\ J$$ ■

3.4 SOLVENT PROPERTIES OF WATER

Water is the ideal biological solvent. It easily dissolves a wide variety of the constituents of living organisms. Examples include ions (e.g., Na^+, K^+, and Cl^-), sugars, and many of the amino acids. Supramolecular structures (e.g., membranes) and numerous biochemical processes (e.g., protein folding) are possible because water cannot dissolve other substances, such as lipids and certain amino acids. This section describes the behavior of hydrophilic and hydrophobic substances in water. The discussion is followed by a brief review of osmotic pressure, one of the colligative properties of water. Colligative properties are physical properties that are affected not by the

specific structure of dissolved solutes, but by their numbers.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay on water, abiotic stress, and compatible solutes.

Hydrophilic Molecules, Cell Water Structuring, and Sol-Gel Transitions

A dipolar structure and the capacity to form hydrogen bonds with electronegative atoms enable water to dissolve both ionic and polar substances. An important aspect of all ionic interactions in aqueous solution is the hydration of ions. Salts such as sodium chloride (NaCl) are held together by ionic forces. Because water molecules are polar, they are attracted to charged ions such as Na^+ and Cl^-. Shells of water molecules, referred to as **solvation spheres**, cluster around both positive and negative ions (**Figure 3.10**). The size of the solvation sphere depends on the charge density of the ion (i.e., size of charge per unit volume). As ions become hydrated, the attractive force between them is reduced, and the charged species dissolves in the water. Organic molecules with ionizable groups and many neutral organic molecules with polar functional groups also dissolve in water, primarily because of the solvent's hydrogen bonding capacity. Such associations form between water and the carbonyl groups of aldehydes and ketones and the hydroxyl groups of alcohols. The capacity of a solvent to reduce the electrostatic attraction between charges is indicated by its *dielectric constant*. Water, sometimes referred to as the *universal solvent* because of the large variety of ionic and polar substances it can dissolve, has a very large dielectric constant.

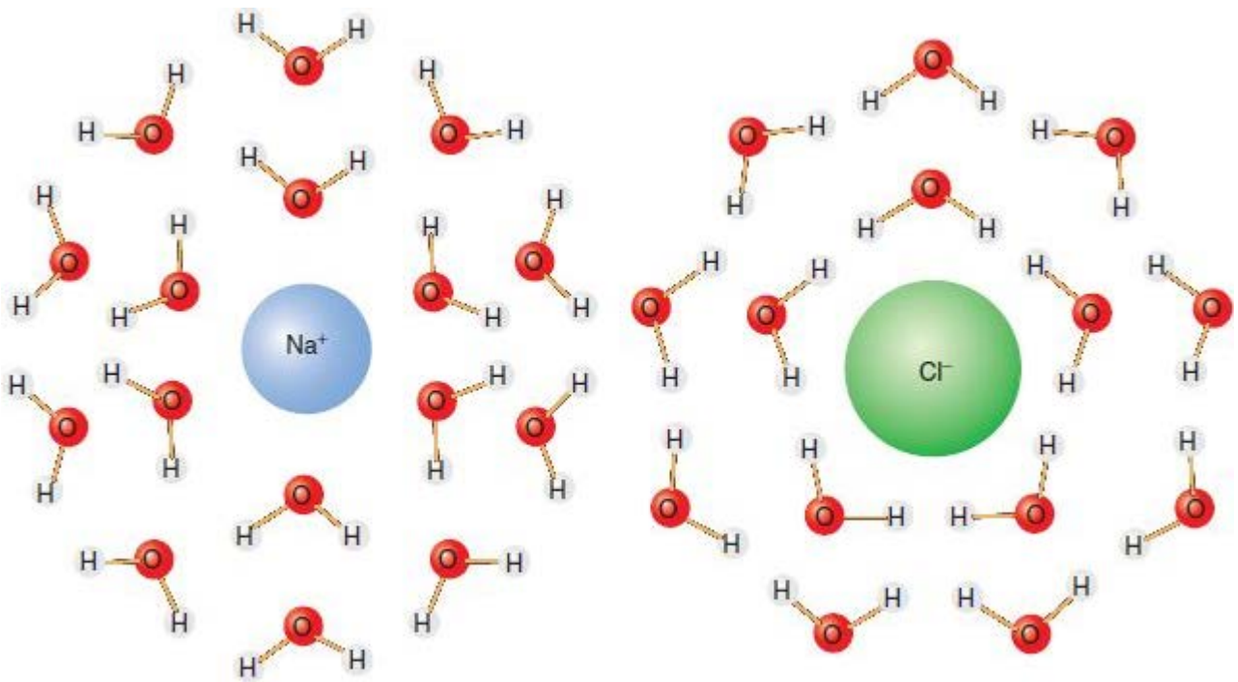


FIGURE 3.10

Solvation Spheres of Water Molecules around Na+ and Cl– Ions

When an ionic compound such as NaCl is dissolved in water, its ions separate because the polar water molecules attract the ions more than the ions attract each other. In reality, the solvation sphere of Na^+ has four

times the volume of that of Cl because of the higher charge density of the sodium ion (the same unit charge distributed over a smaller volume).

STRUCTURED WATER The arrangement of water molecules in living organisms is distinctive. Although organismal water is in liquid form, most water molecules are not in the "bulk water" state (i.e., they do not flow freely). At any given time, most of a cell's billions of water molecules are noncovalently associated with macromolecules and membrane surfaces throughout its densely packed interior. The surfaces of proteins, for example, are studded with positive and negative charges and polar functional groups. Dipolar water molecules readily form hydrogen bonds with such species (**Figure 3.11**). Recall that water molecules are tetrahedral and that each one can form hydrogen bonds with four other water molecules. For this reason, a single layer of water molecules attracts additional water molecules, and an extended three-dimensional network of water molecules forms. In crowded cells, numerous water layers bridge the space between adjacent macromolecules. Moreover, the water molecules in these layers, referred to as *structured water*, are in perpetual motion and constantly rearranging. They exchange with bulk water molecules, farther away from the protein's surface, on time scales that range from femtoseconds (10^{-15} s) to picoseconds (10^{-12} s). The pace of the exchange for an individual water molecule depends on how restricted its motion is. In other words, the closer a water molecule is to a polar surface, the slower its motion. The dynamics of structured water contribute to the structural stability of macromolecules such as proteins. They also facilitate the flexibility required for function.

SOL-GEL TRANSITIONS Cytoplasm, like any water-based material that contains polymers, has the properties of a gel. A *gel* is a colloidal mixture, which is a type of mixture in which small particles are evenly distributed throughout another substance. In cells, the gelatinous cytoplasm is a semisolid composed of biopolymers with polar surfaces in association with adsorbed water. Gelatin desserts are well-known examples of gels with fibers of the protein collagen suspended and hydrated in a large quantity of water. The highly structured solvation layers on a matrix of protein give the viscoelastic properties we associate with Jell-O. The stability of a gel depends on the length and cross-linking of the polymer and the continuity of the adsorbed water. The freedom of solutes to move within this meshwork or gel matrix varies with the *trabecular* (resembling a sponge) arrangement of the protein polymers. If you punch wells in a petri dish filled with solidified gelatin and pour a solution of inorganic ions into the wells, the ions will migrate out into the gel at rates related to their size and degree of hydration. You can observe the results of this sieving effect in just a few minutes. In addition, the water solvating the surface of the gelatin (collagen polymers) is for all practical purposes fixed in position; that is, diffusion is limited.

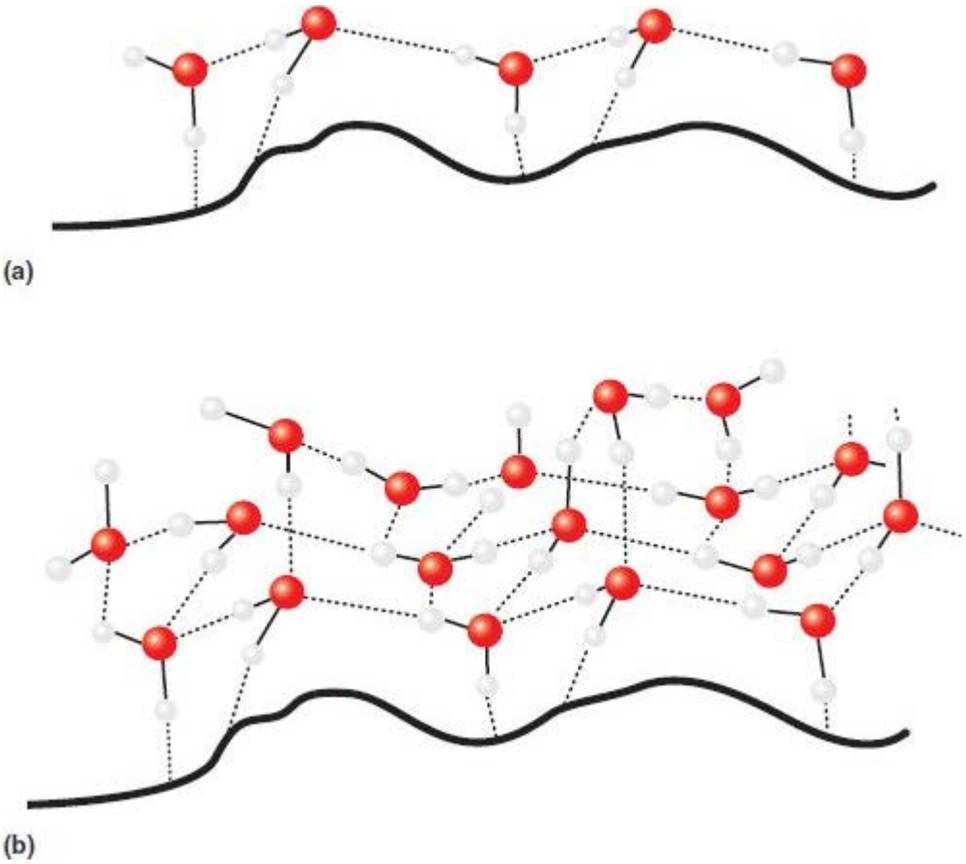


FIGURE 3.11

Diagrammatic View of Structured Water

Polar surfaces of macromolecules attract water molecules: (a) a short segment of a polar macromolecular surface with a single water layer, which attracts additional water molecules that form an extended network (b).

Changes in temperature (and therefore molecular motion), matrix architecture, and inclusion of solutes can lead to a transition from the gel to a "sol" or liquid state. Cells behave in a similar way because of the highly structured solvation surfaces of the polymeric proteins. Transitions from gel to sol (from more solid to less solid) contribute to many aspects of cell function, most notably cell movement. These transitions are caused by the reversible polymerization of G-actin to form F-actin and the subsequent cross-linking of actin filaments. These transitions are carefully regulated by signal transduction mechanisms that affect the concentrations and functions of a group of proteins called *actin-binding proteins* (p. 58). Various actin-binding proteins can inhibit polymerization or they can cross-link or sever actin filaments.

Amoeboid motion provides an example of the highly regulated nature of cellular sol–gel transitions and the forces that such transformations create. The principal feature of amoeboid motion is the protrusion of a cellular extension called a *pseudopodium* (**Figure 3.12**). The pseudopodium moves forward because polymerizing actin filaments in the cell cortex (the *ectoplasm*) in this isolated part of the cell undergo further cross-linking (a sol-to-gel transition). Once this has occurred, actin filaments in the cell's interior (the *endoplasm*) depolymerize, effecting a gel-to-sol transition. Simultaneously, a contractile force is created by the binding of actin filaments to myosin (a motor protein) in the trailing end of the cell. This force squeezes the freely flowing endoplasm, causing it to stream forward into the pseudopodium.

Hydrophobic Molecules and the Hydrophobic Effect

Small amounts of nonpolar substances mixed with water are excluded from the solvation network of the water; that is, they coalesce into droplets. This process is called the *hydrophobic effect*. **Hydrophobic** ("water-hating") molecules, such as the hydrocarbons, are virtually insoluble in water. Their association into droplets (or, in larger amounts, into a separate layer) results from the solvent properties of water, not from the relatively weak attraction between the associating nonpolar molecules. When nonpolar molecules enter an aqueous environment, the water molecules organize into a cage-like structure that drives the hydrophobic region in on itself (a partitioning or exclusion process). The excluded hydrophobic phase is ultimately stabilized by van der Waals interactions between closely spaced nonpolar regions (**Figure 3.13**). The water-caged structure, or *clathrate*, is stabilized when exposure of water to the hydrophobic material is minimized. The hydrophobic effect is responsible for the generation of stable lipid membranes and contributes to the fidelity of protein folding.

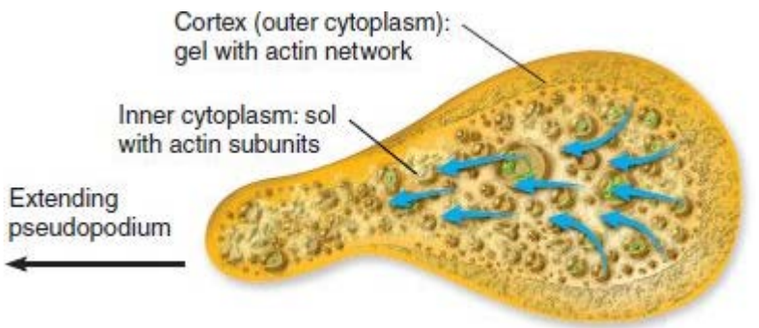


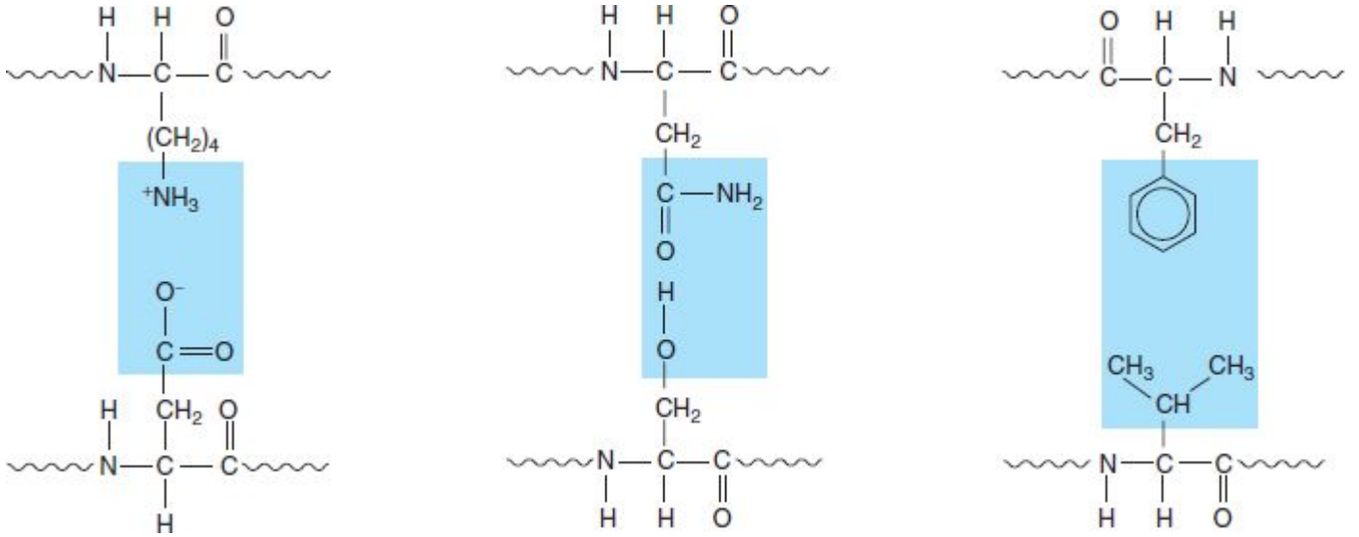
FIGURE 3.12

Amoeboid Motion and Sol-Gel Transitions

The cell moves forward because of the coordination of sol-gel transitions in the cell cortex (ectoplasm) and cytoplasm in the cell's interior (endoplasm). A contractile force in the rear of the cell squeezes the fluid endoplasm forward.

QUESTION 3.2

Proteins are amino acid polymers. Noncovalent bonding plays an important role in determining the three-dimensional structures of proteins. The noncovalent interactions indicated here by shaded areas are typical of the bonding that occurs between amino acid side chains.



Which type of noncovalent bond is primarily responsible for each of the interactions indicated in the figure?

QUESTION 3.3

Collagen, a large fiber-like protein, combined with other molecules, forms a gel-like material found in shock-absorbing body components (e.g., tendons and ligaments). Explain the role of structured water in the function of these tissues. [*Hint*: Water is an incompressible substance.]

Collagen as Shock Absorber

Amphipathic Molecules

A large number of biomolecules, referred to as **amphipathic**, contain both polar and nonpolar groups. This property significantly affects their behavior in water. For example, ionized fatty acids are amphipathic molecules because they contain hydrophilic carboxylate groups and hydrophobic hydrocarbon groups. When they are mixed with water, amphipathic molecules form structures called **micelles** (**Figure 3.14**). In micelles, the charged species (the carboxylate groups), called *polar heads*, orient themselves so that they are in contact with water. The nonpolar hydrocarbon "tails" become sequestered in the hydrophobic interior. The tendency of amphipathic biomolecules to spontaneously rearrange themselves in water is an important feature of numerous cell components. For example, a group of bilayer-forming phospholipid molecules is the basic structural feature of biological membranes (see Chapter 11).

KEY CONCEPTS

- Water's dipolar structure and its capacity to form hydrogen bonds enable water to dissolve many ionic and polar substances.
- Nonpolar molecules cannot form hydrogen bonds with water and are excluded via clathrate formation.

- Amphipathic molecules, such as fatty acid salts, spontaneously rearrange themselves in water to form micelles.

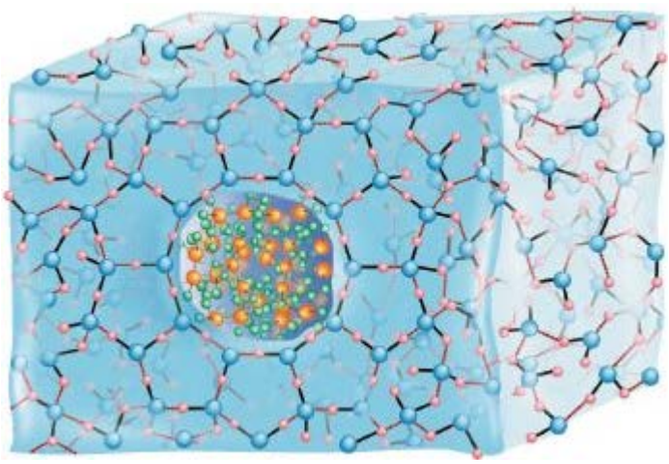


FIGURE 3.13

The Hydrophobic Effect

When nonpolar molecules and water are mixed, a cage of organized hydrogen-bonded water molecules forms to minimize exposure to the hydrophobic substance. Nonpolar molecules, when in close proximity, are attracted to each other by van der Waals forces. However, the driving force in the formation of the cage and exclusion of the hydrophobic substance is the strong tendency of water molecules to form hydrogen bonds among themselves. Nonpolar molecules are excluded because they cannot form hydrogen bonds.

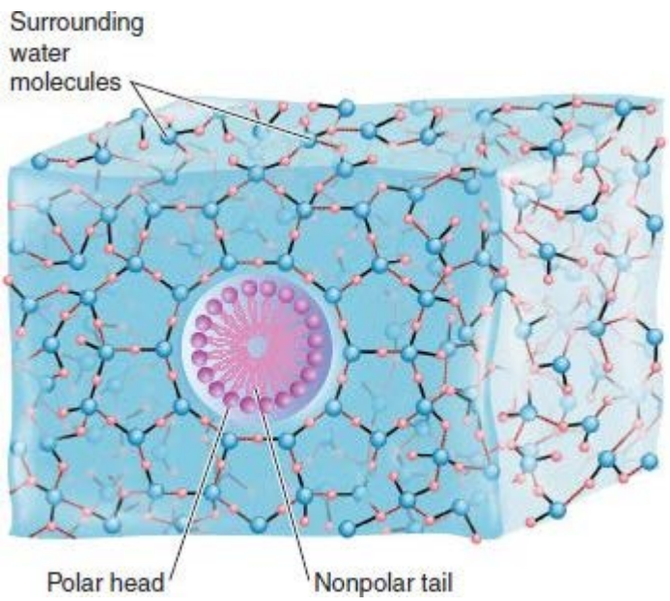


FIGURE 3.14

Formation of Micelles

The polar heads of amphipathic molecules orient themselves so that they are hydrogen-bonded to water. The nonpolar tails aggregate in the center, away from water.

Osmotic Pressure

Osmosis is the spontaneous passage of solvent molecules through a semipermeable membrane that

separates a solution of lower solute concentration from a solution of higher solute concentration. Pores in the membrane are wide enough to allow solvent molecules to pass through in both directions but too narrow for the larger solute molecules or ions to pass. **Figure 3.15** illustrates the movement of solvent across a membrane. As the process begins, there are fewer water molecules on the high-solute concentration side of the membrane. Over time, more water moves from side A (lower solute concentration) to side B (higher solute concentration). The higher the concentration of water in a solution (i.e., the lower the solute concentration), the greater the rate of water flow through the membrane.

Osmosis is the principal cause of water flow across cellular membranes and a driving force in numerous living processes. For example, osmosis appears to be a significant factor in the formation of sap in trees. Cell membranes are not, strictly speaking, osmotic membranes because they permit molecules other than solvent (water) to move across the membrane. The term *dialyzing membrane* would be more accurate.

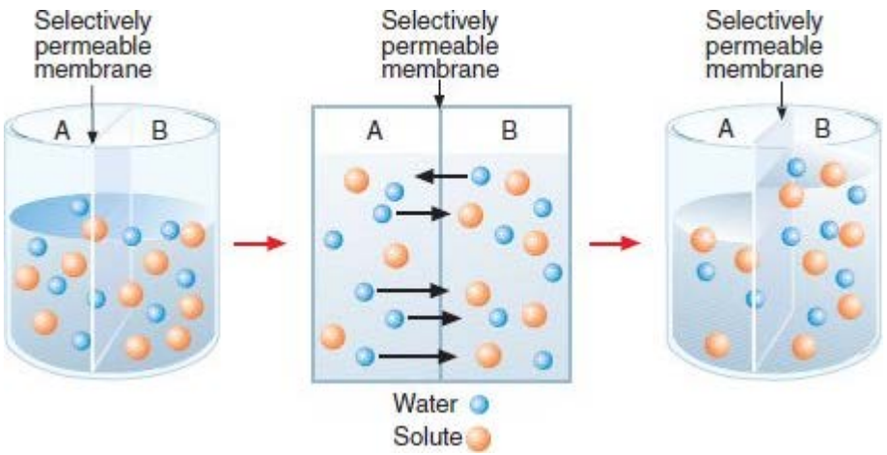


FIGURE 3.15

Osmotic Pressure

Over time, water diffuses from side A (more dilute) to side B (more concentrated). Equilibrium between the solutions on both sides of a semipermeable membrane is attained when there is no net movement of water molecules between side A and side B. Osmotic pressure stops the net flow of water across the membrane.

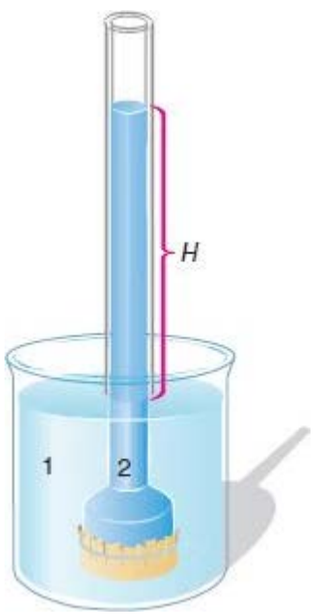


FIGURE 3.16

The Measurement of Osmosis Using an Osmometer

Volume 1 contains pure water. Volume 2 contains a solution of sucrose. The membrane is permeable to water but not to the sucrose. Therefore, there will be a net movement of water into the osmometer. Osmotic pressure, which is proportional to the height *H* of the solution in the tube, can be measured directly with the use of a piston (not shown) connected to a force-generating device that returns volume 2 to the same level of volume 1.

Osmotic pressure is defined as the force necessary to resist the movement of water across a semipermeable membrane. In other words, the net movement of water stops when the pressures exerted on both sides of the membrane become equal. Osmotic pressure, measured with a device called an *osmometer* (**Figure 3.16**), can also be calculated using the following equation, keeping in mind that the final osmotic pressure reflects the contribution of all solutes present.

$$\pi = iMRT$$

where π = osmotic pressure (atm)
i = van't Hoff factor (reflecting the extent of ionization of solutes)
M = molarity (mol/L)
R = gas constant (0.082 L·atm/K·mol)
T = temperature (K)

The concentration of a solution can be expressed in terms of *osmolarity*. The unit of osmolarity is osmoles (osmol) per liter. In the equation $\pi = iMRT$, the osmolarity is equal to iM, where i (the van't Hoff factor) represents the degree of ionization of the solute species, which varies with temperature. The degree of ionization of a 1 M NaCl solution is 90%, with 10% of the NaCl existing as ion pairs. Thus, if we write

$$i = [Na^+] + [Cl^-] + [NaCl]_{un\text{-}ionized} = 0.9 + 0.9 + 0.1 = 1.9$$

the value of i for this solution is 1.9. The value of i approaches 2 for NaCl solutions as they become increasingly more dilute. The value of i for a 1 M solution of a weak acid that undergoes a 10% ionization is 1.1. The value of i for a nonionizable solute is always 1.0. Problems 3.1, 3.2, and 3.3 use the concept of osmotic pressure.

Osmotic pressure creates some critical problems for living organisms. Cells typically contain fairly high concentrations of solutes, that is, small organic molecules and ions, as well as lower concentrations of macromolecules. Consequently, cells may gain or lose water because of the concentration of solute in their environment. If cells are placed in an **isotonic solution** (i.e., the concentration of solute and water is the same on both sides of the selectively permeable plasma membrane), there is no net movement of water in either direction across the membrane (**Figure 3.17**). For example, red blood cells are isotonic to a 0.9% NaCl solution. When cells are placed in a solution with a lower solute concentration (i.e., a **hypotonic solution**), water moves into the cells. When red blood cells are immersed in pure water, for example, they swell and rupture in a process called *hemolysis*. In solutions with higher solute concentrations (i.e., **hypertonic solutions**), cells shrivel because there is a net movement of water out of the cell. The shrinkage of red blood cells in a hypertonic solution (e.g., a 3% NaCl solution) is referred to as *crenation*.

WORKED PROBLEM 3.2

When 0.25 g of urea (MW 60) is diluted to 100 mL with water, what is the osmotic pressure

of the solution? [Assume room temperature, i.e., 25°C (298 K).]

SOLUTION

Calculate the molarity of the urea solution. Urea is a nonelectrolyte, so the van't Hoff factor (i) is 1.

$$\text{Molarity} = \frac{0.10\ \text{g urea} \times 1.0\ \text{mol}}{60\ \text{g}} \times \frac{1}{0.10\ \text{L}} = 1.7 \times 10^{-2}\ \text{mol/L}$$

The osmotic pressure at room temperature is given by

$$\begin{aligned}\pi &= iMRT \\ &= (1)\frac{4.2 \times 10^{-2}\ \text{mol}}{\text{L}}\ \frac{0.0821\ \text{L} \cdot \text{atm}}{\text{K} \cdot \text{mol}}\ (298\ \text{K}) \\ &= 1.0\ \text{atm}\end{aligned}$$

■

WORKED PROBLEM 3.3

Estimate the osmotic pressure of a solution of 0.1 M NaCl at 25°C. Assume 100% ionization of solute.

SOLUTION

A solution of 0.10 M NaCl produces 0.2 mol of particles per liter (0.10 mol of Na^+ and 0.10 mol of Cl^-). The osmotic pressure at room temperature is

$$\begin{aligned}\pi &= \frac{2 \times 0.10\ \text{mol}}{\text{L}}\ \frac{0.0821\ \text{L} \cdot \text{atm}}{\text{K} \cdot \text{mol}}\ (298\ \text{K}) \\ &= 4.9\ \text{atm}\end{aligned}$$

■

WORKED PROBLEM 3.4

Osmotic pressure can be used as a method to estimate the molecular mass of a biomolecule. Determine the molecular mass (m) of the nonionic compound X (i = 1). When 1.0 g of compound X is dissolved in 100 ml of water, the solution has an osmotic pressure of 0.3 atm at 25°C.

SOLUTION

Calculate the molarity (M) of the solution using the formula for osmotic pressure.

$\pi = iMRT$

$M = \pi/iRT$

$M = (0.30 \text{ atm})/[(1)(0.0821 \text{ L} \cdot \text{atm/mol} \cdot \text{K})\ (298 \text{ K})]$

$M = 12.3 \times 10^{-3} \text{ mol/L}$

Use the formula for molarity and the given mass and volume (V) to calculate the molecular mass (m) of compound *X*.

$\text{M} = \text{mass}/[\text{m/V}]$

$\text{m} = \text{mass/MV}$

$\text{m} = 1.0 \text{ g}/[12.3 \times 10^{-3} \text{ mol/L } (0.1 \text{ L})]$

$\text{m} = 8.1 \times 10^{2} \text{ g/mol}$

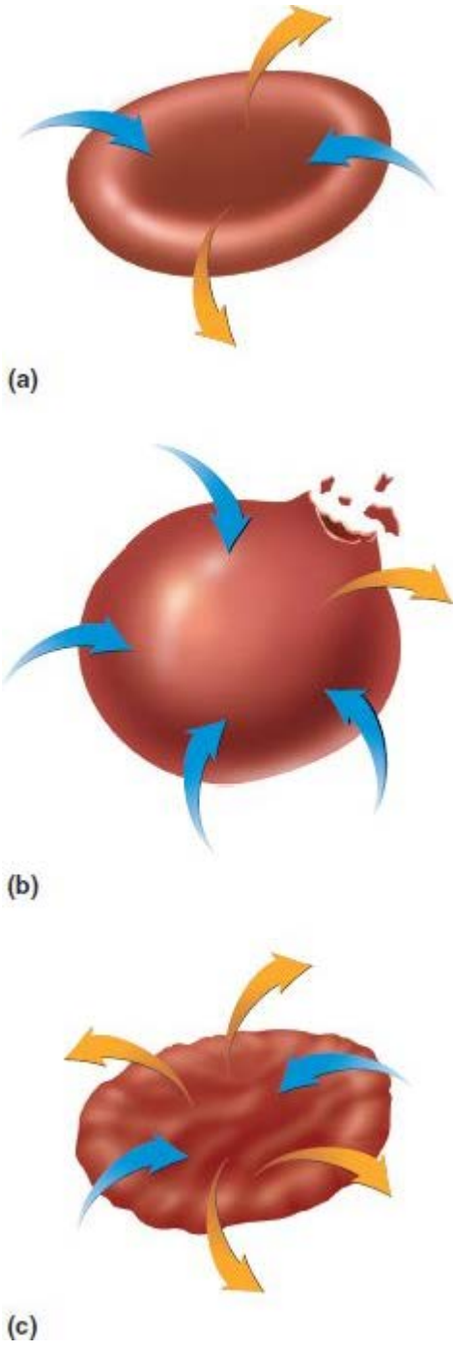


FIGURE 3.17

The Effect of Hypertonic and Hypotonic Solutions on Animal Cells

(a) Isotonic solutions do not change cell volume because water is entering and leaving the cell at the same rate. (b) Hypotonic solutions cause cell rupture. (c) Hypertonic solutions cause cell shrinkage (crenation).

ION DISTRIBUTION ACROSS CELL MEMBRANES Macromolecules have little direct effect on cellular osmolarity because their cellular molar concentrations are relatively low. However, macromolecules such as the proteins contain a large number of ionizable groups. The ions of opposite charge that are attracted to these groups have a substantial effect on intracellular osmolarity. The nature of this effect is determined by the interaction of the structured water associated with proteins with hydrated ions. The size of an ion's solvation sphere is inversely related to its charge density (size of charge per unit volume). For example, sodium and potassium ions have nonhydrated diameters of 1.96 and 2.66 Å, respectively. The hydrated diameters of sodium and potassium ions are 9.0 and 6.0 Å, respectively. Consequently, the hydrated volume of

Na^+ is 3.4 times that of K^+. In addition, since the solvation sphere of K^+ is much smaller than that of Na^+, the potassium ion solvation sphere is easier to remove, thus allowing formation of ion pairs with the surface anions of proteins. It costs significantly more energy to remove the solvation sphere of Na^+, a step that must occur for the sodium ion to move through ion channels. As a result, the ion distribution across the cell membrane is unequal, with the tendency to accumulate inside the cell being much greater for K^+ (159 mM) than for Na^+ (10 mM). This inequality would occur even if specific ion pumps were not present.

Unlike most inorganic ions, the ionizable groups of cellular proteins are fixed within the cell, conferring a significant net negative charge to the intracellular environment. As a consequence, there exists an electronegative gradient across the cell membrane: that is, the ions and negative charge are distributed unequally. The cytoplasmic side of the membrane is strongly negative, an effect that is partly offset by potassium ions. The outside of the membrane is positive because of the relatively large number of extracellular sodium ions. The existence of this asymmetry on the surfaces of cell membranes results in the establishment of an electrical gradient, called a **membrane potential**, which provides the means for electrical conduction, active transport, and even passive transport.

Although hydrated sodium ions tend to be excluded from structured water inside cells, leakage of these ions back across the plasma membrane does occur. Small intracellular increases in $[Na^+]$ cause the cytoplasm to be slightly less negative. As a result, small amounts of K^+ move down their concentration gradient out of the cell. Animals and bacteria control cell volume by opposing this process with ATP-driven Na^+-K^+ pumps. Ion pumping in these cells requires substantial amounts of energy. (Refer to Biochemistry in Perspective: Cell Volume Regulation and Metabolism on p. 103 for an insight into how osmotic pressure affects cell volume.)

KEY CONCEPTS

- Osmosis is the movement of water across a semipermeable membrane from a dilute solution to a more concentrated solution.
- Osmotic pressure is the pressure exerted by water on a semipermeable membrane that prevents the flow of water across the membrane.

3.5 IONIZATION OF WATER

Liquid water molecules have a limited capacity to ionize to form a proton, or hydrogen ion (H^+), and a hydroxide ion (OH^-). Protons do not actually exist in aqueous solution. In water, a proton combines with a water molecule to form H_3O^+, commonly referred to as *hydronium ion* (**Figure 3.18**). For convenience, H^+ will be used to represent the ionization reactions of water.

The disassociation of water

$$H_2O(l) \rightleftharpoons H^+ + OH^-$$

may be expressed as

$$K_{eq} = \frac{[H^+][OH^-]}{H_2O}$$

where K_{eq} is the equilibrium constant for the reaction.

The concentration of unionized water is essentially unchanged and can be thought of as a constant. The equilibrium expression can be rewritten combining the two constants as

$$K_{eq}[H_2O] = [H^+]\,[OH^-]$$

The term $K_{eq}[H_2O]$ is referred to as the *ion product of water* or K_w. After substitution of the term K_w, the preceding equation may be rewritten as

$$K_w = [H^+]\,[OH^-]$$

The K_w for H_2O at 25°C and 1 atm pressure is 1.0×10^{-14} and is a fixed characteristic of water at this temperature and pressure. In pure water, where there are no other contributors of H^+ or OH^-, the concentrations of these ions are equal:

$$[H^+] = [OH^-] = (K_w)^{1/2} = (1.0 \times 10^{-14})^{1/2} = 1.0 \times 10^{-7}\ M$$

A solution that contains equal amounts of H^+ and OH^- is said to be *neutral*. When an ionic or polar substance is dissolved in water, it may change the relative numbers of H^+ and OH^-. Solutions with an excess of H^+ are *acidic*, whereas those with a greater number of OH^- are *basic*. Hydrogen ion concentration varies over a wide range: commonly between 10^0 and 10^{-14} M, which provides the basis of the pH scale ($pH = -\log [H^+]$).

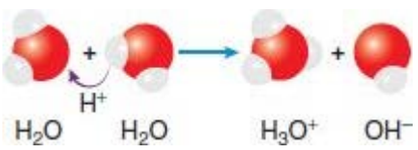


FIGURE 3.18

Hydronium Ion

The hydronium ion (H_3O^+) is formed by the protonation of a water molecule. When a water molecule dissociates to yield a proton (H^+) and a hydroxide ion (OH^-), the electronegative oxygen atom of a nearby water molecule is attracted to the positively charged proton.

3D animation of Hydronium Ion H30+

Acids, Bases, and pH

The concentration of the hydrogen ion, one of the most important ions in biological systems, affects most cellular and organismal processes. For example, the structure and function of proteins and the rates of most biochemical reactions are strongly affected by hydrogen ion concentration. Additionally, hydrogen ions play a major role in processes such as energy generation (see Chapter 10) and endocytosis (p. 53).

Many biomolecules have acidic and/or basic properties. Large polymers and macromolecular complexes usually have amphoteric surfaces; that is, they possess both acidic and basic groups. A side group of a molecule is said to be an **acid** if it is a proton donor and a **base** if it is a proton acceptor.

Strong acids (e.g., HCl) and bases (e.g., NaOH) ionize almost completely in water:

$$HCl \rightarrow H^+ + Cl^-$$
$$NaOH \rightarrow Na^+ + OH^-$$

Many acids and bases, however, do not dissociate completely. Organic acids (compounds with carboxyl groups) are referred to as **weak acids** because of their partial dissociation in water. Organic bases have a small but measurable capacity to combine with hydrogen ions. Many common **weak bases** contain amino groups.

The dissociation of an organic acid is described by the following reaction:

$$\underset{\text{Weak acid}}{HA} \rightleftharpoons H^+ + \underset{\text{base of HA}}{A^-}$$

Note that the deprotonated product of the dissociation reaction is referred to as a **conjugate base**. For example, acetic acid (CH_3COOH) dissociates to form the conjugate base acetate (CH_3COO^-).

The strength of a weak acid (i.e., its capacity to release hydrogen ions) may be determined using the following expression:

$$K_a = \frac{[H^+][A^-]}{HA}$$

where K_a is the acid dissociation constant.

TABLE 3.4 Dissociation Constants and pK_a Values for Common Weak Acids*

Acid	HA	A^-	K_a	pK_a
Acetic acid	CH_3COOH	CH_3COO^-	1.76×10^{-5}	4.76
Carbonic acid	H_2CO_3	HCO_3^-	4.5×10^{-7}	6.35
Bicarbonate	HCO_3^-	CO_3^{2-}	5.61×10^{-11}	10.33
Lactic acid	$CH_3CHCOOH$ (OH)	CH_3CHCOO^- (OH)	1.38×10^{-4}	3.86
Phosphoric acid	H_3PO_4	$H_2PO_4^-$	7.25×10^{-3}	2.14
Dihydrogen	$H\ PO^-$	HPO^{2-}	$6.31 \times$	7.20

phosphate	$H_2PO_4^-$	HPO_4^{2-}	4	10^{-8}

* Equilibrium constants should be expressed in terms of activities rather than concentrations (activity is the effective concentration of a substance in a solution). However, in dilute solutions, concentrations may be substituted for activities with reasonable accuracy.

The larger the value of K_a, the stronger the acid is. Because K_a values vary over a wide range, they are expressed using a logarithmic scale:

$$pK_a = -\log K_a$$

The lower the pK_a, the stronger the acid. Dissociation constants and pK_a values for several common weak acids are given in **Table 3.4**.

The **pH scale** (**Figure 3.19**) can be used to determine hydrogen ion concentration $[H^+]$:

$$pH = -\log [H^+]$$
$$[H^+] = \text{antilog } (-pH)$$

On the pH scale, neutrality is defined as pH 7; that is, $[H^+]$ is equal to 1×10^{-7} M. Acidic solutions have pH values less than 7; that is, $[H^+]$ is greater than 1×10^{-7} M. A pH value greater than 7 indicates a solution that is basic, or alkaline.

It is important to note that, although pK_a and pH appear to be similar mathematical expressions, they are in fact different. At constant temperature, the pK_a value of a substance is a constant. In contrast, the pH values of a system may vary.

Buffers

The regulation of pH is a universal and essential activity of living organisms. Hydrogen ion concentration must typically be kept within narrow limits. For example, normal human blood has a pH of 7.4. It may vary between 7.35 and 7.45, depending on the concentrations of acidic and basic waste products and metabolites. Certain disease processes cause pH changes that, if not corrected, can be disastrous. **Acidosis**, a condition that occurs when human blood pH falls below 7.35, results from excessive production of acid in the tissues, loss of base from body fluids, or failure of the kidneys to excrete acidic metabolites. Acidosis occurs in certain diseases (e.g., diabetes mellitus) and during starvation. If blood pH drops below 7, the central nervous system becomes depressed, resulting in coma and eventually death. When blood pH rises above 7.45, **alkalosis** results. This condition, brought on by prolonged vomiting or by ingestion of excessive amounts of alkaline drugs, overexcites the central nervous system. Muscles then go into a state of spasm. If this situation is uncorrected, convulsions and respiratory arrest develop.

Acidosis

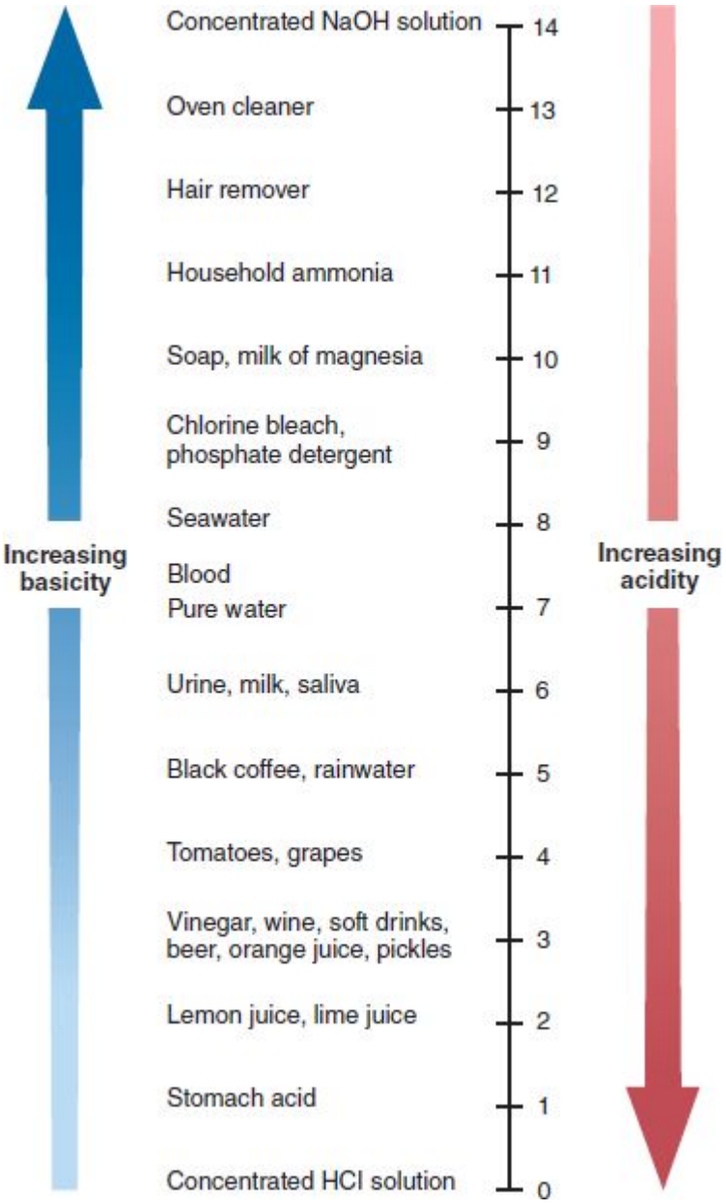


FIGURE 3.19

The pH Scale and the pH Values of Common Fluids

Buffers help maintain a relatively constant hydrogen ion concentration. The most common buffers consist of mixtures of weak acids and their conjugate bases. A buffered solution can resist pH changes because an equilibrium between the buffer's components is established. Therefore, buffers obey **Le Chatelier's principle**, which states that if a stress is applied to a reaction at equilibrium, the equilibrium will be displaced in the direction that relieves the stress. Consider a solution containing acetate buffer, which consists of acetic acid and sodium acetate (**Figure 3.20**). The buffer is created by mixing a solution of sodium acetate with a solution of acetic acid to create an equilibrium mixture of the correct pH and ionic strength.

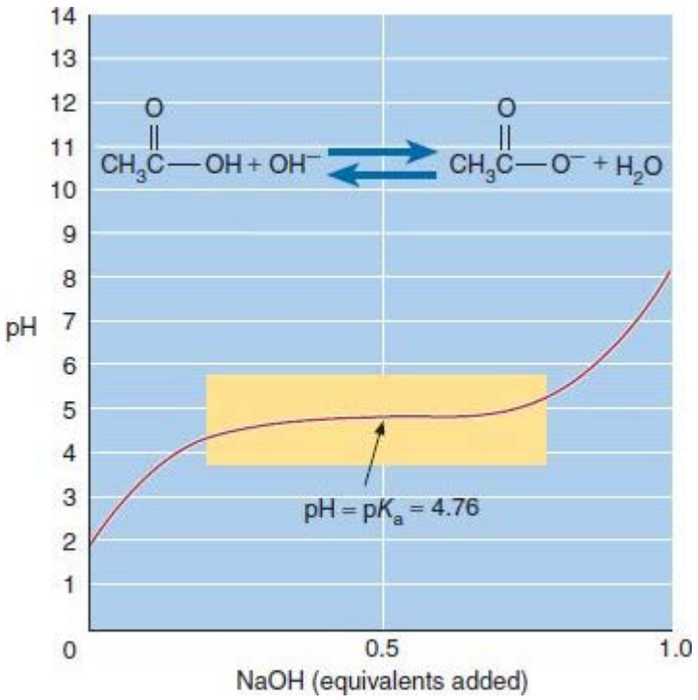


FIGURE 3.20

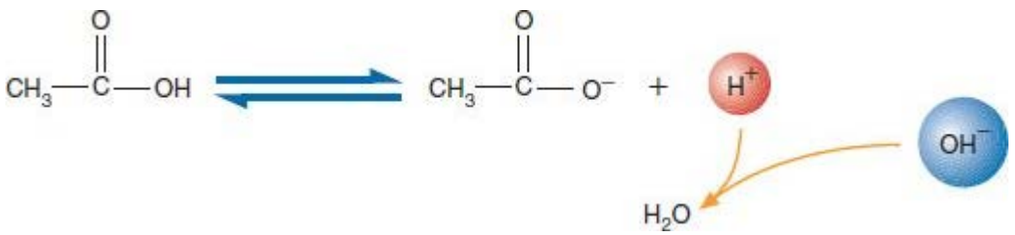
Titration of Acetic Acid with NaOH

The shaded band indicates the pH range over which acetate buffer functions effectively. A buffer is most effective at or near its pK_a value.

If hydrogen ions are added, the equilibrium shifts toward the formation of acetic acid with the $[H^+]$ changing little:

$$H^+ + CH_3COO^- \rightarrow CH_3COOH$$

If hydroxide ions are added, they react with the free hydrogen ions to form water, the equilibrium shifts to the acetate ion, and the pH changes little.



BUFFERING CAPACITY The capacity of a buffer to maintain a specific pH depends on two factors: (1) the molar concentration of the acid–conjugate base pair and (2) the ratio of their concentrations. Buffering capacity is directly proportional to the concentration of the buffer components. In other words, the more molecules of buffer present, the more H^+ and OH^- ions can be absorbed without significantly changing the pH. The concentration of the buffer is defined as the sum of the concentration of the weak acid and its conjugate base. For example, a 0.2 M acetate buffer may contain 0.1 mol of acetic acid and 0.1 mol of sodium acetate in 1 L of H_2O. Such a buffer may also consist of 0.05 mol of acetic acid and 0.15 mol of sodium acetate in 1 L of H_2O. The most effective buffers are usually those that contain equal concentrations of both components or the pH is equal to the pK_a. Biological systems generate acids during metabolism, and buffer

capacity for acid neutralization must be maximized. Consequently, biological buffers often contain a higher concentration of the conjugate base. Bicarbonate buffer (p. 101) is an example of such a buffering system.

HENDERSON–HASSELBALCH EQUATION In choosing or making a buffer, the pH and pK_a concepts are useful. The relationship between these two quantities is expressed in the Henderson–Hasselbalch equation, which is derived from the following equilibrium expression:

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

Solving for $[H^+]$ results in

$$[H^+] = K_a \frac{[HA]}{[A^-]}$$

Taking the negative logarithm of each side, we obtain

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

Defining $-\log [H^+]$ as pH and $-\log K_a$ as pK_a gives

$$\text{pH} = \text{p}K_a - \log\frac{[HA]}{[A^-]}$$

If the log term is inverted, thereby changing its sign, the *Henderson–Hasselbalch equation* is obtained:

$$\text{pH} = \text{p}K_a + \log\frac{[A^-]}{[HA]}$$

Note that when $[A^-] = [HA]$, the equation becomes

$$\begin{aligned}\text{pH} &= \text{p}K_a + \log 1\\ &= \text{p}K_a + 0\end{aligned}$$

Under this circumstance, pH is equal to pK_a. **Figure 3.20** illustrates that buffers are most effective when they are composed of equal amounts of weak acid and conjugate base. The most effective buffering occurs in the portion of the titration curve that has a minimum slope, that is, 1 pH unit above and below the value of pK_a. In the graph, the abscissa displays the equivalents added. Here, an equivalent is the mass of base that can accept 1 mol of H^+ ions; an acid equivalent gives the mass of acid that can donate a mole of protons.

Problems 3.5 through 3.11 are typical buffer problems.

KEY CONCEPTS

- Liquid water molecules have a limited capacity to ionize to form H^+ and OH^- ions.
- The concentration of hydrogen ions is a crucial feature of biological systems primarily because of their effects on biochemical reaction rates and protein structure.
- Buffers, which consist of weak acids and their conjugate bases, prevent changes in pH (a measure of $[H^+]$).

WORKED PROBLEM 3.5

Calculate the pH of a mixture of 0.25 M acetic acid and 0.20 M sodium acetate. The pK_a of acetic acid is 4.76.

SOLUTION

$$\begin{aligned}\text{pH} &= \text{p}K_a + \log\frac{[\text{acetate}]}{[\text{acetic acid}]}\\ &= 4.76 + \log\frac{0.20}{0.25}\\ &= 4.76 - 0.097\\ &= 4.66\end{aligned}$$

■

WORKED PROBLEM 3.6

What is the pH in the preceding problem if the mixture consists of 0.10 M acetic acid and 0.25 M sodium acetate?

SOLUTION

$$\begin{aligned}\text{pH} &= 4.76 + \log\frac{0.25}{0.10}\\ &= 4.76 + 0.40\\ &= 5.16\end{aligned}$$

■

WORKED PROBLEM 3.7

Calculate the ratio of lactic acid and lactate required in a buffer system of pH 4.9. The pK_a of lactic acid is 3.86.

SOLUTION

The equation

$$\text{pH} = \text{p}K_a + \log \frac{[\text{lactate}]}{[\text{lactic acid}]}$$

can be rearranged to

$$\log \frac{[\text{lactate}]}{[\text{lactic acid}]} = \text{pH} + \text{p}K_a$$
$$= 4.90 - 3.86 = 1.04$$

Therefore, the required ratio is

$$\frac{[\text{lactate}]}{[\text{lactic acid}]} = \text{antilog } 1.04$$
$$= 11.0$$

For a lactate buffer to have a pH of 5, the lactate and lactic acid components must be present in a ratio of 11:1. A good buffer is a mixture of a weak acid and its conjugate base present in near equal concentrations, and the buffered pH should be within 1 pH unit of the pK_a. Thus, lactate buffer is a poor choice in this situation. With a pK_a of 4.76, the acetate buffer would be a better choice. ■

WORKED PROBLEM 3.8

What is the pH of a solution of 100 ml of 0.01 M H_3PO_4 and 100 ml of 0.01 M Na_3PO_4?

SOLUTION

First, determine what species are present in the solution. The two reagents will react to give a mixture composed of 0.01 mol NaH_2PO_4 (weak acid) and 0.01 mol Na_2HPO_4 (conjugate base).

Using the Henderson–Hasselbalch equation,

$$\text{pH} = \text{p}K_a + \log [\text{A}^-]/[\text{HA}]$$
$$\text{pH} = \text{p}K_a + \log 0.01 \text{ mol}/0.01 \text{ mol}$$
$$\text{pH} = \text{p}K_a + \log 1$$
$$\text{pH} = \text{p}K_a + 0, \textit{ therefore}$$
$$\text{pH} = \text{p}K_a.$$

Because the pK_a of Na_2HPO_4 is 7.2 (refer to Table 3.4), this is also the pH of the solution. ■

WORKED PROBLEM 3.9

During the fermentation of wine, a buffer system consisting of tartaric acid and potassium hydrogen tartrate is produced by a biochemical reaction. Assuming that at some time the

concentration of potassium hydrogen tartrate is twice that of tartaric acid, calculate the pH of the wine. The pK_a of tartaric acid is 2.96.

SOLUTION

$$\begin{aligned}\text{pH} &= \text{p}K_a + \log\frac{[\text{hydrogen tartrate}]}{[\text{tartaric acid}]}\\ &= 2.96 + \log 2\\ &= 2.96 + 0.30\\ &= 3.26\end{aligned}$$

■

WORKED PROBLEM 3.10

What is the pH of a solution prepared by mixing 150 mL of 0.10 M HCl with 300 mL of 0.20 M sodium acetate (NaOAc) and diluting the mixture to 1 L? The pK_a of acetic acid is 4.76.

SOLUTION

The amount of acid present in the solution is found by multiplying the volume of the solution, in milliliters, by M, the molarity of the solution; it is expressed in millimoles (mmol):

$$150\ \text{mL} \times 0.10\ \text{M} = 15\ \text{mmol acid}$$

The amount of sodium acetate is found using the same equation:

$$300\ \text{mL} \times 0.20\ \text{M} = 60\ \text{mmol base}$$

Each mole of HCl will consume 1 mol of sodium acetate and produce 1 mol of acetic acid. This will give 15 mmol of acetic acid, with 45 mmol remaining of sodium acetate (i.e., 60 mmol – 15 mmol). Substituting these values into the Henderson–Hasselbalch equation gives

$$\begin{aligned}\text{pH} &= 4.76 + \log\frac{45}{15}\\ &= 4.76 + \log 3\\ &= 4.76 + 0.48\\ &= 5.24\end{aligned}$$

Because the log term is a ratio of two concentrations, the volume factor can be eliminated and the molar amounts can be used directly. ■

WORKED PROBLEM 3.11

What would be the effect of adding an additional 50 mL of 0.10 M HCl to the solution in

Problem 3.10 before dilution to 1 L?

SOLUTION

Using the same equation as in Problem 3.10, the amount of HCl would be

$$200 \text{ mL} \times 0.10 \text{ M} = 20 \text{ mmol acid}$$

which is also equal to the concentration of acetic acid.

The amount of sodium acetate would be

$$60 \text{ mmol} - 20 \text{ mmol} = 40 \text{ mmol}$$

Substituting into the Henderson–Hasselbalch equation gives

$$\begin{aligned} \text{pH} &= 4.76 + \log\frac{40}{20} \\ &= 4.76 + \log 2 \\ &= 4.76 + 0.30 \\ &= 5.06 \end{aligned}$$

Adding an additional amount of strong acid results in only a small change in pH (from pH 5.24 to 5.06) because of the acetic acid/acetate buffer. ■

WEAK ACIDS WITH MORE THAN ONE IONIZABLE GROUP Some molecules contain more than one ionizable group. Phosphoric acid (H_3PO_4) is a weak polyprotic acid; that is, it can donate more than one hydrogen ion (in this case, three hydrogen ions). During titration with NaOH (**Figure 3.21**) these ionizations occur in a stepwise fashion with 1 proton being released at a time:

The pK_a for the most acidic group is referred to as pK_1. The pK_a for the next most acidic group is pK_2. The third most acidic pK_a value is pK_3.

At low pH most molecules are fully protonated. As NaOH is added, protons are released in the order of decreasing acidity, with the least acidic proton (with the largest pK_a value) ionizing last. When the pH is equal to pK_1, equal amounts of H_3PO_4 and $H_2PO_4^-$ exist in the solution.

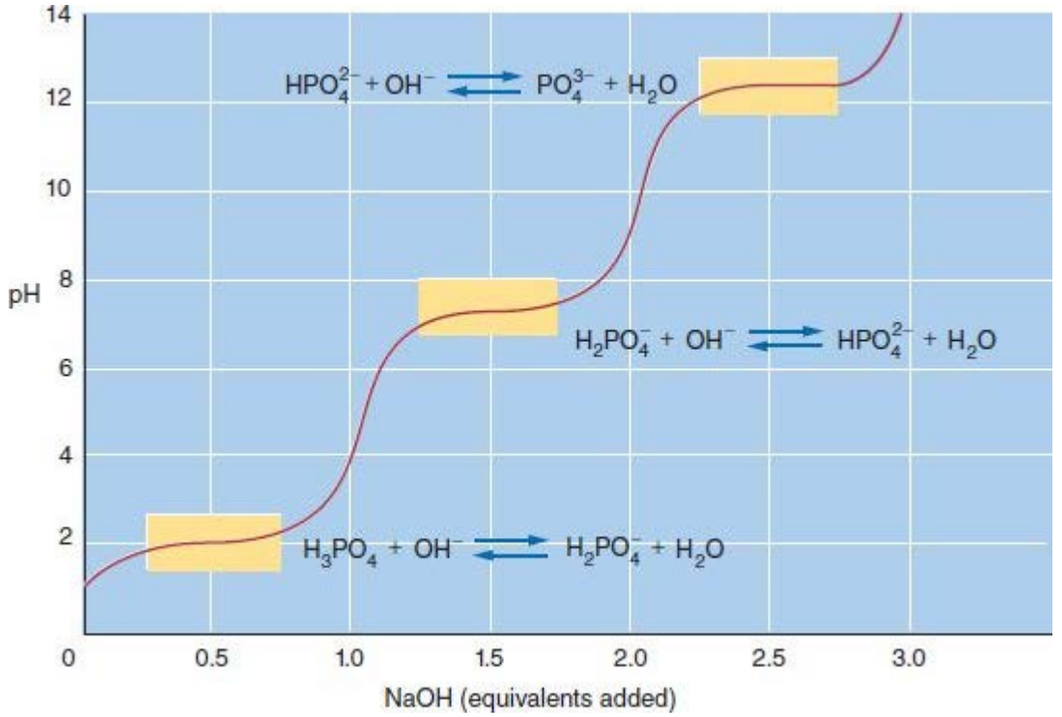
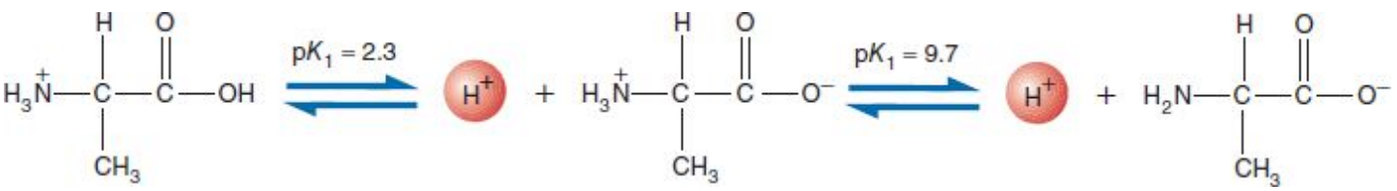


FIGURE 3.21

Titration of Phosphoric Acid with NaOH

Phosphoric acid (H_3PO_4) is a polyprotic acid that releases three protons sequentially upon titration with NaOH.

Amino acids are biomolecules that contain several ionizable groups. Like all amino acids, alanine contains both a carboxyl group and an amino group. At low pH, both of these groups are protonated. As the pH rises during a titration with NaOH, the acidic carboxyl group (COOH) loses its proton to form a carboxylate group (COO^-). The addition of more NaOH eventually causes the ionized amino group to release its proton:



Certain amino acids also possess side chains with ionizable groups. Because of their structures, alanine, lysine, and the other amino acids can act as effective buffers at or near their respective pK_a values [e.g., lysine: pK_1 = 2.0, pK_2 = 9.0, pK_3 (R group) = 10.7]. See Chapter 5 for a description of the titration of the R group of an amino acid.

Physiological Buffers

The three most important buffers in the body are the bicarbonate buffer, the phosphate buffer, and the protein buffer. Each is adapted to solve specific physiological problems in the body.

BICARBONATE BUFFER Bicarbonate buffer, one of the more important buffers in blood, has three components. The first of these, carbon dioxide, reacts with water to form carbonic acid:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3$$
Carbonic acid

Carbonic acid then rapidly dissociates to form H^+ and HCO_3^- ions:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$
Bicarbonate

Because the concentration of H_2CO_3 is very low in blood, the preceding equations may be simplified to

$$CO_2 + H_2O \rightleftharpoons H^+ + HCO_3^-$$

Recall that buffering capacity is greatest at or near the pK_a of the acid–conjugate base pair. Carbonic acid is a diprotic acid (it can donate two hydrogen ions) with a pK_1 of 6.3. In blood, there is a critical need to maintain the pH at the high end of the buffering range of this acid and to maximize buffering capacity for acid. Therefore, it is optimal for the concentration of the conjugate base, bicarbonate, to be high compared to H_2CO_3 (or CO_2), typically 20 to 1. This ratio, which differs from the ideal weak acid:conjugate base ratio of 1:1, indicates that the bicarbonate buffer is operating in blood at the limit of its buffering capacity. Nevertheless, the bicarbonate buffer is effective for two reasons: there is a high bicarbonate concentration in blood, and the components are under physiological control, as described next.

The uncatalyzed conversion of CO_2 to HCO_3^- and H^+ is a slow process:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^- + H^+$$

In blood, this reaction is catalyzed by the enzyme carbonic anhydrase. With rates as high as 10^6 molecules of CO_2 converted to bicarbonate per second per enzyme molecule, carbonic anhydrase is one of the most efficient enzymes known. The CO_2 level is kept low and is regulated through changes in the respiratory rate. The bicarbonate level stays high because the kidneys excrete H^+. When excessive amounts of HCO_3^- are produced, the kidney excretes bicarbonate. As acid, a metabolic waste product, is added to the body's bicarbonate system, the concentration of HCO_3^- decreases and CO_2 is formed. Because the excess CO_2 is exhaled, the ratio of HCO_3^- to CO_2 remains essentially unchanged.

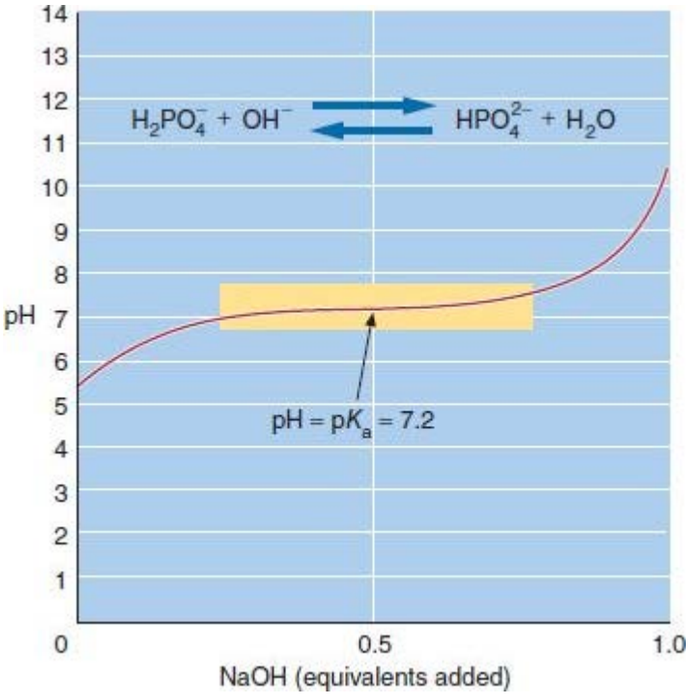


FIGURE 3.22

Titration of $H_2PO_4^-$ by Strong Base

The shaded band indicates the pH range over which the weak acid–conjugate base pair $H_2PO_4^-/HPO_4^{2-}$ functions effectively as a buffer.

PHOSPHATE BUFFER Phosphate buffer consists of the weak acid–conjugate base pair $H_2PO_4^-/HPO_4^{2-}$ (**Figure 3.22**):

$$\underset{\text{Dihydrogen phosphate}}{H_2PO_4^-} \rightleftharpoons H^+ + \underset{\text{Hydrogen phosphate}}{HPO_4^{2-}}$$

With pK_a 7.2, it would appear that phosphate buffer is an excellent choice for buffering the blood. Although the blood pH of 7.4 is well within this buffer system's capability, the concentrations of $H_2PO_4^-$ and HPO_4^{2-} in blood are too low to have a major effect. Instead, the phosphate system is an important buffer in intracellular fluids where its concentration is approximately 75 milliequivalents (mEq) per liter. Phosphate concentration in extracellular fluids such as blood is about 4 mEq/L. Because the normal pH of cell fluids is approximately 7.2 (the range is from 6.9 to 7.4), an equimolar mixture of $H_2PO_4^-$ and HPO_4^{2-} is typically present. Although cells contain other weak acids, these substances are unimportant as buffers. Their concentrations are quite low, and their pK_a values are significantly lower than intracellular pH. For example, lactic acid has a pK_a of 3.86.

PROTEIN BUFFER Proteins are a significant source of buffering capacity. Composed of amino acids linked together by peptide bonds, proteins contain several types of ionizable groups in side chains that can donate or accept protons. Protein molecules are powerful buffers because they are present in significant concentration in living organisms. For example, the oxygen-carrying protein hemoglobin is the most abundant biomolecule in red blood cells. Hemoglobin plays a major role in

maintaining blood pH because of its structure and high cellular concentration. Also present in blood in high concentrations and buffering capacity are the serum albumins and other proteins.

Hemoglobin and pH

KEY CONCEPT

The most important buffers in the body are the bicarbonate buffer (blood), the phosphate buffer (intracellular fluids), and the protein buffer.

QUESTION 3.4

Severe diarrhea is one of the most common causes of death in young children. One of the principal effects of diarrhea is the excretion of large quantities of sodium bicarbonate. In which direction does the bicarbonate buffer system shift under this circumstance? What is the resulting condition called?

Diarrhea

Biochemistry IN PERSPECTIVE

Cell Volume Regulation and Metabolism

Is there a relationship between metabolism and cell volume? Living cells are in constant danger. Even the smallest changes in the balance of solutes between their interiors and their surroundings make cells vulnerable to potentially damaging changes in osmotic pressure. Any inability to manage osmotic balance can lead to distortions in shape and volume that compromise cell function. In multicellular organisms such as animals, however, individual cells are usually not exposed to significant fluctuations in the osmolarity of their surroundings. Instead, it is now realized, they are continuously challenged by internal variations that are created by normal metabolic processes. Routine tasks such as the uptake of nutrients (e.g., sugars, fatty acids, and amino acids), the excretion of waste products (e.g., H^+ and CO_2), and metabolic processes such as the synthesis and degradation of macromolecules (e.g., proteins and glycogen) cause osmotic imbalances.

Research efforts have revealed that cells possess several sophisticated mechanisms that together rapidly correct even the most minor changes in osmolarity. The best understood of these

mechanisms is the exchange of inorganic ions across membranes (**Figure 3A**). For example, when a cell is engaged in the synthesis of protein, the resulting reduction in the concentration of amino acids causes water to flow out of the cell. The cell responds by importing K^+, Na^+, and Cl^- (in exchange for HCO_3^-) through specialized membrane channel complexes. The osmotic gradient created by this process results in the flow of water into the cell, thus restoring the cell's normal volume. When protein is degraded, the opposite process occurs. The increased concentration of osmotically active amino acids causes the cell to swell. Ions (e.g., K^+, Cl^-, and HCO_3^-) followed by water then move across the plasma membrane out of the cell, and cell volume is restored.

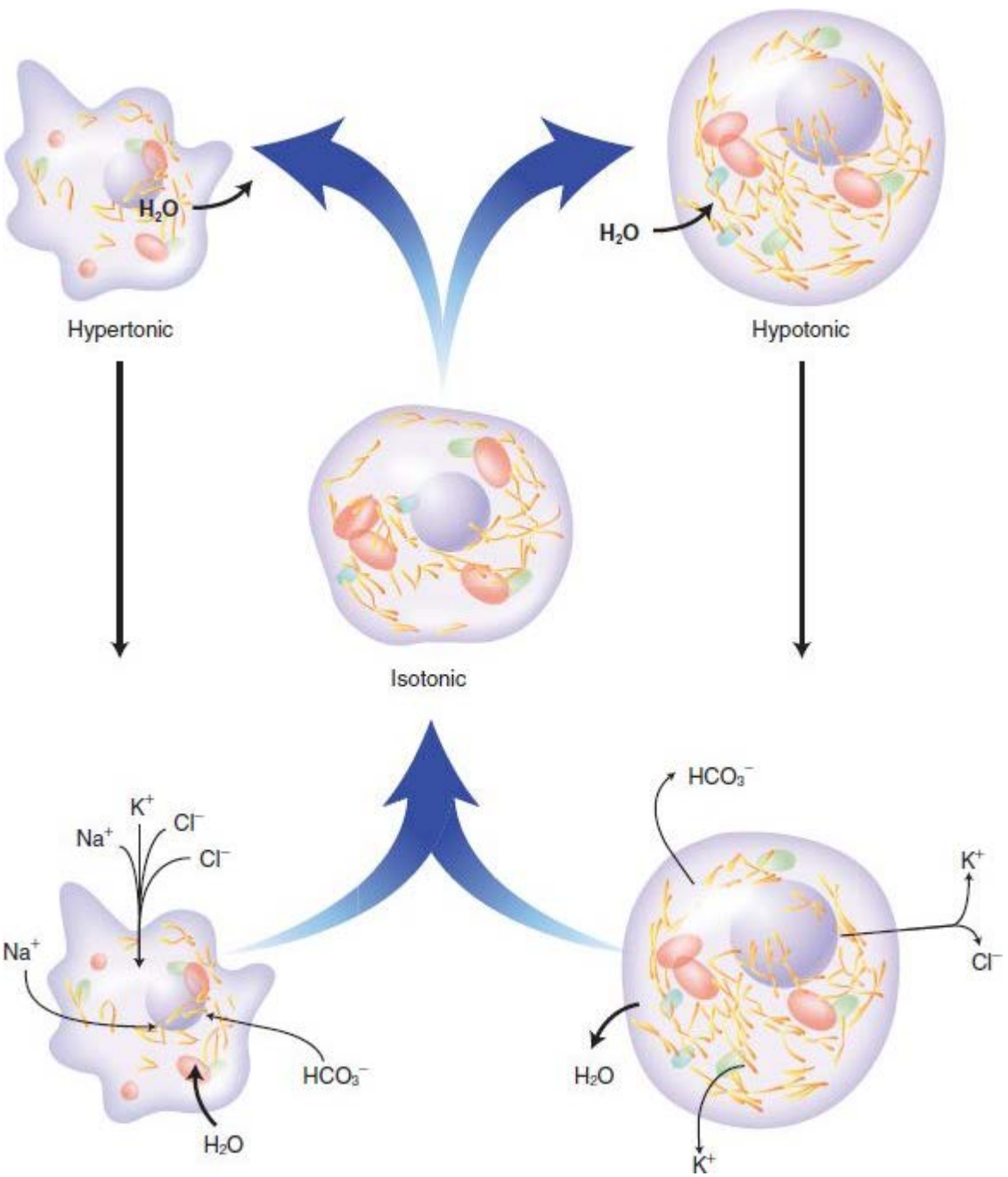


FIGURE 3A

Osmotic Pressure and Cell Volume Changes

Cells shrink when they are exposed to a hypertonic medium or when biochemical processes reduce the number of osmotically active particles. The cell's osmotic balance is restored when inorganic ions such as Na^+, K^+, and Cl^- enter via anion and cation channels and pumps. Ions can exchange with like-charged

ions or be carried along a Na^+ or K^+ gradient. The cell returns to its normal volume as water then flows back into the cell. Cells swell when they are placed in a hypotonic medium or they increase their concentration of osmotically active particles through transport or the degradation of macromolecules. Osmotic balance is restored with the expulsion of inorganic ions, followed by the outflow of water.

Cell volume can also be controlled by the synthesis of quantities of osmotically active substances called *osmolytes*. For example, when confronted with osmotic stress, the cells of some organisms produce large amounts of alcohols (e.g., sorbitol; see p. 253), amino acids, or amino acid derivatives such as taurine (see p. 481). Cells have also been observed to restore osmotic balance by synthesizing or degrading macromolecules such as glycogen. The precise means by which cells manage osmotic balance are not yet resolved. It is known that cell volume changes signaled by distortions of the cytoskeleton cause alterations in the expression of genes, some of which code for the synthesis of membrane channel proteins and osmolytes.

SUMMARY: Living cells are constantly engaged in managing the balance of solutes across their membranes. Nutrient uptake, waste product excretion, and metabolic processes such as macromolecule synthesis affect this balance. Any significant failure to correct imbalances can cause potentially lethal cell volume changes.

Chapter Summary

1. Water molecules (H_2O) are composed of two atoms of hydrogen and one of oxygen. Each hydrogen atom is linked to the oxygen atom by a single covalent bond. The oxygen–hydrogen bonds are polar, and water molecules are dipoles. One consequence of water's polarity is that water molecules are attracted to each other by the electrostatic force between the oxygen of one molecule and the hydrogen of another. This attraction is called a hydrogen bond.
2. Noncovalent bonds are relatively weak and, therefore, easily disrupted. They play a vital role in determining the physical and chemical properties of water and biomolecules. Although each hydrogen bond is not especially strong in comparison to covalent bonds, large numbers of them have a significant effect on the molecules involved. Ionic interactions occur between charged atoms or groups. Van der Waals forces, either attractive or repulsive, occur between permanent and/or induced dipoles.
3. Water has an exceptionally high heat capacity. Its boiling and melting points are significantly higher than those of compounds of comparable structure and molecular weight. Hydrogen bonding is responsible for this anomalous behavior.
4. Water is a remarkable solvent. Water's dipolar structure and its capacity to form hydrogen bonds enable it to dissolve many ionic and polar substances.
5. Most water molecules in living organisms are structured; that is, they are noncovalently associated with macromolecules and membrane surfaces. The networks of water molecules that form act as bridges between macromolecules in the densely crowded cytoplasm. Cytoplasm has the properties of a gel, a semisolid viscoelastic substance that resists flow and stores mechanical energy. Gels can undergo reversible transitions to a liquid or sol state.
6. Hydrophobic molecules are virtually insoluble in water. When nonpolar molecules enter an aqueous environment, they form droplets surrounded by water molecules that rearrange into their most energetically favorable configuration.
7. Amphipathic molecules contain both polar and nonpolar groups. Fatty acids are amphipathic molecules that form structures called micelles when they are placed in water.
8. Several physical properties of liquid water change when solute molecules are dissolved. The most

important of these for living organisms is osmotic pressure, the pressure that prevents the flow of water across cellular membranes. Macromolecules have little direct effect on cellular osmolarity. The large number of ionizable groups on these molecules attracts ions of opposite charge. The structured water network that surrounds macromolecules such as proteins tends to exclude Na^+ because of its relatively large hydrated volume. The charge asymmetry across the cell membrane (negative on the inside and positive on the outside) creates an electrical gradient called a membrane potential.

9. Liquid water molecules have a limited capacity to ionize to form a hydrogen ion (H^+) and a hydroxide ion (OH^-). When a solution contains equal amounts of H^+ and OH^- ions, it is said to be neutral. Solutions with an excess of H^+ are acidic, whereas those with a greater number of OH^- are basic. Because organic acids do not completely dissociate in water, they are referred to as weak acids. The acid dissociation constant K_a is a measure of the strength of a weak acid. Because K_a values vary over a wide range, pK_a values ($-\log K_a$) are used instead.
10. The hydrogen ion is one of the most important ions in biological systems. The pH scale conveniently expresses hydrogen ion concentration. pH is defined as the negative logarithm of the hydrogen ion concentration.
11. Because hydrogen ion concentration affects living processes so profoundly, it is not surprising that regulating pH is a universal and essential activity of living organisms. Hydrogen ion concentration is typically kept within narrow limits. Because buffers combine with H^+ ions, they help maintain a relatively constant hydrogen ion concentration. The ability of a solution to resist pH changes is called buffering capacity. Most buffers consist of mixtures of a weak acid and its conjugate base.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on water to help you prepare for exams.

Chapter 3 Review Quiz

Suggested Readings

Ball P. 2005. Water as an active constituent in cell biology. Chem Rev 108(1):74–108.

Diamond J. 2005. Collapse: how societies choose to fail or succeed. New York (NY): Penguin.

Gerstein M, Levitt M. 1998. Simulating water and the molecules of life. Sci Am 279(5):100–5.

Ho M-W. 2012. Living rainbow H_2O. Singapore: World Scientific.

Kemsley J. 2013. Hydrogen bonds visualized. Chem Engineer News 91(39):5.

Kollipara P. 2013. Diving deeper into souring oceans. Chem Engineer News 91(41):33–5.

Lang F, Waldegger S. 2001. Regulating cell volume. Am Sci 85:456–63.

Leterrier J-F. 2001. Water and the cytoskeleton. Cell Mol Biol (Noisy-le-Grand) 47(5):901–23.

Pollack GH. 2001. Cells, gels and the engine of life. Seattle (WA): Ebner & Sons.

Key Words

acid, *93*
acidosis, *94*
alkalosis, *94*
amphipathic molecule, *88*
base, *93*
buffer, *95*
conjugate base, *93*
covalent bond, *80*
dipole, *79*
electrostatic interaction, *80*
hydrogen bond, *79*
hydrophobic interaction, *87*
hypertonic solution, *90*
hypotonic solution, *90*
isotonic solution, *90*
Le Chatelier’s principle, *95*
London dispersion force, *82*
membrane potential, *92*
micelle, *88*
osmolyte, *104*
osmosis, *89*
osmotic pressure, *90*
pH scale, *94*
polar, *79*
salt bridge, *81*
solvation sphere, *85*
van der Waals force, *82*
weak acid, *93*
weak base, *93*

Review Questions

SECTION 3.1

Comprehension Questions

1. Define the following terms:
    a. dipole
    b. hydrogen bond
    c. electrostatic interaction
    d. covalent bond
    e. hydration

2. Which of the following molecules have a dipole moment?

    a. CCl_4
    b. $CHCl_3$
    c. H_2O
    d. CH_3OCH_3
    e. CH_3CH_3
    f. H_2

Fill in the Blanks

3. The largest atom in a water molecule is _____.
4. In water, two of the four _____ orbitals of oxygen are occupied by two lone pairs of electrons.
5. _____ occur between any two opposite partial charges or full charges.
6. The bond angle between the two hydrogens in water molecules is _____.

Short-Answer Questions

7. Describe why hydrogen bonds form around water molecules.
8. Why do the hydrogens and oxygens of water molecules have partial charges?
9. List four biological processes on which water is an indispensable component.
10. Why are CO_2 molecules not dipoles?

Critical-Thinking Question

11. Some of the water on Earth was formed by reactions of iron and silicon oxides and methane. Write the balanced equation for the reaction of ferric oxide and methane.

SECTION 3.2

Comprehension Questions

12. Define the following terms:

    a. salt bridges
    b. ionic interactions
    c. van der Waals forces
    d. dipole–dipole interaction
    e. London dispersion forces

13. What interactions occur between the following molecules and ions?

    a. water and ammonia
    b. lactate and ammonia ion
    c. benzene and octane
    d. carbon tetrachloride and chloroform
    e. chloroform and diethyl ether

Fill in the Blanks

14. The _____ interactions that occur between charged atoms are nondirected.
15. A _____ dipole induces a transient dipole.
16. The stacking of bases in DNA is an example of _____ interactions.
17. The attraction between a carbonyl-containing molecule and an aromatic ring is a form of _____ interactions.
18. The relative ease with which _____ respond to an electric field determines the magnitude of van der Waals forces.

Short-Answer Questions

19. A water molecule can form hydrogen bonds with a maximum of four other molecules. Draw this interaction.
20. How many hydrogen bonds form between methanol molecules? Draw this interaction.

Critical-Thinking Questions

21. Examine the elements in the second row of the periodic table. Note that HF, H_2O, NH_3 and CH_4 are sp^3 hybridized. The first three are all capable of hydrogen bonding. Methane, however, is not capable of hydrogen bonding. Explain.
22. Chloroform ($HCCl_3$) is capable is capable of weak hydrogen bonding. Explain.
23. Water forms stronger hydrogen bonds than ammonia. Suggest a reason for this.

SECTION 3.3

Comprehension Question

24. Define the following terms:
    a. heat of fusion
    b. heat of vaporization
    c. heat capacity
    d. boiling point of water
    e. freezing point of water

Fill in the Blanks

25. _____ is responsible for the unusual thermal properties of ammonia.
26. The _____ is defined as the energy required to melt ice.
27. The _____ is defined as the energy required to vaporize one mole of a liquid at a pressure of one atmosphere.
28. _____ is defined as the energy that must be added or removed to change the temperature by one degree Celsius.

Short-Answer Questions

29. Calculate how much energy is required to raise the temperature of methanol (C = 2.5 J/g °C) by 10°C.

Critical-Thinking Questions

30. Explain why ice is less dense than water. If ice were not less dense than water, how would the oceans be affected? How would the development of life on Earth be affected?
31. The heat absorbed or liberated by a substance (q) can be calculated using the following equation: $q = g \quad C \quad \Delta T$, where g is the mass in grams, C is the heat capacity per unit mass, and ΔT is the change in temperature. Use the following values to calculate the energy required to convert one gram of ice at 0°C to one gram of steam at 100°C. The heat capacity of water is 4.25 J/g °C. The heat of fusion (the amount of heat required to change a solid into a liquid at its melting point) of ice is 335 J/g. The heat of vaporization of water is 2258 J/g.
32. Calculate the energy required to convert solid hydrogen sulfide (H_2S) to a gas. The heat capacity of H_2S is 1.03 J/g °C. Its heats of fusion and vaporization are 69.9 and 549 J/g, respectively. Compare your answer with the value for water.

SECTION 3.4

Comprehension Questions

33. Define the following terms:

    a. solvation sphere
    b. amphipathic
    c. micelle
    d. hydrophobic effect
    e. clathrate

34. Define the following terms:

    a. osmosis
    b. osmotic pressure
    c. membrane potential
    d. isotonic solution
    e. structured water

35. Define the following terms:

    a. osmolarity
    b. sol
    c. gel
    d. ectoplasm
    e. endoplasm

36. Define the following terms:

    a. hydrocarbon tails
    b. polar head

c. dialyzing membrane
d. van’t Hoff factor
e. osmometer

37. Define the following terms:

a. semipermeable membrane
b. hypotonic
c. hypertonic
d. crenation
e. hemolysis

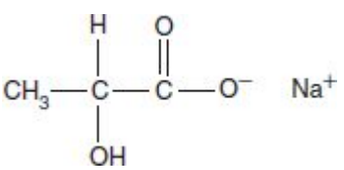
Fill in the Blanks

38. Osmotic pressure is measured with a _____.
39. _____ are shells of water molecules that cluster around ions.
40. The capacity of a solvent to reduce the electrostatic attraction between charges is measured by its _____.
41. _____ water molecules are noncovalently associated on the surfaces of macromolecules and membranes.
42. A _____ is a type of mixture in which small particles are evenly distributed throughout another substance.
43. Amoeboid motion is made possible by highly regulated _____ transitions.
44. A semisolid viscoelastic substance that resists flow and stores mechanical energy is called a _____.

Short-Answer Questions

45. A solution containing 56 mg of a protein in 30 ml of distilled water exerts an osmotic pressure of 0.01 atm at $T = 25°C$. Determine the molecular weight of the unknown protein.
46. A dialysis bag containing a 3 M solution of the sugar fructose is placed in the following solutions. In each case, give the direction in which water flows.

a. 1 M sodium lactate
b. 3 M sodium lactate
c. 4.5 M sodium lactate



What is the osmolarity of a 1.3 M solution of sodium phosphate (Na_3PO_4)? Assume 85% ionization for this solution.

47. Considering the hydration spheres of sodium and potassium ions, would you expect lithium ions to easily penetrate the cell or remain outside the cell?

Critical-Thinking Questions

49. Gelatin consists of collagen that is suspended in and hydrated by water. What happens to

gelatin when NaCl is sprinkled over its surface?

50. Gelatin is composed of a small amount of protein in a large amount of water. Explain why gelatin is a solid.
51. A well is punched into a slab of gelatin, and a solution of equimolar amounts of NaCl and KCl is placed in it. After an hour, the concentration of Na^+ and K^+ present in the well is measured. Which ion, Na^+ or K^+, would be present in greatest abundance in the well? Explain your answer.
52. The sodium and potassium ions have identical charges, yet that hydrated volume of Na^+ is 3.4 times that of K^+. Explain.
53. In many cells that can survive severe dehydration, certain sugars replace water. These sugars interact with and protect membrane surfaces and prevent protein aggregation. What structural feature of the sugar molecules is responsible for this phenomenon?
54. Water has been described as the universal solvent. If this statement were strictly true, could life have arisen in a water medium? Explain.
55. A cell's ATP-driven sodium-potassium pump fails. Will the cell shrivel or rupture?
56. Alcohols (ROH) are structurally similar to water. Why are alcohols not as powerful a solvent as water for ionic compounds? [*Hint*: Methanol is a better solvent for ionic compounds than is propanol.]
57. Many fruits can be preserved by candying. The fruit is immersed in a highly concentrated sugar solution, and then the sugar is allowed to crystallize. How does the sugar preserve the fruit?
58. During stressful situations, some cells in the body convert glycogen to glucose. What effect does this conversion have on cellular osmotic balance? Explain how cells handle this situation.
59. Why can't seawater be used to water plants?

SECTION 3.5

Comprehension Questions

60. Define the following terms:
    a. hydronium ion
    b. acid
    c. base
    d. K_a
    e. pK_a
61. Define the following terms:
    a. buffer
    b. acidosis
    c. alkalosis
    d. Le Chatelier's principle
    e. hydroxide ion
62. Define the following terms:
    a. Henderson–Hasselbalch equation
    b. bicarbonate ion
    c. protein buffer

d. phosphate buffer
e. osmolytes

Fill in the Blanks

63. The rule concerning chemical processes stating that stress applied to a reaction at equilibrium results in the displacement of the equilibrium so as to relieve the stress is referred to as _____.
64. The ability of a solution to resist pH changes is referred to as its _____.
65. Osmotically active substances are called _____.
66. The conjugate acid of water is _____.
67. A buffer is composed of a weak acid and its _____ base.
68. When blood pH rises above 7.45, _____ results.

Short-Answer Questions

69. Why are the bicarbonate and phosphate buffers the main buffers in the blood and cells, respectively, and not vice versa?
70. If the total concentration of a buffer is known (and not the individual concentrations of the weak acid and its conjugate base), can the pH be calculated?
71. Describe how you can increase the buffering capacity of a 0.1 M acetate buffer.
72. What would be the pH of 1 L of water if 1 ml of 1 M HCl were added?
73. Calculate the ratio of dihydrogen phosphate to hydrogen phosphate in blood at pH 7.4. The K_a is 6.3×10^{-8}.
74. Calculate the pH of a solution prepared by mixing 300 ml of 0.25 M sodium hydrogen ascorbate and 150 ml of 0.2 M HCl. The pK_{a1} of ascorbic acid is 4.04.

Critical-Thinking Questions

75. Consider the following ion series:

$$Mg^{2+} > Ca^{2+} > Na^{+} > K^{+} > Cl^{-} > NO_3^{-}$$

Ions to the left are more strongly hydrated than ions to the right. Indicate whether ions such as Mg^{2+} and Cl^{-} would move easily into the structured water that is associated with cellular macromolecules.

76. When acetic acid ionizes under normal conditions, both the departing proton and the acetate anion are solvated by water molecules. In the absence of water, would you expect the pK_a of acetic acid to be larger or smaller? Explain your answer.
77. The pH scale is only valid for water. Explain.
78. Explain how the acids produced in metabolism are transported to the liver without greatly affecting the pH of the blood.

MCAT Study Questions

79. Noncovalent bonds include all of the following except:

a. van der Waals interaction
b. ionic bonds

c. carbon–oxygen double bond
d. hydrogen bond

80. Which of the following bonding types is the most important in the nonaqueous interfaces between protein subunits?

a. van der Waals interactions
b. the hydrophobic effect
c. ionic bonds
d. disulfide bridges

81. Calculate the molecular mass of nonionic compound X given the following information: One gram of compound X is dissolved in 100 ml of water. The solution has an osmotic pressure of 0.4 atm at 25°C.

a. 16.3×10^3 g/mol
b. 8.1×10^2 g/mol
c. 6.1×10^2 g/mol
d. Not enough information is given.

82. What is the osmotic pressure when 0.3 g of nonelectrolyte compound A (MW 75) is diluted to 100 ml of water at a temperature of 25°C?

a. 1 atm
b. 10 atm
c. 0.86 atm
d. 8.6 atm

83. Hydrogen bonds are a type of _____ bond or interaction.

a. ionic
b. dipole–dipole
c. dipole-induced dipole
d. London dispersion

CHAPTER 4 Energy

Energy Transformation Animals convert the chemical bond energy in food molecules into the energy required to sustain life. This energy is used for growth, repair, reproduction, and activities such as seeking food. Bioenergetics is the branch of biochemistry that investigates how cells transform and utilize energy.

OUTLINE

ENERGY AND LIFE'S DEEP, DARK SECRETS

4.1 THERMODYNAMICS
First Law of Thermodynamics
Second Law of Thermodynamics

4.2 FREE ENERGY
Standard Free Energy Changes
Coupled Reactions
The Hydrophobic Effect Revisited

4.3 ENERGY CURRENCY MOLECULES
Energy and Reducing Power

Biochemistry in Perspective
Nonequilibrium Thermodynamics and the Evolution of Life

AVAILABLE ONLINE:

Biochemistry in Perspective
The Extremophiles: Organisms That Make a Living in Hostile Environments

Energy and Life's Deep, Dark Secrets

Chris and Steve are microbiologists on a mission: to retrieve rock from the sizzling depths of a South African gold mine (**Figure 4.1**). Wearing coveralls, boots, and hard hats fitted with lamps, and carrying their sterilized tools and specimen bags and a 4-L supply of water, the two scientists descend along with the miners 3 km into the hot, oppressive darkness of the deepest excavation on Earth. Chris and Steve then work quickly to remove rock samples. The threat of heat stress limits them to only four hours in the mine because mine temperatures often exceed 49°C. Work is often interrupted by the removal of liters of sweat from boots and gloves. Their goal is to find microbes with unique biochemical capabilities (e.g., radiation resistance), most notably for use in *bioremediation* (microbes used to remove toxic substances from contaminated soil or water).

How and why did Chris and Steve find their way into such a hostile workplace? Recent research efforts have overturned the longstanding conviction that the biosphere is limited largely to Earth's surface. Not only is the hot and toxic planet's crust suffused with life, but also its biochemically diverse microbial inhabitants have vital roles in the biogeochemical cycles that maintain the health of the entire biosphere.

Biogeochemical cycles are pathways driven by solar and geothermal energy in which chemical elements move throughout Earth's biotic (biosphere) and abiotic compartments [*lithosphere* (crust and upper mantle), *hydrosphere* (surface water, such as lakes, rivers, and oceans), and atmosphere]. Because Earth's supply of elements is fixed and finite, the cyclic transport and chemical transformations of elements are critical for living organisms. In the carbon cycle, for example, CO_2 is incorporated into organic biomolecules by light-driven photosynthesis, and animals use O_2-based respiration to convert consumed plant biomolecules back into CO_2, which is released (along with CO_2 generated from the decomposition of dead organisms and fossil fuel combustion) into the atmosphere. Carbon is also incorporated into carbonate minerals in a solar energy-driven process in which CO_2 and water react with silicate minerals (e.g., $CaSiO_3$) to form calcium carbonate ($CaCO_3$) and silicon dioxide (SiO_2). As the result of erosion, $CaCO_3$ and other minerals are then washed into the oceans where they become buried in sediments that also contain the calcium carbonate–containing shells of a variety of dead marine organisms along with organic detritus formed from decomposing organisms. With life spans of thousands of years, subseafloor microbes slowly decompose this organic matter, releasing CO_2 into ocean water for reuse in photosynthesis or evaporation into the atmosphere.

FIGURE 4.1

To Hell and Back

In their investigations of some of the Earth's most inaccessible microorganisms, scientists must endure searing heat in deep mines where rock temperatures can reach 60°C (140°F) and the ever-present threat of dehydration. Research into the unique biochemistry of organisms that survive these conditions has revealed that some microbes generate energy by reducing geologically produced sulfate with hydrogen released when water molecules are split by radiation emitted from radioactive elements such as uranium and thorium.

Although it is well known that CO_2 is also released into the atmosphere via volcanism, the role of subsurface microorganisms has only been appreciated recently. Volcanoes are the inevitable result of the slow recycling of oceanic crust tectonic plates that is driven by geothermal energy (heat energy partially generated by radioactive element decay). New oceanic crust is formed at a mid-ocean ridge (the boundary between crust and continental plates) by the upwelling of magma (molten rock) from the underlying mantle. As the magma seeps through the cracks between the two plates, it solidifies and displaces existing oceanic crust. At the opposite end of each of the two enlarging oceanic plates, there is a collision with a continental plate in which the denser crust slab and any surface sediment will sink (subduct) beneath the continental plate. Trapped by the subducting slab, the sediment microbes continue to release CO_2 from any remaining organic matter until the slab's temperature becomes incompatible with life. As subduction proceeds, crust melts to form magma as it is forced deeper into the Earth. Some of the melted rock flows into magma chambers of volcanoes near crustal plate edges. The intense heat within the mantle converts $CaCO_3$ back into $CaSiO_3$ and CO_2, which then escapes with other volcanic gases into the atmosphere, thereby completing the cycle.

Overview

ENERGY! IT IS CERTAINLY ESSENTIAL TO LIFE, BUT WHAT IS IT AND WHY IS IT SO VITAL TO LIVING ORGANISMS? ENERGY IS THE BASIC CONSTITUENT of the universe. The relationship between matter and its energy equivalent is defined by Einstein's famous equation $E = mc^2$. In other words, energy and matter are interconvertible: matter is condensed energy. The total energy (E) in joules ($kg \cdot m^2/s^2$) in a particle is equal to the mass (m) in kilograms of the particle multiplied by the speed of light (c = 3.0 × 108 m/s) squared. Energy is more commonly defined, however, as the capacity to do work. **Work** is organized molecular motion that causes the displacement or movement of an object by the application of force and results in a specific physical change (e.g., the force of flowing water turns the blades of a turbine in a hydroelectric plant). Energy comes in many interconvertible forms: gravitational, nuclear, radiant, chemical, mechanical, electrical, and thermal (heat).

Electromagnetic radiation, electrical energy, and chemical energy are high-quality energy sources, whereas heat is lower-quality energy. Compare, for example, the work potential of electricity as it enters a building carried by electrical wire and the heat energy radiating from a light bulb. In the right circumstances, however, heat can be useful energy. For example, heat flow from the Earth's core to the mantle and within the mantle by the movement of magma is an important contributing force in biogeochemical cycles.

Energy flows continuously through the biosphere. Originating as either solar or geothermal energy, it drives the flow of matter (e.g., nutrients) and biochemical processes in living organisms. Living organisms make use of three energy-generating mechanisms: photosynthesis, chemoorganotrophy,

and chemolithotrophy. Photosynthesis (Chapter 13) is a process that converts light energy into chemical bond energy (ATP). Chemoorganotrophs and chemolithotrophs generate ATP by oxidizing organic and inorganic compounds, respectively. All of these methods of capturing and transforming energy involve oxidation-reduction (redox) reactions (described in Chapter 9) in which electrons are transferred from an **electron donor** to an **electron acceptor**. Living organisms use the energy provided by ATP to power thousands of molecular machines. Work performed by these machines includes maintenance of concentration gradients and the synthesis of biomolecules.

The investigation of energy transformations that accompany physical and chemical changes in matter is called **thermodynamics**. **Bioenergetics**, a branch of thermodynamics, is the study of energy transformations in living organisms. It is especially useful in determining the direction and extent to which specific biochemical reactions occur. These reactions are affected by three factors. Two of these, **enthalpy** (total heat content) and **entropy** (disorder), are related to the first and second laws of thermodynamics, respectively. The third factor, called **free energy** (energy available to do chemical work and a measure of the spontaneity of chemical reactions), is explained by a mathematical relationship between enthalpy and entropy.

The chapter begins with some basic thermodynamic concepts and their relationship to biochemical reactions. This is followed by a discussion of free energy, a useful measure of the spontaneity of biochemical reactions. The chapter ends with a review of the structure and function of ATP and other high-energy compounds.

4.1 THERMODYNAMICS

The modern concept of energy is an invention of the Industrial Revolution. In the nineteenth century, investigations of the relationship between mechanical work and heat led to the discovery of a set of rules called the *laws of thermodynamics* that describe energy transformations:

1. **The first law of thermodynamics:** The total amount of energy in the universe is constant. Energy can be neither created nor destroyed, but it can be transformed from one form into another.
2. **The second law of thermodynamics:** The disorder of the universe always increases. Chemical and physical processes occur spontaneously only when the disorder of the universe increases.
3. **The third law of thermodynamics:** As the temperature of a perfect crystalline solid approaches absolute zero (0 K), disorder approaches zero.

The first two laws are powerful tools that biochemists use to investigate the energy transformations in living systems.

Thermodynamics considers heat and energy transformations in a “universe” composed of a system and its surroundings (**Figure 4.2**). A system is defined according to the interests of the investigator: an entire organism, or a single cell, or a reaction occurring in a flask. In an *open system*, matter and energy are exchanged between the system and its surroundings. If only energy can be exchanged with the surroundings, then the system is said to be *closed*. Living organisms, which consume nutrients from their surroundings and release waste products into it, are open systems.

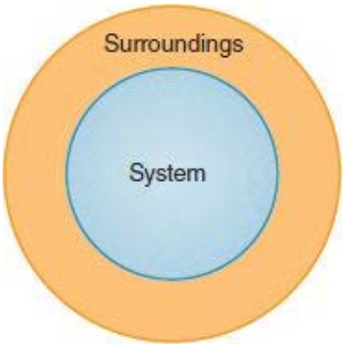


FIGURE 4.2

A Thermodynamic Universe

A universe consists of a system and its surroundings.

Thermodynamic functions include enthalpy, entropy, and free energy. Knowledge of these functions enables biochemists to predict whether a process is spontaneous (*thermodynamically favorable*). Spontaneity alone does not indicate that a reaction will occur, only that it can occur under the right set of conditions. Reactions occur only if there is sufficient energy available to the system. These reactions are described as *kinetically favorable*.

Several thermodynamic properties are *state functions*, which depend only on their initial and final states. State functions are independent of the pathway taken to get from the initial state to the final state. For example, the energy contained in a glucose molecule is inherent in its molecular structure. How the energy of glucose is distributed when it is degraded, however, is not fixed but is governed by the system or pathway undergoing change. Living cells use some of the energy in glucose molecules to perform cellular work such as muscle contraction. The remainder is released as disordered heat energy. If glucose molecules are instead ignited in a laboratory dish, the overall reaction is the same. However, all of the chemical bond energy in the glucose is transformed directly into heat, and little or no measurable work is performed. The energy content of the glucose molecules is the same in each process. The work accomplished by each process is different. In other words, work and heat are not state functions; their values vary with the pathway.

The exchange of energy between a system and its surroundings can occur in only two ways. Heat (q), random molecular motion, may be transferred to or from the system. Alternatively, the system may do work (w) on its surroundings or have work done on it by its surroundings. Energy is transferred as heat when the system and its surroundings are at different temperatures. Energy is transferred as work when an object is moved by force.

First Law of Thermodynamics

The first law of thermodynamics expresses the relationship between the internal energy (E) of a closed system and the heat (q, or disorganized motion) and work (w, or organized motion) transferred between the system and its surroundings (**Figure 4.2**). It is an alternative statement of the law of *conservation of energy*: the total energy of an isolated system (e.g., our universe) is constant. In other words, for a closed system

$$\Delta E = q + w \tag{1}$$

where ΔE = the change in energy of the system
q = the heat absorbed or released by the system
w = the work done by or to the system

Chemists define enthalpy (H), a measure of the system's internal energy:

$$H = E + PV \quad (2)$$

where PV = pressure-volume work, that is, the work done on or by a system that involves changes in pressure and volume

In biochemical systems, pressure is nearly constant and volume changes are negligible. Changes in enthalpy are then essentially equal to changes in internal energy:

$$\Delta H = \Delta E \quad (3)$$

If ΔH is negative ($\Delta H < 0$), the reaction or process gives off heat and is referred to as **exothermic**. If ΔH is positive ($\Delta H > 0$), heat is absorbed from the surroundings, and the process by which it is emitted is called **endothermic**. In **isothermic** processes ($\Delta H = 0$), heat is not exchanged with the surroundings.

Equation (3) indicates that the total energy change of a biological system is equivalent to the heat evolved or absorbed by the system. Because the enthalpy of a reactant or product is a state function (independent of pathway), then the enthalpy change for a particular reaction forming that substance can be used to calculate the ΔH of any other reaction involving that substance. If the sum of the ΔH values ($\sum \Delta H$) for both the reactants and the products is known, then the enthalpy change for the reaction can be calculated using the following equation:

$$\Delta H_{\text{reaction}} = \Delta H_{\text{products}} - \Delta H_{\text{reactants}} \quad (4)$$

The standard enthalpy of formation per mole (25°C, 1 atm), symbolized by $\Delta H_f°$, is commonly used in enthalpy calculations; $H_f°$ is the energy evolved or absorbed when 1 mol of a substance is formed from its most stable elements under standard conditions. Note that Equation (4) cannot predict the direction of any chemical reaction. It determines only the heat flow. Problems 4.1 and 4.2 give standard enthalpy calculations.

KEY CONCEPTS

- At constant pressure, a system's enthalpy change ΔH is equal to the flow of heat energy.
- If ΔH is negative, the reaction or process is exothermic. If ΔH is positive, the reaction or process is endothermic. In isothermic processes, no heat is exchanged with the surroundings.

WORKED PROBLEM 4.1

Given the following $\Delta H_f°$ values, where $\Delta H_f°$ is the energy change required to produce a compound from its elements, calculate $\Delta H_f°$ for the reaction

$$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2$$

	$\Delta H_f°$	
	kcal/mol	**kJ/mol**
$C_6H_{12}O_6$	−304.7	−1274.9
CO_2	−94.0	−393.3

H_2O	−68.4	−286.2
O_2	0	0

The units in the table have the following definitions: 1 kcal is the energy required to raise the temperature of 1000 g of water 1°C; the joule (J) is a unit of energy that is replacing the calorie (cal) in scientific usage (1 cal = 4.184 J).

SOLUTION

The total enthalpy for a reaction is equal to the sum of enthalpy values of the products minus those of the reactants.

$$\begin{array}{rcl} 6CO_2 + 6H_2O & \rightarrow & C_6H_{12}O_6 + 6O_2 \\ 6(-393.3) + 6(-286.2) & & -1274.9 + 6(0) \\ -2359.8 + (-1717.2) & & -1274.9 \end{array}$$

$$\Delta H = -1274.9 - (-4077.0) = 2802.1 \text{ kJ/mol}$$

The positive ΔH indicates that the reaction is endothermic. ■

WORKED PROBLEM 4.2

Given the following data, calculate the change in mileage of a car when it is converted from using gasoline (*n*-octane) as fuel to ethanol. Assume that your car has a mileage rating of 7.92 mi/L (30 mi/gal) of gasoline and that it burns gasoline with 100% efficiency to yield only CO_2 and H_2O.

ΔH_f (kj/mol)		**Density (g/mL)**		**Mass (g/mol)**
CO_2	−393.5	CH_3CH_2OH	0.80	46.1
H_2O	−285.8	Octane	0.70	114.2
CH_3CH_2OH	−277.7			
Octane	−250.1			

SOLUTION

1. Calculate the energy evolved per mole for each type of fuel molecule.

$$\begin{array}{lccccc} CH_3CH_2OH & + & O_2 & \rightarrow & 2CO_2 & + & 3H_2O \\ -277.7 & & 0 & & 2(-393.5) & & 3(-285.8) \end{array}$$

$$-277.7 \qquad -787.0 \qquad -857.4$$

$$\Delta H = -1644.4 - (-277.7) = -1366.7 \text{ kJ/mol}$$

$$\begin{array}{lccccc} C_8H_{18} & + & O_2 & \rightarrow & 8CO_2 & + & 9\,H_2O \\ -250.1 & & 0 & & 8(-393.5) & & 9(-285.8) \end{array}$$

$$-250.1 \qquad -3148.0 \qquad -2572.2$$

$$\Delta H = -5720.2 - (-250.1) = -5470.1 \text{ kJ/mol}$$

2. Calculate the number of moles per liter for each fuel burned.

CH_3CH_2OH (0.80 g/mL)(1000 mL)/46.1 g/mol = 17.4 mol
n-Octane (0.70 g/ml)(1000 ml)/114.2 g/mol = 6.1 mol

3. Calculate the amount of energy produced per liter of each fuel and the change in mileage when the fuel is switched from gasoline to ethanol.

CH_3CH_2OH (17.4 mol)(1366.7 kJ/mol) = −23,780.6 kJ
n-Octane (6.1 mol)(5470.1 kJ/mol) = −33,367.6 kJ
Relative heat production (ethanol/*n*-octane) = 23,780.6kJ/−33,367.6kJ
= 0.7

Assuming that a car using gasoline has a mileage rating of 7.9 mi/L, then switching to ethanol would give (0.7)(7.9 mi/L) = 5.5 mi/L or 21.0 mpg. ■

Second Law of Thermodynamics

The first law accounts for the energy changes that can occur during a process, but it cannot predict to what extent the process will occur. In some circumstances, whether processes occur appears obvious: for example, the behavior of ice at room temperature or gasoline in an internal combustion engine. Experience tells us that ice melts at temperatures above 0°C and that gasoline molecules can be converted to energy in the presence of oxygen to form CO_2 and H_2O. Physical or chemical changes that occur with the release of energy are said to be **spontaneous**. Nonspontaneous processes are those that occur when a constant input of energy is required to support a change. Experience convinces us that certain processes will not occur: ice will not form at temperatures above 0°C, and gasoline molecules cannot form from an engine's exhaust fumes. In other words, we intuitively understand that there is a direction to these processes and that we can easily predict their outcome. When experience cannot be relied on to allow us to make predictions concerning spontaneity and direction, the second law can be used. According to the second law, all spontaneous processes occur in the direction that increases the total disorder of the universe (a system and its surroundings) (**Figure 4.3**). As a result of spontaneous processes, matter and energy become more disorganized. Gasoline molecules, for example, are hydrocarbons in which carbon atoms are linked in an orderly arrangement. When gasoline burns, the carbon atoms in the gaseous products are randomly dispersed (**Figure 4.4**). Similarly, the energy that is released as gasoline burns becomes more disordered; it becomes less concentrated and less useful. In a car engine, increased gas pressure in the cylinders drives the pistons and causes the car to move. When we compare the chemical energy in the gasoline molecules and the kinetic energy that moves the car, it becomes apparent that a significant amount of energy does no useful work. Rather, it is dissipated (dispersed) into the surroundings, producing a

hot engine and exhaust fumes.

A system's degree of disorder is measured by the state function called *entropy* (S). The more disordered a system is, the greater is its entropy value. According to the second law, the entropy change of the universe is positive for every spontaneous process. The increase may take place in any part of the universe, either the system or its surroundings (ΔS_{sys} or ΔS_{surr}):

$$\Delta S_{univ} = \Delta S_{sys} + \Delta S_{surr}$$

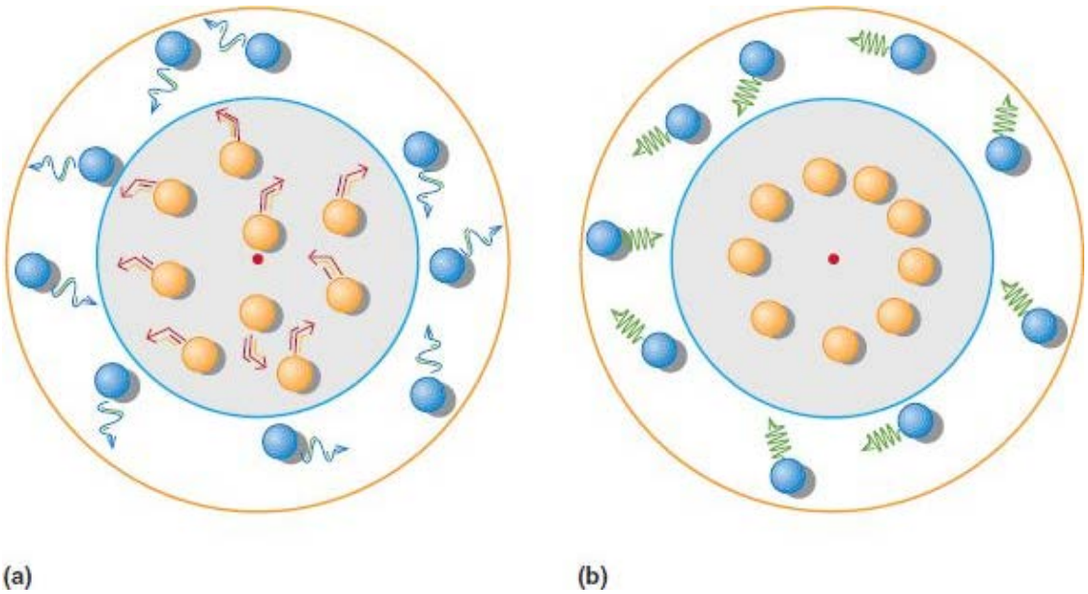


FIGURE 4.3

A Living Cell as a Thermodynamic System

(a) The molecules of the cell and its surroundings are in a relatively disordered state. (b) Heat is released from the cell as a consequence of reactions that create order among the molecules inside the cell. This energy increases the random motion, and therefore the disorder, of the molecules outside the cell (indicated by tighter springs on the outer molecules). This process causes a net positive entropy change. The cell's decrease in entropy is more than offset by an increase in the entropy of the surroundings.

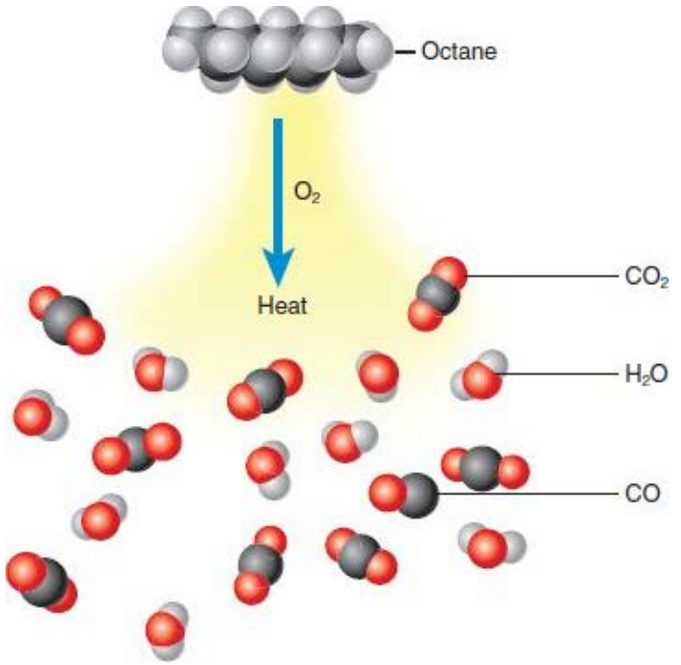


FIGURE 4.4

Gasoline Combustion

When hydrocarbons such as octane are burned, the release of energy is accompanied by the conversion of highly ordered reactant molecules into relatively disorganized gaseous products such as CO_2 and H_2O. However, gasoline combustion is inefficient; that is, other substances such as the environmental pollutant carbon monoxide (CO) are also released.

Living cells do not increase their internal disorder when they consume and metabolize nutrients. Instead, the organism's surroundings increase in entropy. For example, humans consume food to provide the energy and structural molecules needed to maintain their complex bodies. The food molecules are converted into vast amounts of disordered waste products (e.g., CO_2, H_2O, and heat) that are discharged into their surroundings.

Although entropy may be considered unusable energy, the formation of entropy is not a useless activity. Some reactions are said to be entropy-driven because the increase in entropy in the system overrides a gain in enthalpy to result in a spontaneous reaction. (By definition, a spontaneous process will occur. The rate at which it occurs, however, may be very rapid or very slow.) In irreversible processes, processes that proceed in only one direction, entropy and enthalpy are driving forces. Entropy directs a system toward equilibrium with its surroundings. Once a process has reached equilibrium (i.e., there is no net change in either direction), there is no longer any driving force to propel it.

To predict whether a process is spontaneous, the sign of ΔS_{univ} must be known. For example, if the value of ΔS_{univ} for a process is positive (i.e., the entropy of the universe increases), then the process is spontaneous. If ΔS_{univ} is negative, the process does not occur, but the reverse process takes place spontaneously. If ΔS_{univ} is zero, neither process tends to occur. Organisms that are at equilibrium with their surroundings are dead.

KEY CONCEPTS

- The second law of thermodynamics states that the universe tends to become more disorganized.
- Entropy increases may take place anywhere in the system's universe.
- For processes in living organisms, the increase in entropy takes place in the surroundings.

4.2 FREE ENERGY

Although the entropy of the universe always increases in a spontaneous process, measuring it is often impractical because both ΔS_{sys} and ΔS_{surr} must be known. A more convenient thermodynamic function for predicting the spontaneity of a process is free energy, which can be derived from the expression for ΔS_{univ}:

$$\Delta S_{univ} = \Delta S_{surr} + \Delta S_{sys}$$

The ΔS_{surr} is defined as the quantity of heat exchanged per kelvin (K) of temperature in the course of a specific chemical or physical change. The ΔS_{surr} can therefore be defined as

$$\Delta S_{surr} = -\Delta H/T$$

By substitution

$$\Delta S_{univ} = -\Delta H/T + \Delta S_{sys}$$

Multiply both sides by $-T$:

$$-T\Delta S_{univ} = \Delta H - T\Delta S_{sys}$$

Josiah Gibbs defined the state function $-T\Delta S_{univ}$, now known as *Gibbs free energy change* or ΔG:

$$\Delta G = \Delta H - T\Delta S_{sys}$$

At constant temperature and pressure, the change in free energy is negative when ΔS_{univ} is positive, which reflects a spontaneous reaction said to be **exergonic** (**Figure 4.5**). If the ΔG is positive, the process is characterized as **endergonic** (nonspontaneous). When the ΔG is zero, the process is at equilibrium. As with other thermodynamic functions, ΔG provides no information about reaction rates. Reaction rates depend on the precise mechanism by which a process occurs and are dealt with under the study of kinetics (Chapter 6).

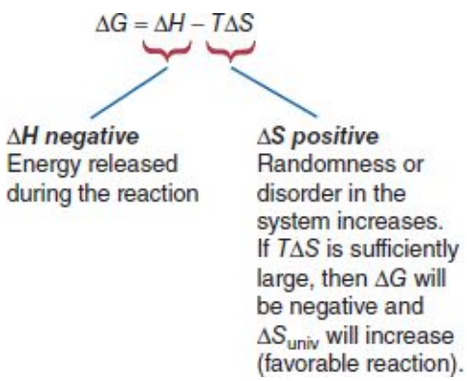


FIGURE 4.5

The Gibbs Free Energy Equation

At constant pressure, enthalpy (H) is essentially equal to the total energy content of the system. A process is spontaneous if it decreases free energy. At constant temperature and pressure, free energy changes (ΔG) are negative if enthalpy decreases or if the entropy term $T\Delta S$ is sufficiently large.

Standard Free Energy Changes

The *standard state* provides a uniform basis for free energy calculations. The standard free energy, $\Delta G°$, is defined for reactions at 25°C (298 K) and 1.0 atm pressure with all solutes at a concentration of 1.0 M.

The standard free energy change is related to the reaction's equilibrium constant, K_{eq}. This is the value of the reaction quotient at equilibrium when the forward and reverse reaction rates are equal. For a reaction

$$aA + bB \rightleftharpoons cC + dD$$

the equilibrium constant is related to the reaction's equilibrium concentrations:

$$K_{eq} = \frac{[C]^c\,[D]^d}{[A]^a\,[B]^b}$$

Based on the observations that the free energy of an ideal gas depends on its pressure (concentration) and that the state function G can be manipulated in the same way as the state function H, the following equation was derived:

$$\Delta G = \Delta G° + RT\,\ln\frac{[C]^c\,[D]^d}{[A]^a\,[B]^b}$$

If the reaction is allowed to go to equilibrium, the ΔG is 0 and the expression reduces simply to

$$\Delta G° = -RT\,\ln K_{eq}$$

This equation allows the calculation of $\Delta G°$ if the K_{eq} is known. Because most biochemical reactions take place at or near pH 7 ($[H+] = 1.0 \times 10^{-7}$ M), this exception is made in the 1.0 M solute rule in bioenergetics and the free energy change is expressed as $\Delta G°'$.

WORKED PROBLEM 4.3

For the reaction $HC_2H_3O_2 \rightleftharpoons C_2H_3O_2^- + H+$ (the ionization of acetic acid), calculate $\Delta G°$ and $\Delta G°'$. Assume that T = 25°C. The ionization constant for acetic acid is 1.8×10^{-5}. Is this reaction spontaneous? The gas constant R is 8.314 J/mol·K. Recall that

$$K_{eq} = \frac{[C_2H_3O_2^-][H^+]}{[HC_2H_3O_2]}$$

SOLUTION

1. Calculate $\Delta G°$.

$$\Delta G^\circ = -RT \ln K_{eq}$$
$$= -(8.314 \text{ J/mol·K})(298 \text{ K}) \ln(1.8 \times 10^{-5})$$
$$= 27{,}067 = 27.1 \text{ kJ/mol}$$

The ΔG° indicates that under these conditions the reaction is not spontaneous.

2. Calculate $\Delta G^{\circ\prime}$. Use the relation between free energy change and standard free energy change. For this example, the expression becomes

$$\Delta G^{\circ\prime} = \Delta G^\circ + RT \ln[H+]$$

Substituting values, we have

$$\Delta G^{\circ\prime} = 27{,}067 \text{ J/mol} + (8.314 \text{ J/mol.K})(298 \text{ K})(\ln 10^{-7})$$
$$= 27{,}071 - 39{,}934$$
$$= -12867.54 = -12.9 \text{ kJ/mol}$$

Under the conditions specified for ΔG° (i.e., 1 M concentrations for all reactants including $H+$) the ionization of acetic acid is not spontaneous, as indicated by the positive ΔG°. When the pH value is 7, however, the reaction becomes spontaneous. A low $[H+]$ makes the ionization of a weak acid such as acetic acid a more likely process, as indicated by the negative $\Delta G^{\circ\prime}$. ■

Coupled Reactions

Many chemical reactions within living organisms have positive $\Delta G^{\circ\prime}$ values. Fortunately, free energy values are additive in any reaction sequence.

$A + B \rightleftharpoons C + D$	$\Delta G^{\circ\prime}_{\text{reaction 1}}$	(1)
$C + E \rightleftharpoons F + G$	$\Delta G^{\circ\prime}_{\text{reaction 2}}$	(2)
$A + B + E \rightleftharpoons D + F + G$	$\Delta G^{\circ\prime}_{\text{overall}} = \Delta G^{\circ\prime}_{\text{reaction 1}} + \Delta G^{\circ\prime}_{\text{reaction 2}}$	(3)

Note that reactions (1) and (2) are coupled (i.e., they have a common intermediate, C). If the net $\Delta G^{\circ\prime}$ value ($\Delta G^{\circ\prime}_{\text{overall}}$) is sufficiently negative, forming the products F and G is an exergonic process.

The conversion of glucose-6-phosphate to fructose-1,6-bisphosphate illustrates the principle of coupled reactions (**Figure 4.6**). The common intermediate in this reaction sequence is fructose-6-phosphate. Because the formation of fructose-6-phosphate from glucose-6-phosphate is endergonic ($\Delta G^{\circ\prime}$ is +1.7 kJ/mol), the reaction is not expected to proceed as written (at least under standard conditions). The conversion of fructose-6-phosphate to fructose-1,6-bisphosphate is strongly exergonic because it is coupled to the cleavage of the phosphoanhydride bond of ATP. (The cleavage of ATP's phosphoanhydride bond to form ADP yields approximately −30.5 kJ/mol. ATP in living organisms is discussed in Section 4.3.) Because $\Delta G^{\circ\prime}_{\text{overall}}$ for the coupled reactions is negative, the reactions do proceed in the direction written.

KEY CONCEPT

Free energy is a thermodynamic function that can be used to predict the spontaneity of a process.

Spontaneous reactions are exergonic ($-\Delta G$). Nonspontaneous reactions are endergonic ($+\Delta G$).

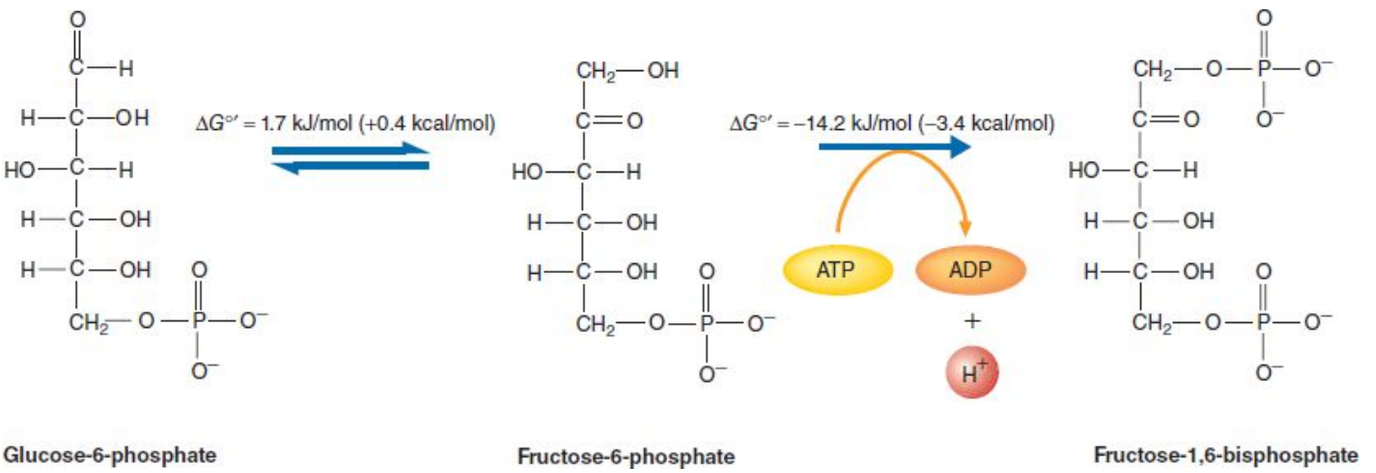


FIGURE 4.6

A Coupled Reaction

The net $\Delta G°'$ value for the two reactions is −12.5 kJ/mol (−3.0 kcal/mol).

3D animation of Fructose-1,6-bisphosphate

WORKED PROBLEM 4.4

Glycogen is synthesized from glucose-1-phosphate. To be incorporated into glycogen, glucose-1-phosphate is first converted to a derivative of the nucleotide uridine diphosphate (UDP). The UDP serves as an excellent leaving group in the condensation reaction to form the glycogen polymer. The reaction is

$$\text{Glucose-1-phosphate} + \text{UTP} + \text{H}_2\text{O} \rightarrow \text{UDP-glucose} + \text{PP}_\text{i}$$

where PP_i is the inorganic compound pyrophosphate.

If the $\Delta G°'$ value for this reaction is approximately zero, is this reaction favorable? If PP_i is hydrolyzed, then

$$\text{PP}_\text{i} + \text{H}_2\text{O} \rightarrow 2\text{P}_\text{i}$$

where P_i is the inorganic compound orthophosphate.

The loss in free energy ($\Delta G°'$) is −33.5 kJ. What is the overall reaction? Determine the $\Delta G°'_{\text{overall}}$ value. How does this second reaction affect the first one?

SOLUTION

The overall reaction is

Glucose-1-phosphate + UTP → UDP-glucose + $2P_i$

$$\Delta G^{\circ\prime}_{\text{overall}} = \Delta G^{\circ\prime}_{\text{reaction 1}} + \Delta G^{\circ\prime}_{\text{reaction 2}}$$
$$= 0 + (-33.5 \text{ kJ})$$
$$= -33.5 \text{ kJ}$$

The hydrolysis of PP_i drives the formation of UDP-glucose from glucose-1-phosphate and UTP. ■

WORKED PROBLEM 4.5

Consider the following reaction in which the sugar fructose reacts with ATP to form fructose-6-phosphate:

ATP + Fructose → ADP + Fructose-6-phosphate

Calculate the equilibrium constant for the reaction given the following free energy values for the two half-reactions:

	$\Delta G^{\circ\prime}$(kJ/mol)
ATP → ADP + P_i	−30.5
Fructose + P_i → Fructose-6-phosphate	+15.9

SOLUTION

1. Sum the free energy values for the two reactions

ATP + Fructose → Fructose-6-phosphate + ADP
−30.5 kJ/mol + 15.9 kJ/mol = −14.6 kJ/mol

2. Determine the value of K_{eq} using the equation

$$\log K_{eq} = -14{,}600/-5706.5$$
$$\log K_{eq} = 2.6$$
$$K_{eq} = 398.1$$

■

QUESTION 4.1

In living cells, the concentrations of ATP and the products of its hydrolysis (ADP and P_i) are significantly lower than the standard 1 M concentrations. Therefore, the actual free energy of hydrolysis of ATP ($\Delta G'$) differs from the standard free energy ($\Delta G^{\circ\prime}$). Unfortunately, it is difficult to obtain an accurate measure of the concentrations of cellular components. For this reason, only estimates can be made. The following equation includes a correction for nonstandard concentrations:

$$\Delta G' = \Delta G^{\circ\prime} + RT \ln \frac{[\text{ADP}][P_i]}{[\text{ATP}]}$$

The temperature is 37°C. Assume that the pH is 7. In a liver cell, the approximate concentrations (mM) are as follows:

$$ATP = 4.0, \quad ADP = 1.35, \quad P_i = 4.65$$
$$\Delta G^{\circ\prime} = -30.5 \text{ kJ/mol}$$

What is the actual $\Delta G'$ for the hydrolysis of ATP under these conditions?

The Hydrophobic Effect Revisited

Understanding the spontaneous aggregation of nonpolar substances in water is enhanced by consideration of thermodynamic principles. When nonpolar molecules are mixed with water, they disrupt water's energetically favorable hydrogen-bonded interactions. The hydrogen bonds that stabilize the highly ordered cage-like structures around clusters of nonpolar molecules restrict the motion of the water molecules, thus resulting in a decrease in entropy. Consequently, the free energy of dissolving nonpolar molecules is unfavorable (i.e., ΔG is positive because ΔH is positive and $-T\Delta S$ is strongly positive). The decrease in entropy, however, is proportional to the surface area of contact between nonpolar molecules and water. The aggregation of nonpolar molecules significantly decreases the surface area of their contact with water, and thus the water becomes less ordered (i.e., the entropy change, ΔS, is now positive). Because $-T\Delta S$ becomes negative, the free energy of the process is negative, and therefore, it proceeds spontaneously. The spontaneous exclusion of nonpolar groups or molecules by the interaction of water with other water molecules or nearby polar groups is a major factor in biological processes such as protein folding and the assembly of supramolecular structures such as membranes.

Biochemistry IN PERSPECTIVE

Nonequilibrium Thermodynamics and the Evolution of Life

How does thermodynamic theory relate to energy flow in living organisms? The thermodynamic concepts described in this chapter, referred to as classical thermodynamics, were discovered during investigations of internal combustion engines in the nineteenth century. Classical thermodynamics explains energy flow in ideal systems in or near equilibrium. Living organisms, however, are open systems that are never at equilibrium until they die. In contrast to stable systems that are in thermodynamic equilibrium, systems that are far from equilibrium are inherently unstable. Thus, a critical question arises: How can an organized living system (a living organism) not in equilibrium remain structurally stable for an extended period of time?

The properties of a Benard cell provide a clue to the phenomenon of order within a universe that favors disorder. A Benard cell is a fluid-filled insulated container that is fitted with a cold reservoir on top and a heat source at the bottom. The liquid placed in the container at the beginning of the experiment has a uniform temperature. As the temperature of the liquid at the bottom of the container gradually increases, a temperature gradient is created. Warm (less dense) liquid begins to rise and cooler (more dense) liquid moves downward. As a specific temperature threshold is reached,

convection currents that are organized, rotating, and dynamically stable form spontaneously. The term *dissipative* is used to describe the capacity of far-from-equilibrium systems such as the Benard cell to form ordered structures under the influence of an energy gradient.

Living organisms are **dissipative systems** that facilitate the reduction in the enormous energy gradient between the sun and the Earth. Energy dissipation begins when photosynthetic organisms capture a fraction of the total solar radiation (about 10^{18} kJ/day). Phototrophs dissipate a portion of the captured energy by producing their own ordered structure. The dissipation process continues as animals and other heterotrophs consume phototrophs. Eventually, all of the energy captured from the sun is released as heat (disorganized energy). Living organisms can be compared to the phenomenon of the Benard cell only in the limited sense that the energy-driven creation of ordered structures is reminiscent of the organized convection currents in the Benard cell.

The critical property of energy flow in living organisms is that equilibrium is never reached because the key feature of the living state is the ability to dissipate energy. The ordered structure of living systems, made possible by the uninterrupted flow of high-quality energy, is maintained by the capacity to release heat and more disordered waste products into the surroundings. The maintenance of dissipative systems requires that continuous work be done on the system because otherwise all natural processes will proceed toward equilibrium. In living organisms, this far-from-equilibrium state is maintained by transport, chemical, and mechanical work.

The evolution of living systems is driven by the size of the energy gradient to be dissipated and the dictates of the second law of thermodynamics, namely, the required increase in entropy in the universe. But although the second law determines the direction of living processes, it is insufficient to explain the precise molecular mechanisms that sustain life. The mechanisms by which energy flow is coupled to the performance of work to build and maintain living organisms evolved over several billion years, and the precise details of the mechanism of energy dissipation by living organisms have yet to be resolved. Through trial and error, living organisms, taking advantage of the physical and chemical properties of elements such as carbon, nitrogen, and oxygen, have developed gigantic and complex energy-dissipating biochemical and information-processing networks. The vast species diversification observed on Earth can, therefore, be viewed as a means to provide the greatest number of pathways for energy dissipation. The region of the planet that exhibits the greatest number of species of living organisms is at the equator, where the energy gradient between the sun and Earth is at its largest. Unsurprisingly, nonequilibrium thermodynamics is an active area of research.

SUMMARY Living organisms are far-from-equilibrium dissipative structures. They create and maintain internal organization via a continuous flow of energy.

4.3 ENERGY CURRENCY MOLECULES

Two types of reactions together provide the energy that drives metabolic processes: phosphorylation and redox reactions. In phosphorylation reactions, an inorganic phosphate group (P_i) bonds with another molecule. ATP, the most important source of immediate energy in living cells, is the product of the reaction of adenosine diphosphate with P_i. In redox reactions (p. 18) electrons, along with the energy they contain, are transferred from reduced molecules to oxidized molecules. The principal electron carriers in living cells are NADH, NADPH, and $FADH_2$.

Adenosine triphosphate is a nucleotide that plays an extraordinarily important role in living cells. ATP hydrolysis (**Figure 4.7**) immediately and directly provides the free energy to drive an immense variety of endergonic biochemical reactions. Produced from ADP and P_i with energy released by the

breakdown of food molecules and the light reactions of photosynthesis, ATP drives processes of several types (**Figure 4.8**). These include (1) biosynthesis of biomolecules, (2) active transport of substances across cell membranes, and (3) mechanical work such as muscle contraction.

ATP is ideally suited to its role as universal energy currency because of its structure (**Figure 4.9**). ATP is a nucleotide composed of adenine, ribose, and a triphosphate unit. Its two terminal phosphoryl groups ($—PO_3^{2-}$) are linked by phosphoanhydride bonds. Although anhydrides are easily hydrolyzed, the phosphoanhydride bonds of ATP are sufficiently stable under mild intracellular conditions. Specific enzymes facilitate ATP hydrolysis.

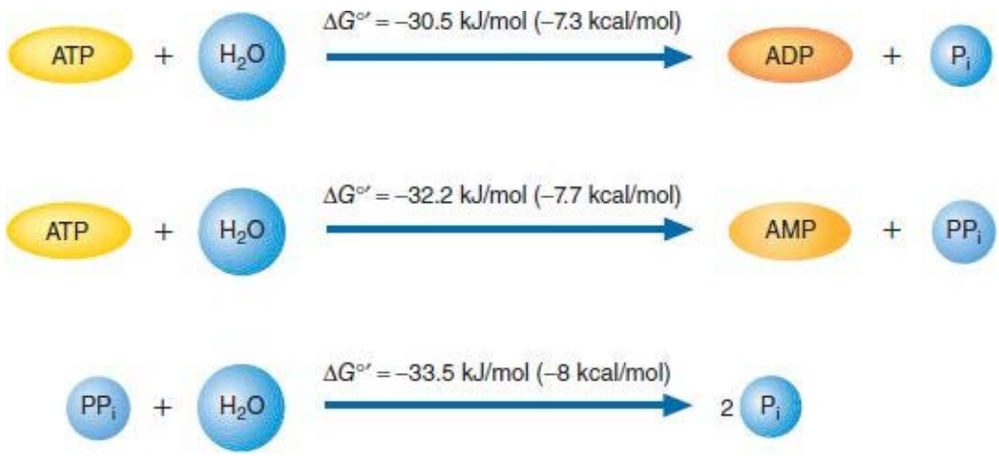


FIGURE 4.7

Hydrolysis of ATP

ATP may be hydrolyzed to form ADP and P_i (orthophosphate) or AMP (adenosine monophosphate) and PP_i (pyrophosphate). Pyrophosphate may be subsequently hydrolyzed to orthophosphate, releasing additional free energy. The hydrolysis of ATP to form AMP and pyrophosphate is often used to drive reactions with high positive $\Delta G^{\circ\prime}$ values or to ensure that a reaction goes to completion.

3D animation of ATP

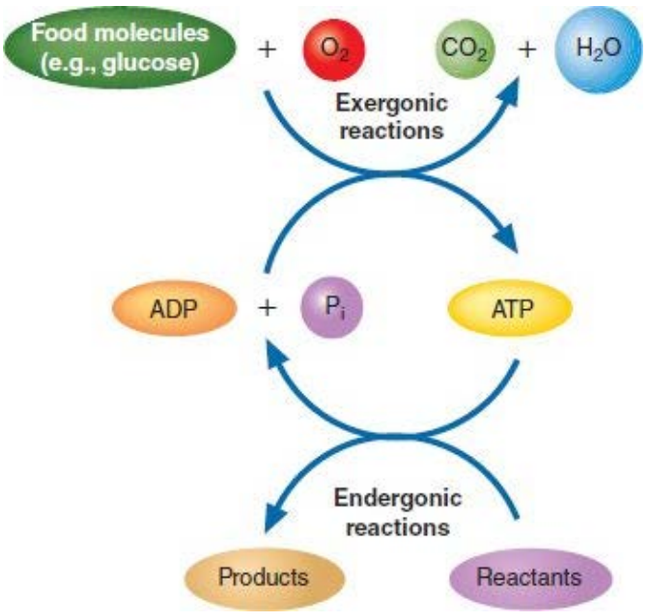


FIGURE 4.8

The Role of ATP

ATP is an intermediate in the flow of energy from food molecules to the biosynthetic reactions of metabolism.

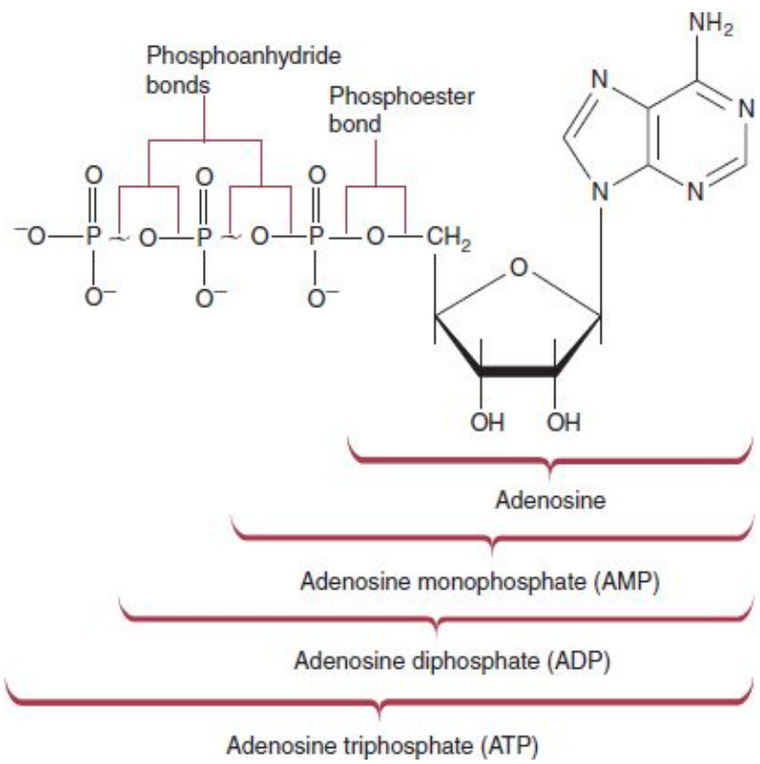


FIGURE 4.9

The Structure of ATP

The squiggles (~) in ATP indicate that the bonds so connected are easily hydrolyzed.

The tendency of ATP to undergo hydrolysis, also referred to as its **phosphoryl group transfer potential**, is not unique. A variety of biomolecules can transfer phosphate groups to other compounds. **Table 4.1** lists several important examples.

Phosphorylated compounds with high negative $\Delta G°'$ values of hydrolysis have larger phosphoryl group transfer potentials than those compounds with smaller negative values. Because ATP has an

intermediate phosphoryl group transfer potential, it can be an intermediate carrier of phosphoryl groups from higher-energy molecules such as phosphoenolpyruvate to lower-energy molecules such as in the synthesis of glucose-6-phosphate (**Figure 4.10**). ATP is therefore the "energy currency" for living systems because cells usually transfer phosphate by coupling reactions to ATP hydrolysis. The two phosphoanhydride bonds of ATP are often referred to as "high energy." The term *high-energy bond* is now considered inappropriate, however, because it denotes instability of the bond and, therefore, its ability to participate in reactions rather than the quantitative value of the bond energy. To understand why ATP hydrolysis is so exergonic, several factors must be considered.

TABLE 4.1 Standard Free Energy of Hydrolysis of Selected Phosphorylated Biomolecules

	$\Delta G°'$	
Molecule	**kcal/mol**	**kJ/mol**
Glucose-6-phosphate	−3.3	−13.8
Fructose-6-phosphate	−3.8	−15.9
Glucose-1-phosphate	−5	−20.9
ATP → ADP + P_i	−7.3	−30.5
ATP → AMP + PP_i	−7.7	−32.2
PP_i → $2P_i$	−8.0	−33.5
Phosphocreatine	−10.3	−43.1
Glycerate-1,3-bisphosphate	−11.8	−49.4
Carbamoyl phosphate	−12.3	−51.5
Phosphoenolpyruvate	−14.8	−61.9

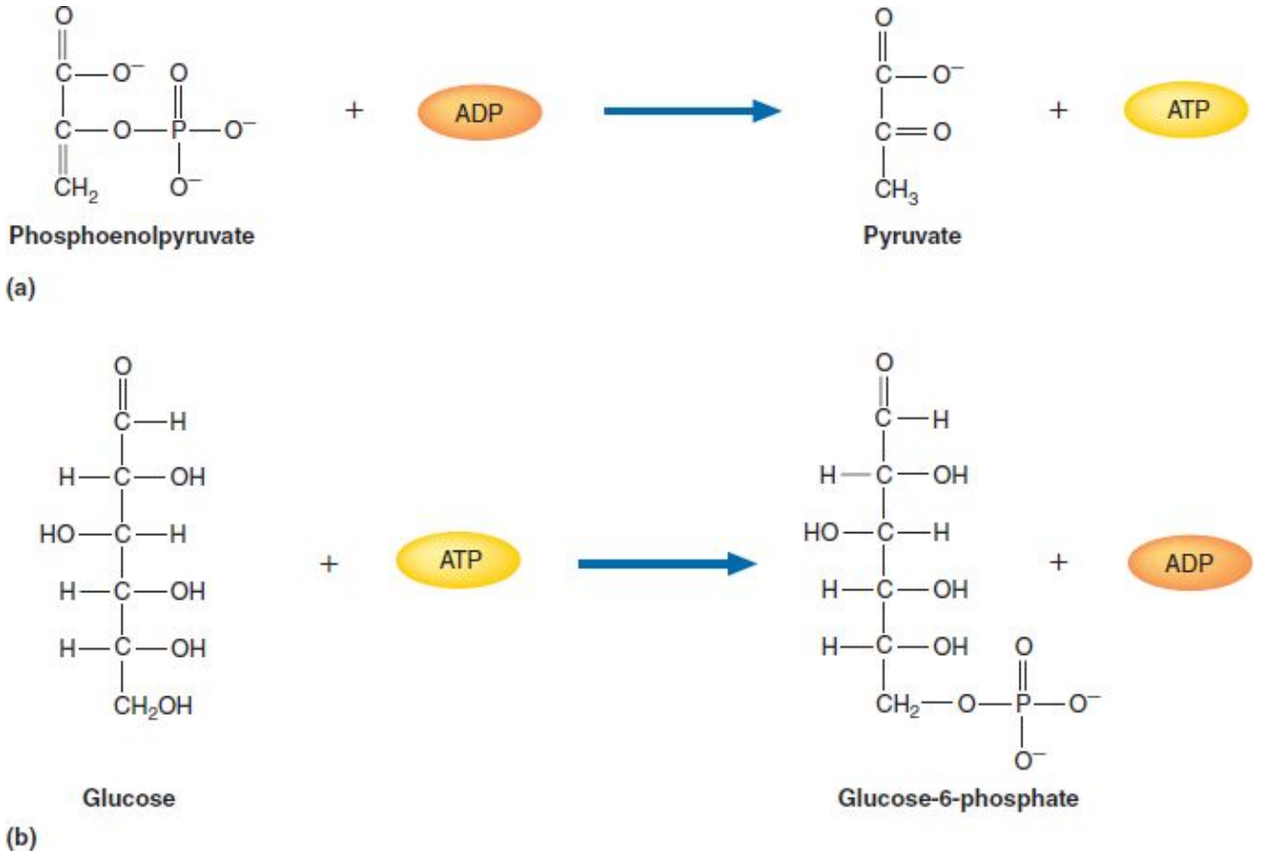


FIGURE 4.10

Transfer of Phosphoryl Groups

(a) Transfer of a phosphoryl group from phosphoenolpyruvate to ADP. As discussed in Chapter 8, this reaction is one of two steps that form ATP during glycolysis, the reaction pathway that breaks down glucose. (b) Transfer of a phosphoryl group from ATP to glucose. The product of this reaction, glucose-6-phosphate, is the first intermediate formed during glycolysis.

3D animation of glucose-6-phosphate

FIGURE 4.11

Contributing Structure of the Resonance Hybrid of Phosphate

At physiological pH, orthophosphate is HPO_4^{2-}. In this illustration, H+ is not assigned permanently to any of the four oxygen atoms.

1. At typical intracellular pH values, the triphosphate unit of ATP carries three or four negative charges that repel each other. Hydrolysis of ATP reduces electrostatic repulsion.
2. Because of *resonance stabilization*, the products of ATP hydrolysis are more stable than a resonance-restricted ATP. When a molecule has two or more alternative structures that differ

only in the position of electrons, the result is called a **resonance hybrid**. The electrons in a resonance hybrid with several contributing structures possess much less energy than those with fewer contributing structures. The contributing structures of the phosphate resonance hybrid are illustrated in **Figure 4.11**.

3. The hydrolyzed products of ATP, either ADP and P_i or AMP and PP_i, are more easily solvated than ATP. Recall that the water molecules that form the solvation spheres around ions shield them from one another. The resulting decrease in the repulsive force between phosphoryl groups drives the hydrolytic reaction.
4. There is an increase in disorder caused by an increase in the number of molecules. ATP is converted into two molecules (ADP and P_i), both of which now move randomly.

KEY CONCEPTS

- The hydrolysis of ATP immediately and directly provides the free energy to drive an immense variety of endergonic biochemical reactions.
- Because ATP has an intermediate phosphoryl group transfer potential, it can carry phosphoryl groups from high-energy molecules to low-energy molecules.
- ATP is the energy currency for living systems.

QUESTION 4.2

Walking consumes approximately 100 kcal/mi. In the hydrolysis of ATP (ATP → ADP + P_i), the reaction that drives muscle contraction, $\Delta G°'$ is −7.3 kcal/mol (−30.5 kJ/mol). Calculate how many grams of ATP must be produced to walk a mile. ATP synthesis is coupled to the oxidation of glucose ($\Delta G°'$ = −686 kcal/mol). How many grams of glucose are actually metabolized to produce this amount of ATP? (Assume that only glucose oxidation is used to generate ATP and that 40% of the energy generated from this process is used to phosphorylate ADP. The gram molecular weight of glucose is 180 g and that of ATP is 507 g.)

Energy and Reducing Power

In addition to ATP, there are three other energy currency molecules in cells: NADH, NADPH and $FADH_2$. All three molecules are electron carriers with considerable reducing power; that is, they donate high-energy electrons in cellular redox reactions. NADH (whose structure and properties are described on pp. 335–36) has a major role in the energy captured in the complete oxidation of glucose (pp. 380–83) and fatty acids (pp. 455–56), processes that yield large amounts of ATP. NADPH (Figure 9.5 on p. 336) donates high-energy electrons in the biosynthesis of molecules such as fatty acids and cholesterol. In photosynthesis (pp. 507–09), the reducing power of NADPH is used to convert CO_2 to sugar molecules. NADPH is also an important antioxidant, a molecule that protects cells from oxygen radicals (pp. 385–88). $FADH_2$, the reduced form of FAD (flavin adenine dinucleotide; Figure 9.6 on p. 337) has functions in a large variety of redox reactions. $FADH_2$'s most prominent roles are in the citric acid cycle (a biochemical pathway that oxidizes acetyl groups to yield, in addition to $FADH_2$, CO_2 and NADH) and as a product of fatty acid degradation.

Chapter Summary

1. All living organisms unrelentingly require energy. Bioenergetics, the study of energy transformations, can be used to determine the direction and extent to which biochemical reactions proceed. Enthalpy (a measure of heat content) and entropy (a measure of disorder) are related to the first and second laws of thermodynamics, respectively. Free energy (the portion of total energy that is available to do work) is related to a mathematical relationship between enthalpy and entropy.
2. Energy and heat transformations take place in a "universe" composed of a system and its surroundings. In an open system, matter and energy are exchanged between the system and its surroundings. If energy, but not matter, can be exchanged with the surroundings, then the system is said to be closed. Living organisms are open systems.
3. Several thermodynamic quantities are state functions; that is, their value does not depend on the pathway used to make or degrade a specific substance. Examples of state functions are total energy, free energy, enthalpy, and entropy. Quantities such as work and heat depend on the pathway and thus are not state functions.
4. Free energy, a state function that relates the first and second laws of thermodynamics, represents the maximum useful work obtainable from a process. Exergonic processes, that is, processes in which free energy decreases ($\Delta G < 0$), are spontaneous. If the free energy change is positive ($\Delta G > 0$), the process is called endergonic. A system is at equilibrium when the free energy change is zero. The standard free energy ($\Delta G°$) is defined for reactions at 25°C, 1 atm pressure, and 1 M solute concentrations. The standard pH in bioenergetics is 7. The standard free energy change $\Delta G°'$ at pH 7 is used in this textbook.
5. ATP hydrolysis provides most of the free energy required for living processes. ATP is ideally suited to its role as universal energy currency because it is a sufficiently stable phosphoanhydride and has an intermediate phosphoryl group transfer potential that allows it to carry phosphoryl groups from high-energy biomolecules to lower-energy biomolecules.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on energy to help you prepare for exams.

Chapter 4 Review Quiz

Suggested Readings

Falkowski PG, Fenchel T, Delong EF. 2008. The microbial engines that drive Earth's biogeochemical cycles. Science 320:1034–9.

Hanson RW. 1989. The role of ATP in metabolism. Biochem Educ 17:86–92.

Harold FM. 2005. Molecules into cells: specifying spatial architecture. Microbiol Mol Biol Rev 69(4):544–64.

Kleidon A. 2010. Non-equilibrium thermodynamics, maximum entropy production and Earth-system evolution. Phil Trans R Soc 368(1910):181–96.

Mascarelli AL. 2009. Low life. Nature 459:770–3.

Rubi JM. 2008. The long arm of the second law. Sci Am 299(5):62–7.

Schneider ED, Sagan D. 2005. Into the cool: energy flow, thermodynamics and life. Chicago (IL): University of Chicago.

Schrödinger E. 1944. What is life? Cambridge (UK): Cambridge University.

Wallace DC. 2010. Bioenergetics: the origins of complexity and the ascent of man. Proc Nat Acad Sci 107:8947–53.

Key Words

bioenergetics, *113*
biogeochemical cycle, *111*
dissipative system, *123*
electron acceptor, *113*
electron donor, *113*
endergonic process, *119*
endothermic reaction, *115*
enthalpy, *113*
entropy, *113*
exergonic process, *119*
exothermic reaction, *115*
free energy, *113*
isothermic reaction, *115*
phosphoryl group transfer potential, *125*
resonance hybrid, *127*
spontaneous chemical changes, *117*
thermodynamics, *113*
work, *112*

Review Questions

SECTION 4.1

Comprehension Questions

1. Define the following terms:
    a. thermodynamics
    b. bioenergetics
    c. entropy
    d. enthalpy
    e. free energy
2. Define the following terms:
    a. work
    b. kinetically favorable
    c. thermodynamically favorable
    d. closed system
    e. open system
3. Define the following terms:
    a. endothermic reaction
    b. exothermic reaction
    c. isothermic reaction
    d. spontaneous process
    e. nonspontaneous process

Fill in the Blanks

4. An _____ reaction liberates heat.
5. Energy is defined as the capacity to do _____.
6. Entropy is a measure of the _____ of a system.
7. According to the _____ law of thermodynamics, the disorder of the universe always increases.
8. The exchange of energy between a system and its surroundings can occur in two ways: _____ and _____.
9. Thermodynamic state functions include enthalpy, entropy, and _____.
10. $\Delta H = 0$ is a(n) _____ process.
11. Spontaneous processes occur with the release of _____.

Short-Answer Questions

12. It is thermodynamically favorable for methane to autoignite and burn at room temperature, but this does not occur. Explain.
13. What is the difference between exothermic and exergonic reactions?
14. Living organisms within Earth's crust must be tolerant of high temperatures. What is the major source of the heat?
15. What are the first and second laws of thermodynamics, and what are their defining equations?
16. Which of the following thermodynamic quantities is not a state function? Explain.

    a. work
    b. entropy
    c. enthalpy
    d. free energy

17. The first and second laws of thermodynamics are useful for biochemists who investigate chemical reactions in living organisms. Explain why the third law is not useful.
18. Define the thermodynamic term "work." Provide two physiological examples of work.

Critical-Thinking Questions

19. Many salts produce heat when dissolved in water and the solution becomes warm, reflecting the fact that the ΔH term is positive. In other cases, the solution becomes quite cold, and the enthalpy term is negative. What must be true about the Gibbs equation for these reactions to be spontaneous?
20. Thermodynamics is based on the behavior of large numbers of molecules. Yet within a cell there may only be a few molecules of a particular type at a time. Do the laws of thermodynamics apply under these circumstances?
21. When ammonium chloride dissolves in water, the solution becomes cold. Endothermic processes are usually not spontaneous, so why does this salt dissolve in water?
22. Matter has been described as condensed energy. Using Einstein's equation, calculate the amount of energy in 1 mg of dust. How much coal would have to be burned (yielding 393.3 kJ/mol) to produce an equivalent amount of energy? Assume that coal is pure carbon.
23. After reviewing the discussion on pp. 111–12, outline the phases of the carbon cycle as described.
24. Balance the following reaction and calculate its ΔH value:

    $$C_{17}H_{35}COOH + O_2 \rightarrow CO_2 + H_2O$$

    where the ΔH values (kcal/mol) are as follows:

    $C_{17}H_{35}COOH$ (−211.4)

O_2 (0)

CO_2 (−94)

H_2O (−68.4)

SECTION 4.2

Comprehension Questions

25. Define the following terms:

    a. free energy
    b. exergonic reaction
    c. endergonic reaction
    d. standard free energy charge
    e. $\Delta G^{o\prime}$

26. Define the following terms:

    a. Gibbs free energy equation
    b. K_{eq}
    c. ΔS_{univ}
    d. coupled reactions
    e. dissipative system

Fill in the Blanks

27. The _____ free energy is defined for reactions at 298 K and 1.0 atm pressure with all solutes at a concentration of 1.0 M.
28. The cleavage of ATP's phosphoanhydride bond to form ADP yields approximately _____ kJ/mol.
29. When the value of $\Delta G°'_{overall}$ for coupled reactions is_____, the reactions proceed in the direction as written at standard conditions.

Short-Answer Questions

30. The equilibrium constant for the dissociation of acetic acid is 1.8×10^{-5}. What is the free energy change for the reaction?
31. Under standard conditions, which statements are true?

    a. $\Delta G = \Delta G°$
    b. $\Delta H = \Delta G$
    c. $\Delta G = \Delta G^o + RT \ln K_{eq}$
    d. $\Delta G° = \Delta H - T\Delta S$
    e. $P = 1$ atm
    f. $T = 273\ K$

32. Which statements are true and which are false? Modify each false statement so that it reads correctly.

    a. In a closed system, neither energy nor matter is exchanged with the surroundings.
    b. State functions are independent of the pathway.

c. A process is isothermic if $\Delta H = 0$.
d. The sign and magnitude of ΔG give important information about the direction and rate of a reaction.
e. At equilibrium, $\Delta G = \Delta G°$.
f. For two reactions to be coupled, they must have a common intermediate.

33. What statements concerning free energy change are true or false?

a. Free energy change is a measure of the rate of a reaction.
b. Free energy change is a measure of the maximum amount of work available from a reaction.
c. Free energy change is a constant for a reaction under any conditions.
d. Free energy is related to the equilibrium constant for a specific reaction.
e. Free energy change is equal to zero at equilibrium.

34. Consider the following reaction:

Glucose-1-phosphate → Glucose-6-phosphate

$\Delta G° = -7.1$ kJ/mol

What is the equilibrium constant for this reaction at 25°C?

35. Given the following equation:

Glycerol-3-phosphate → Glycerol + P_i

$\Delta G°' = 9.7$ kJ/mol

At equilibrium, the concentration of both glycerol and orthophosphate are 1 mM. Under these conditions, calculate the final concentration of glycerol-3-phosphate.

36. The $\Delta G°'$ value for the hydrolysis of glucose-6-phosphate is −13.8 kJ/mol. Assuming a concentration of 4 mM for this molecule, what would the phosphate concentration be at equilibrium?
37. The $\Delta G°'$ value for glucose-1-phosphate is -20.9 kJ/mol. If glucose and phosphate are both at 4.8 mM, what is the equilibrium concentration of glucose-1-phosphate?
38. If glucose, phosphate, and glucose-6-phosphate are combined in concentrations of 4.8, 4.8, and 0.25 mM, respectively, what is the equilibrium constant for the hydrolysis of glucose-6-phosphate at a temperature of 25°C?
39. Of the three thermodynamic quantities, enthalpy, free energy, and entropy, which provides the most useful indicator of spontaneity in a reaction? Explain.

Critical-Thinking Questions

40. In the reaction ATP + glucose → ADP + glucose-6-phosphate, $\Delta G°$ is -16.7 kJ/mol. Assume that both ATP and ADP have a concentration of 1 M and $T = 25°C$. What ratio of glucose-6-phosphate to glucose would allow the reverse reaction to occur?
41. Given the following data, calculate K_{eq} for the denaturation reaction of the protein β-lactoglobin at 25°C:

$\Delta H° = -88$ kJ/mol

$\Delta S° = 0.3$ kJ/mol.

The free energy of hydrolysis of ATP in systems free of Mg^{2+} is −35.7 kJ/mol. When the concentration of this ion is 5 mM, $\Delta G°_{observed}$ is approximately −31 kJ/mol at pH 7 and 38°C. Suggest a possible reason for this effect.

42. The free energy of hydrolysis for acetic anhydride is −21.8 kJ/mol. The conversion of ATP to ADP also involves the cleavage of an anhydride bond. Its free energy of hydrolysis is −30

kJ/mol. Explain the difference in these values.

SECTION 4.3

Comprehension Questions

44. Define the following terms:

    a. phosphoryl group transfer potential
    b. adenosine
    c. adenosine monophosphate
    d. high-energy bond
    e. resonance hybrid

45. Define the following terms:

    a. resonance stabilization
    b. electrostatic repulsion
    c. phosphoanhydride bond
    d. phosphodiester bond
    e. orthophosphate

Fill in the Blanks

46. The hydrolysis of ATP to yield AMP and PP_i yields_____ kJ/mol.
47. In biochemical processes, the hydrolysis of _____ensures that a reaction goes to completion.
48. In living organisms, _____ is the high-energy molecule that links food molecule degradation reactions with biosynthetic reactions.
49. Phosphorylated molecules with high _____ values of hydrolysis have large phosphoryl group transfer potentials.

Short-Answer Questions

50. Consider the following reaction:

    $ATP \rightarrow AMP + 2\ P_i$

    Calculate the equilibrium constant (K_{eq}) given the following $\Delta G°'$ values:

    $ATP \rightarrow AMP + PP_i$ (−32.2 kJ/mol)

    $PP_i \rightarrow 2P_i$ (−33.5 kJ/mol)

51. Which of the following compounds would you expect to liberate the least free energy when hydrolyzed? Explain.

    a. ATP
    b. ADP
    c. AMP
    d. phosphoenolpyruvate
    e. phosphocreatine

52. The free energy of hydrolysis ($\Delta G°'$) of pyrophosphate (PP_i) at pH 7 is 19.2 kJ/mol. Does the value of $\Delta G°'$ change with pH? If so, why? If not, why not?
53. What factors make ATP suitable as an energy currency for living organisms?

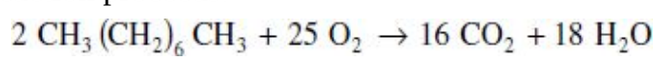
54. Describe why ATP, the molecule that serves as the energy currency in the body, has an intermediate phosphoryl group transfer potential.
55. How does a common intermediate couple two chemical reactions?
56. List four reasons that account for why ATP hydrolysis is sufficiently exergonic to drive biochemical processes.

Critical-Thinking Questions

57. Pyruvate oxidation yields carbon dioxide and water and liberates energy at the rate of 1142.2 kJ/mol. If electron transport also occurs, approximately 12.5 ATP molecules are produced. The free energy of hydrolysis for ATP is −30.5 kJ/mol. What is the apparent efficiency of ATP production?
58. In many ways, arsenate (AsO_4^{3-}) is very similar to phosphate (PO_4^{3-}), yet it does not substitute for phosphate in biomolecules. After reviewing the essential atomic characteristics of the element arsenic, explain this phenomenon.
59. Magnesium ion (Mg^{2+}) forms complexes with the negative charges of the phosphate in ATP. In the absence of Mg^{2+}, would ATP have more, less, or the same stability as when the ion is present?
60. Glucose-1-phosphate has a $\Delta G^{\circ\prime}$ value of −20.9 kJ/mol, whereas that for glucose-6-phosphate is −12.5 kJ/mol. After reviewing the molecular structures of these compounds, explain why there is such a difference in these values.

MCAT Study Questions

61. Consider the following reaction and determine if the entropy is increasing or decreasing as the reaction proceeds.



$$2\ CH_3(CH_2)_6CH_3 + 25\ O_2 \rightarrow 16\ CO_2 + 18\ H_2O$$

a. increasing
b. decreasing
c. decreasing if the temperature is 400°C or above
d. Not enough information is given.

62. What conclusion concerning spontaneity can be determined about a reaction if the value for ΔH is positive and the value of ΔS is negative and the temperature is 30°C?

a. The reaction will be at equilibrium.
b. The reaction will be spontaneous at a higher temperature.
c. The reaction is spontaneous.
d. The reaction will never be spontaneous.

63. Which of the following statements concerning ATP is true?

a. The free energy value for the hydrolysis of ATP is nearly the same for ADP.
b. The free energy value for the hydrolysis of ATP is greater than that for ADP.
c. ATP hydrolysis is more likely at pH 5 than at pH 7.
d. One mole of glycerate-1,3-bisphosphate can phosphorylate one mole of AMP to yield ATP.

64. Which of the following molecules is an energy currency molecule required in fatty acid synthesis?

a. NADH
b. NAD+
c. NADPH

d. NADP+

65. All of the following are state functions except

a. work
b. enthalpy
c. entropy
d. internal energy

CHAPTER 5 Amino Acids, Peptides, and Proteins

A Spider's Orb Web Constructed with Silk Fibers The orb (circular) web is an efficient means of ensnaring insect prey, which become stuck to the sticky spiral, made of capture silk. Many spiders are nocturnal. The next day, if a web is heavily damaged, the spider will eat the silk and then construct a new web. The amino acid sequence of spider silk protein and the spider's silk fiber-spinning process combine to make spider silk one of the strongest materials on Earth.

OUTLINE

SPIDER SILK: A BIOSTEEL PROTEIN

5.1 AMINO ACIDS

Amino Acid Classes

Biologically Active Amino Acids

Modified Amino Acids in Proteins
Amino Acid Stereoisomers
Amino Acid Ionizable Side Chain Groups and pH
Amino Acid Reactions

5.2 PEPTIDES

5.3 PROTEINS
Functions
Shape and Composition
Protein Structure
The Folding Problem
Fibrous Proteins
Globular Proteins

Biochemistry in Perspective

Spider Silk and Biomimetics

Biochemistry in the Lab

Protein Technology

AVAILABLE ONLINE:

Biochemistry in Perspective
Protein Poisons

Biochemistry in Perspective
Protein Folding and Human Disease

Biochemistry in Perspective
Molecular Machines and Motor Proteins

Biochemistry in Perspective
Myosin: A Molecular Machine

Biochemistry in the Lab
Protein Sequence Analysis: The Edman Degradation

Spider Silk: A Biosteel Protein

Spiders have evolved over 400 million years into exceptionally successful predators. These invertebrate animals are a class of arthropods, called the arachnids, which have an exoskeleton, a segmented body, and jointed appendages. Although spiders possess an efficient venom injection system, their most impressive feature is the production of silk, a multiuse protein fiber. Spun through spinnerets at the end of the spider's abdomen, silk is used in locomotion, mating, and offspring protection. The most prominent use of spider silk, however, is for prey capture. The spiral, wheel-shaped orb web, oriented vertically to intercept fast-moving flying prey, is the best-known method. Spider silk's mechanical properties ensure that the web readily absorbs impact energy so that prey is retained until the spider can subdue it.

Humans have also long appreciated spider webs for their physical properties. Examples range from the ancient Greeks, who used spider webs to treat wounds, to the Australian aborigines who used spider silk to make fishing lines. In modern times, spider silk has served as crosshairs in scientific equipment and gun sights. In the past several decades, spider silk and orb webs have attracted the attention of life scientists, bioengineers, and material scientists as they have begun to appreciate the unique mechanical properties of

this remarkable protein.

There are eight different types of spider silk, although no spider makes all of them. Dragline silk, a very strong fiber, is used for frame and radial lines in orb webs and as a safety line (to break a fall or escape other predators). Capture silk, an elastic and sticky fiber, is used in the spiral of webs.

Spider silk is a lightweight fiber with impressive mechanical properties. *Toughness*, a combination of stiffness and strength, is a measure of how much energy is needed to rupture a fiber. Spider silk is about five times as tough as high-grade steel wire of the same weight and about twice as tough as synthetic fibers such as Kevlar (used in body armor). Spider silk's *tensile strength*, the resistance of a material to breaking when stretched, is as great as that of Kevlar and greater than that of high-grade steel wire. *Torsional resistance*, the capacity of a fiber to resist twisting (an absolute requirement for draglines used as safety lines), is higher for spider silk than for all textile fibers, including Kevlar. Spider silk also has superior *elasticity* and *resilience*, the capacity of a material when it is deformed elastically to absorb and then release energy. Scientists estimate that a 2.54 cm (1 in)–thick rope made of spider silk could be substituted for the flexible steel-arresting wires used on aircraft carriers to rapidly stop a jet plane as it lands.

Overview

PROTEINS ARE MOLECULAR TOOLS THAT PERFORM AN ASTONISHING VARIETY OF FUNCTIONS. IN ADDITION TO SERVING AS STRUCTURAL materials in all living organisms (e.g., actin and myosin in animal muscle cells), proteins are involved in such diverse functions as catalysis, metabolic regulation, transport, and defense. Proteins are composed of one or more polypeptides, unbranched polymers of 20 different amino acids. The genomes of most organisms specify the amino acid sequences of thousands or tens of thousands of proteins.

Proteins are a diverse group of macromolecules (**Figure 5.1**). This diversity is directly related to the combinatorial possibilities of the 20 amino acid monomers. Amino acids can be theoretically linked to form protein molecules in any imaginable size or sequence. Consider, for example, a hypothetical protein composed of 100 amino acids. The total possible number of combinations for such a molecule is an astronomical 20^{100}. However, of the trillions of possible protein sequences, only a small fraction (possibly no more than 2 million) is actually produced by all living organisms. An important reason for this remarkable discrepancy is demonstrated by the complex set of structural and functional properties of naturally occurring proteins that have evolved over billions of years in response to selection pressure. Among these properties are (1) structural features that allow newly synthesized proteins to rapidly fold into their biologically functional conformations (shapes), (2) the presence of binding sites that are specific for one or a small group of ligands (molecules that serve a biological purpose when bound to a larger molecule), (3) an appropriate degree of flexibility that is appropriate to their functions, (4) surface structure that is appropriate for a protein's immediate environment (i.e., hydrophobic in membranes and hydrophilic in cytoplasm), and (5) vulnerability of proteins to degradation reactions when they become damaged or no longer useful.

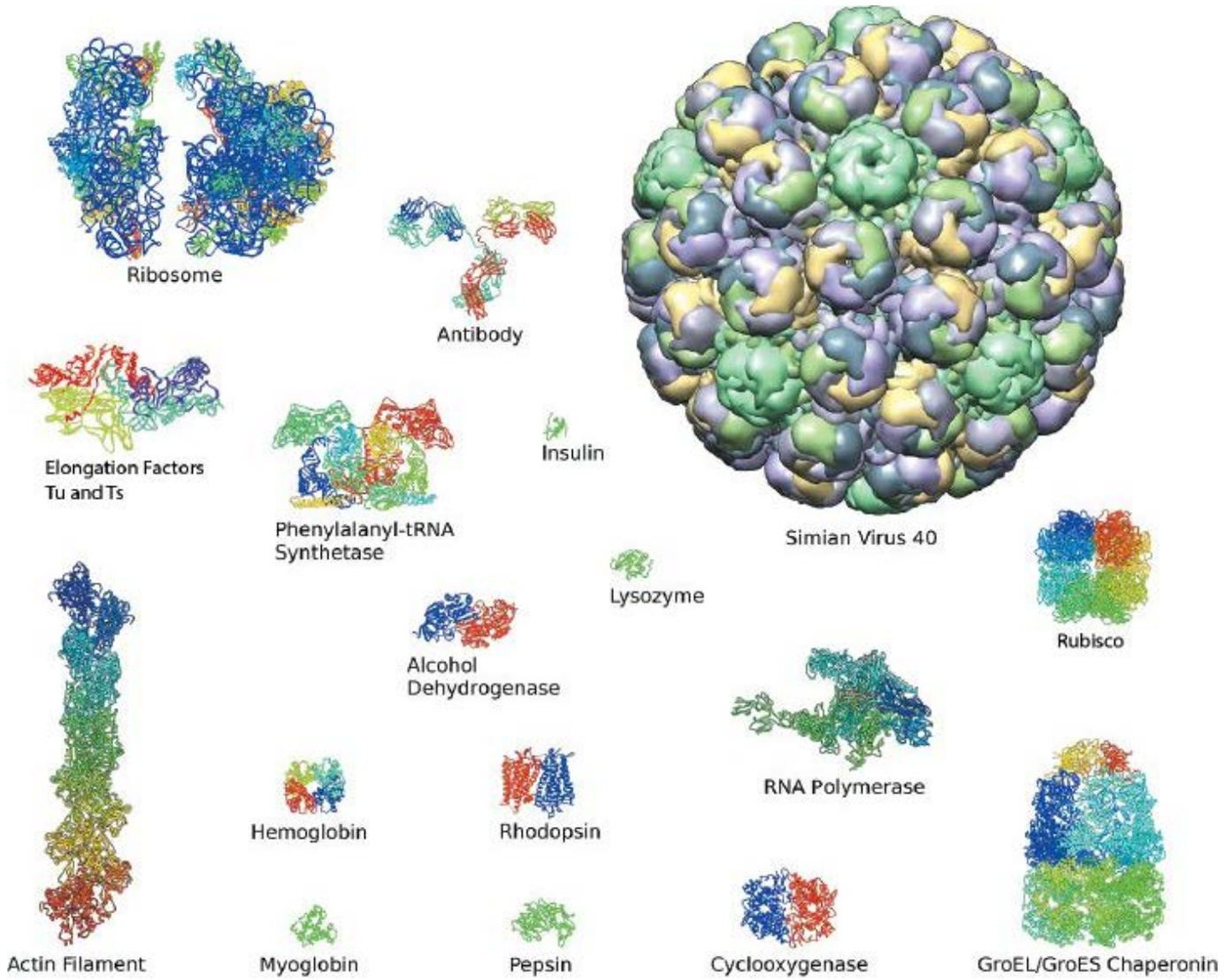


FIGURE 5.1

Protein Diversity

Proteins occur in an enormous diversity of sizes, shapes, and functions. Both the small and large subunits of ribosomes, the molecular machines that synthesize polypeptides, contain large numbers of proteins. Antibodies are proteins produced by the immune system that protect against viruses, bacteria, and other foreign substances. Elongation factors Tu and Ts are prokaryotic proteins that facilitate the binding of aminoacyl-tRNAs to the ribosome. Phenylalanine t-RNA synthetase is an enzyme that links the amino acid phenylalanine to its cognate t-RNA during protein synthesis. Insulin is a protein hormone that regulates several metabolic processes. Simian Virus 40 is a virus in which its DNA genome is enclosed in a protein shell composed of 360 copies of one protein (VP-1). Actin filaments are components of the cytoskeletons of eukaryotes. Alcohol dehydrogenase is an enzyme that interconverts aldehydes and alcohols. Lysozyme is an enzyme that destroys bacterial cell walls, most notably in tears. Rubisco is the enzyme that incorporates CO_2 into organic molecules. Hemoglobin is the protein that carries O_2 from the lungs to the body's tissues. Rhodopsin is a light-sensitive protein found in the rod cells in the retina of the eye. RNA polymerase is the core protein in the protein complex that transcribes DNA to yield RNA. Myoglobin is a protein that stores O_2 in muscle cells. Pepsin is a protein-degrading enzyme produced in stomach cells. Cyclooxygenase is an enzyme that catalyzes the first reaction in a pathway that synthesizes fatty acid derivatives such as the prostaglandins. GroEL/GroES chaperonin is a large protein complex in E. coli that promotes the accurate folding of polypeptides.

Proteins can be distinguished based on their number of amino acids (called **amino acid residues**), their overall amino acid composition, and their amino acid sequence. Molecules with molecular weights ranging from several thousand to several million daltons are called **polypeptides**. Those

with low molecular weights, typically consisting of fewer than 50 amino acids, are called **peptides**. The term **protein** describes molecules that consist of one or more polypeptide chains.

This chapter begins with a review of the structures and chemical properties of the amino acids. This discussion is followed by descriptions of the structural and functional features of peptides and proteins and the protein folding process. The emphasis throughout is on the relationship between the structure and function of polypeptides. In Chapter 6, the functioning of the enzymes, an especially important group of proteins, is discussed. Protein synthesis is covered in Chapter 19.

5.1 AMINO ACIDS

The hydrolysis of each polypeptide yields a set of amino acids, referred to as the molecule's *amino acid composition*. The structures of the 20 amino acids that are commonly found in naturally occurring polypeptides are shown in **Figure 5.2**. These amino acids are referred to as *standard* amino acids. Common abbreviations for the standard amino acids are listed in **Table 5.1**. Note that 19 of the standard amino acids have the same general structure (**Figure 5.3**). These molecules contain a central carbon atom (the α-carbon) to which an amino group, a carboxylate group, a hydrogen atom, and an R (side chain) group are attached. The exception, proline, differs from the other standard amino acids in that its amino group is secondary, formed by ring closure between the R group and the amino nitrogen.

Nonstandard amino acids consist of amino acid residues that have been chemically modified after incorporation into a polypeptide or amino acids that occur in living organisms but are not found in proteins. Nonstandard amino acids found in proteins are usually the result of *posttranslational modifications* (chemical changes that follow protein synthesis). Selenocysteine, an exception to this rule, is discussed in Chapter 19.

At a pH of 7, the carboxyl group of an amino acid is in its conjugate base form (—COO^-), and the amino group is in its conjugate acid form ($-NH_3^+$). Thus, each amino acid can behave as either an acid or a base. The term **amphoteric** is used to describe this property. Molecules that bear both positive and negative charges are called **zwitterions**. The R (side chain) group gives each amino acid its unique properties.

Amino Acid Classes

The sequence of amino acids determines the three-dimensional configuration of each protein. Their structures are therefore examined carefully in the next four subsections. Amino acids are classified according to their capacity to interact with water. Using this criterion, we may distinguish four classes: (1) nonpolar, (2) polar, (3) acidic, and (4) basic.

NONPOLAR AMINO ACIDS The nonpolar amino acids contain mostly hydrocarbon R groups that do not bear positive or negative charges. Nonpolar (i.e., hydrophobic) amino acids play an important role in maintaining the three-dimensional structures of many proteins because they interact poorly with water. Two types of hydrocarbon side chains are found in this group: aliphatic and aromatic. The term **aliphatic** refers to nonaromatic hydrocarbons such as methane and cyclohexane. Glycine, alanine, valine, leucine, isoleucine, and proline have aliphatic R groups. A sulfur atom appears in the thioether-containing aliphatic side chain (—S—CH_3) of methionine. Its derivative *S*-adenosylmethionine (SAM, p. 546) is an important metabolite that serves as a methyl donor in numerous biochemical reactions. It should be noted that glycine with a hydrogen atom side chain instead of a hydrocarbon side chain is slightly hydrophilic. Its small side chain introduces structural flexibility into proteins. **Aromatic hydrocarbons** contain cyclic structures that constitute a class of

unsaturated hydrocarbons with planar conjugated π electron clouds. Benzene is one of the simplest aromatic hydrocarbons (p. P-25). Phenylalanine and tryptophan contain aromatic rings.

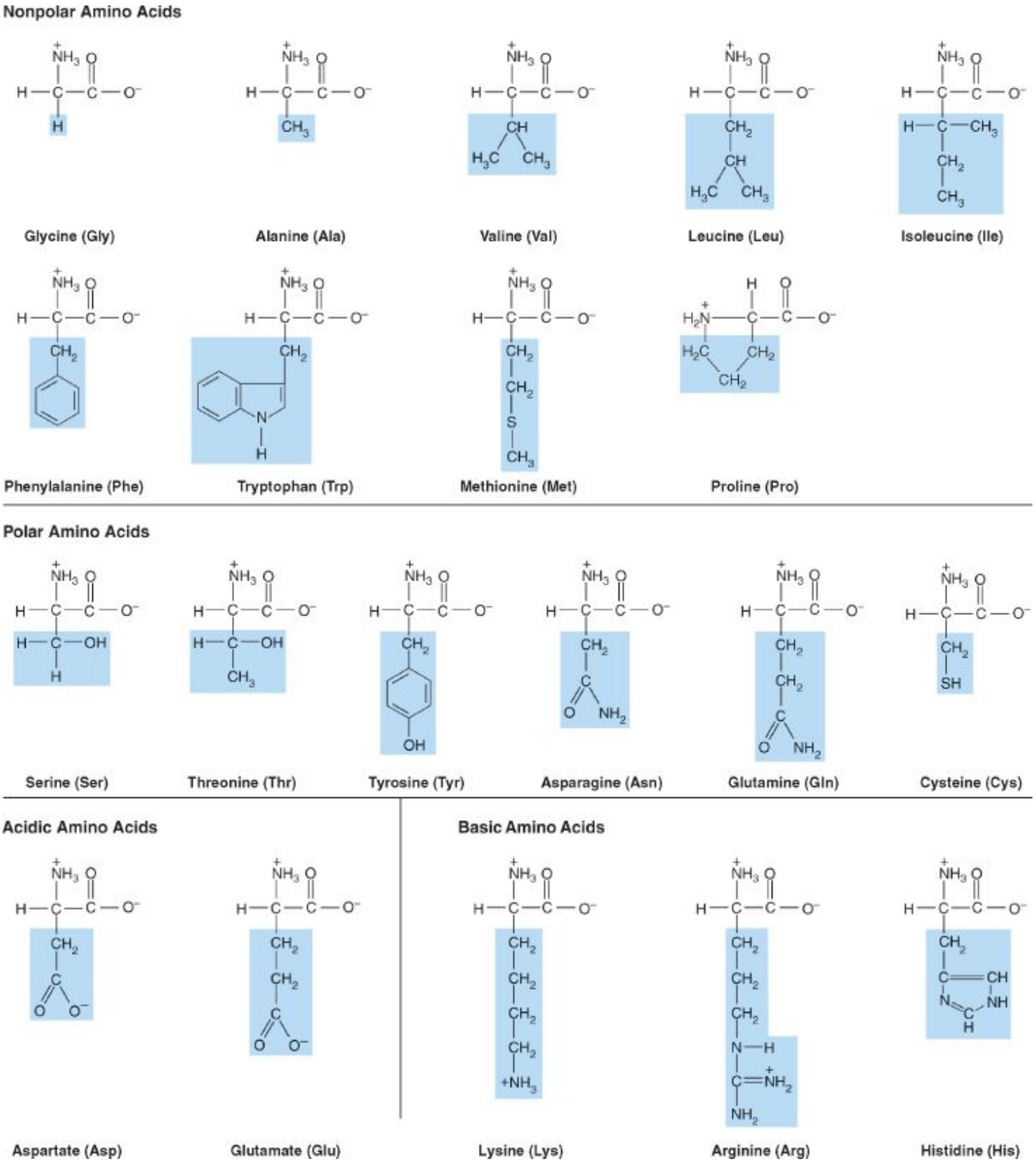


FIGURE 5.2

The Standard Amino Acids

The ionization state of the amino acid molecules in this illustration represents the dominant species that occur at a pH of 7. The side chains are indicated by shaded boxes.

Glycine (Gly)

Alanine (Ala)

Valine (Val)

Leucine (Leu)

Isoleucine (Ile)

Phenylalanine (Phe)

Tryptophan (Trp)

Methionine (Met)

Proline (Pro)

Serine (Ser)

Threonine (Thr)

Tyrosine (Tyr)

Asparagine (Asn)

Glutamine (Gln)

Cysteine (Cys)

Asparatate (Asp)

Glutamate (Glu)

Lysine (Lys)

Arginine (Arg)

Histidine (His)

TABLE 5.1 Names and Abbreviations of the Standard Amino Acids

Amino Acid	Three-Letter Abbreviation	One-Letter Abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N

Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

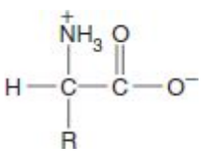


FIGURE 5.3

General Structure of the *α*-Amino Acids

POLAR AMINO ACIDS The functional groups of polar amino acids easily interact with water through electrostatic interactions, such as hydrogen bonding. Serine, threonine, tyrosine, asparagine, and glutamine belong to this category. Serine, threonine, and tyrosine contain a polar hydroxyl group, which enables them to participate in hydrogen bonding, an important factor in protein structure. The hydroxyl groups serve other functions in proteins. For example, the formation of the phosphate ester of tyrosine is a common protein regulatory mechanism. Additionally, the —OH groups of serine and threonine are points for attaching carbohydrate groups. Asparagine and glutamine are amide derivatives of the acidic amino acids aspartic acid and glutamic acid, respectively. Because the amide functional group is highly polar, the hydrogen-bonding capability of asparagine and glutamine has a significant effect on protein stability. The sulfhydryl group (—SH) of cysteine is highly reactive and is an important component of many enzymes. It also binds metals (e.g., iron and copper ions) in proteins. Additionally, the sulfhydryl groups of two cysteine molecules oxidize easily in the extracellular compartment to form a disulfide compound called

cystine. (See p. 147 for a discussion of this reaction.)

ACIDIC AMINO ACIDS Two standard amino acids have side chains with carboxylate groups. Aspartic acid and glutamic acid are often referred to as aspartate and glutamate because carboxyl groups are negatively charged at physiological pH.

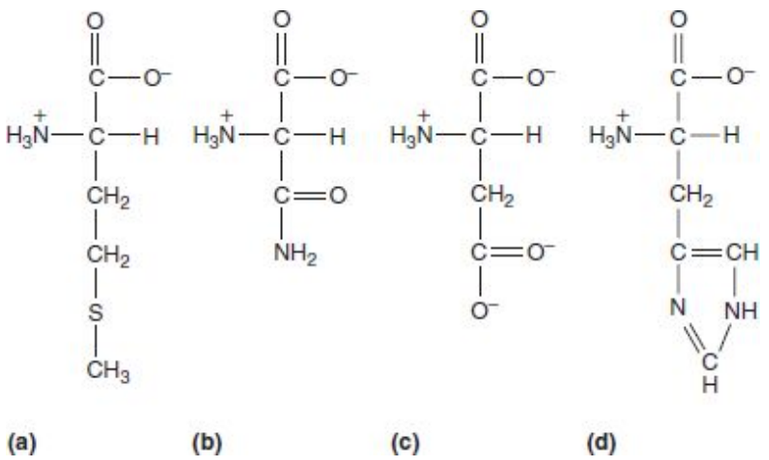
BASIC AMINO ACIDS Basic amino acids bear a positive charge at physiological pH. They can therefore form ionic bonds with acidic amino acids. Lysine, which has a side chain amino group, accepts a proton from water to form the conjugate acid ($-NH_3^+$). When lysine side chains in collagen fibrils, a vital structural component of ligaments and tendons, are oxidized and subsequently condensed, strong intramolecular and intermolecular cross-linkages are formed. Because the guanidino group of arginine has a pK_a range of 11.5 to 12.5 in proteins, it is permanently protonated at physiological pH and, therefore, does not function in acid–base reactions. (See pp. 93–96 for a description of pH and pK_a.) The imidazole side chain of histidine, however, is a weak base that is only partially ionized at pH 7 because its pK_a is approximately 6. Histidine's capacity under physiological conditions to accept or donate protons in response to small changes in pH plays an important role in the catalytic activity of numerous enzymes.

KEY CONCEPTS

Amino acids are classified according to their capacity to interact with water. This criterion may be used to distinguish four classes: nonpolar, polar, acidic, and basic.

QUESTION 5.1

Classify these standard amino acids according to whether their structures are nonpolar, polar, acidic, or basic.



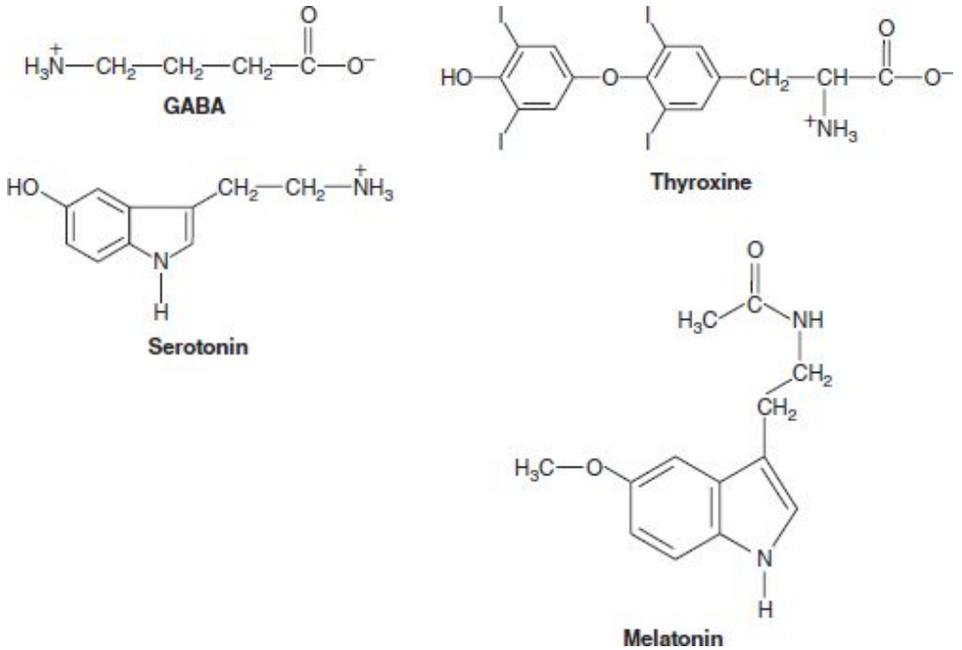


FIGURE 5.4

Some Derivatives of Amino Acids

Biologically Active Amino Acids

In addition to their primary function as components of protein, amino acids have several other biological roles.

1. Several α-amino acids or their derivatives act as chemical messengers **(Figure 5.4)**. For example, glycine, glutamate, γ-amino butyric acid (GABA, a derivative of glutamate), and serotonin and melatonin (derivatives of tryptophan) are **neurotransmitters**, substances released from one nerve cell that influence the function of a second nerve cell or a muscle cell. Thyroxine (a tyrosine derivative produced in the thyroid gland of animals) is a **hormone**—a chemical signal molecule produced in one cell that regulates the function of other cells.
2. Amino acids are precursors of a variety of complex nitrogen-containing molecules. Examples include the nitrogenous base components of nucleotides and the nucleic acids, heme (the iron-containing organic group required for the biological activity of several important proteins), and chlorophyll (a pigment of critical importance in photosynthesis).
3. Several standard and nonstandard amino acids act as metabolic intermediates. For example, arginine **(Figure 5.2)**, citrulline, and ornithine **(Figure 5.5)** are components of the urea cycle (Chapter 15). The synthesis of urea, a molecule formed in vertebrate livers, is the principal mechanism for the disposal of nitrogenous waste.

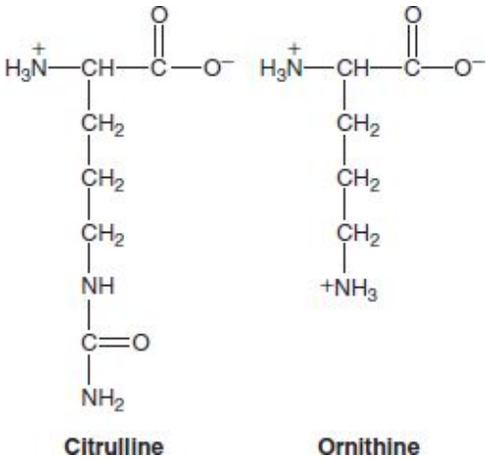


FIGURE 5.5

Citrulline and Ornithine Structure

Both of these molecules are intermediates in the urea cycle, which is described in Chapter 15.

3D animation of ornithine

3D animation of citrulline

Modified Amino Acids in Proteins

Several proteins contain amino acid derivatives that are formed after a polypeptide chain has been synthesized. Among these modified amino acids is γ-carboxyglutamic acid (**Figure 5.6**), a calcium-binding amino acid residue found in the blood-clotting protein prothrombin. Both 4-hydroxyproline and 5-hydroxylysine are important structural components of collagen, the most abundant protein in connective tissue. Phosphorylation of the hydroxyl-containing amino acids serine, threonine, and tyrosine is often used to regulate the activity of proteins. For example, the synthesis of glycogen is significantly curtailed when the enzyme glycogen synthase is phosphorylated.

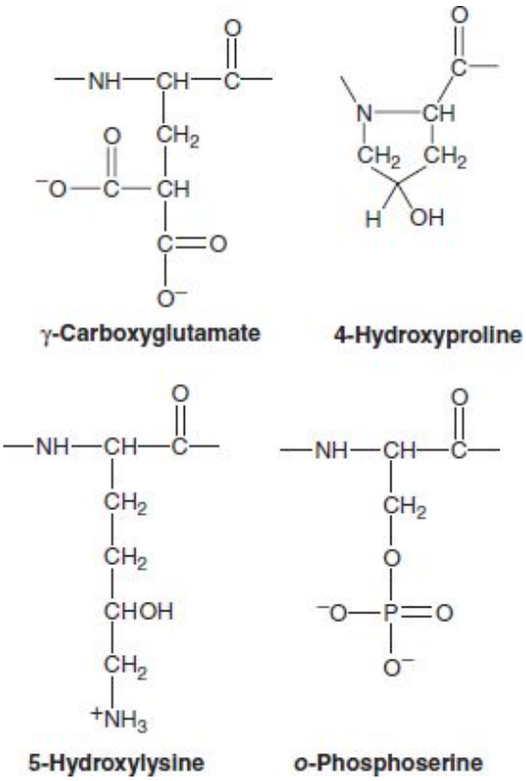


FIGURE 5.6

Some Modified Amino Acid Residues Found in Polypeptides

Amino Acid Stereoisomers

Because the α-carbons of 19 of the 20 standard amino acids are attached to four different groups (i.e., a hydrogen, a carboxyl group, an amino group, and an R group), they are referred to as **asymmetric**, or **chiral**, **carbons**. (Glycine is a symmetrical molecule because its α-carbon is attached to two hydrogens.) Molecules with chiral carbons can exist as **stereoisomers**, molecules that differ only in the spatial arrangement of their atoms. Three-dimensional representations of amino acid stereoisomers are illustrated in **Figure 5.7**. Notice in the figure that the atoms of the two isomers are bonded together in the same pattern except for the position of the ammonium group and the hydrogen atom. These two isomers are mirror images of each other. Such molecules, called **enantiomers**, cannot be superimposed on each other. Enantiomers have identical physical properties except that they rotate plane-polarized light in opposite directions. (Plane-polarized light is produced by passing unpolarized light through a special filter; the emitted light waves vibrate in only one plane.) Molecules that possess this property are called **optical isomers**.

Glyceraldehyde is the reference compound for optical isomers (**Figure 5.8**). One glyceraldehyde isomer rotates the light beam in a clockwise direction and is said to be dextrorotatory (designated by +). The other glyceraldehyde isomer, referred to as levorotatory (designated by −), rotates the beam in the opposite direction to an equal degree. Optical isomers are often designated as D or L (e.g., D-glucose, L-alanine) to indicate the similarity of the arrangement of atoms around a molecule's asymmetric carbon to the asymmetric carbon in either of the glyceraldehyde isomers.

Most biomolecules have more than one chiral carbon. As a result, the letters D and L refer only to a molecule's structural relationship to either of the glyceraldehyde isomers, not to the direction in which it rotates plane-polarized light. Most asymmetric molecules found in living organisms occur in only one stereoisomeric form, either D or L. For example, with few exceptions, only L-amino acids are found in proteins.

Chirality has a profound effect on the structural and functional properties of biomolecules. For

example, the right-handed helices observed in proteins (pp. 156–57) result from the exclusive presence of L-amino acids. Polypeptides synthesized in the laboratory from a mixture of both D- and L-amino acids do not form helices. In addition, because the enzymes are chiral molecules, most bind substrate (reactant) molecules in only one enantiomeric form. Proteases, enzymes that degrade proteins by hydrolyzing peptide bonds, cannot degrade artificial polypeptides composed of D-amino acids.

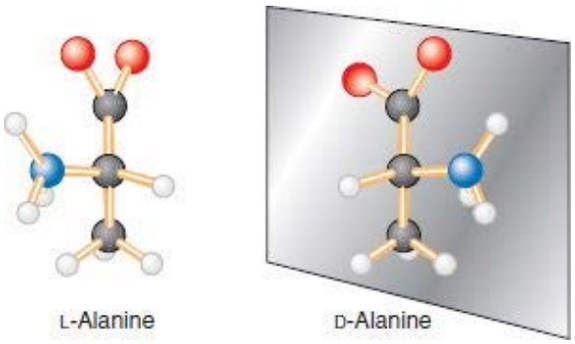


FIGURE 5.7

Two Enantiomers

L-Alanine and D-alanine are mirror images of each other. (Nitrogen = large blue ball; hydrogen = small gray balls; carbon = black balls; oxygen = red balls)

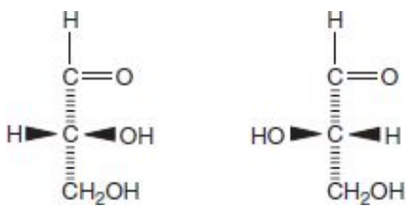


FIGURE 5.8

D- and L-Glyceraldehyde

These molecules are mirror images of each other.

QUESTION 5.2

Certain bacterial species have outer layers composed of polymers made of D-amino acids. Immune system cells, whose task is to attack and destroy foreign cells, cannot destroy these bacteria. Suggest a reason for this phenomenon.

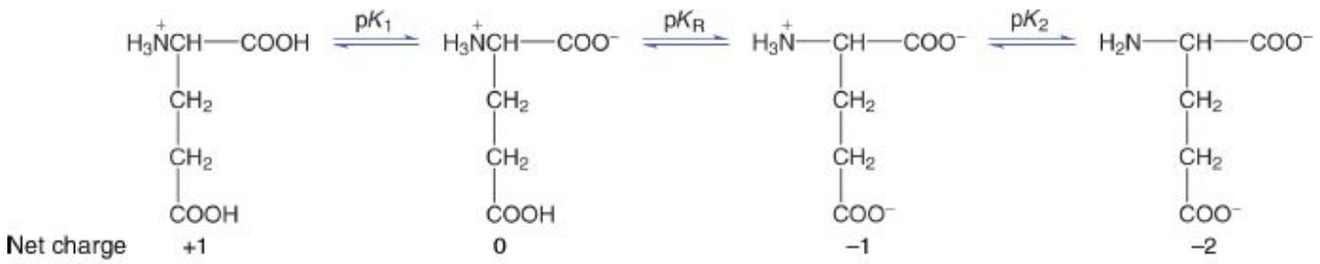
Amino Acid Ionizable Side Chain Groups and pH

The predominant ionic form of the ionizable groups of amino acids in solution depends on pH (**Table 5.2**). Seven standard amino acids have ionizable side chains: Asp, Glu, His, Lys, Arg, Tyr, and Cys. These ionizable groups, which play vital roles in protein structure, stability, binding properties, and function (e.g., enzyme catalysis, pp. 207–09), are sensitive to pH changes. Consider the titration of glutamate (**Figure 5.9**), an amino acid with a side chain carboxyl group, with a strong base such as NaOH. In a strongly acid solution (pH 0), the α-carboxyl, α-amino and side chain groups are

protonated, and the molecule's net charge is +1. (Proton loss is determined by an ionizable group's pK_a.) As NaOH is gradually added to the solution, the α-carboxyl group loses its proton to become a negatively charged carboxylate group. Glutamate now has no net charge. As the titration continues, the side chain carboxyl group loses its proton, yielding a molecule with a −1 charge. Finally, the α-ammonium group loses its charge with the addition of more NAOH, giving glutamate a −2 charge. Since glutamate side chains within proteins are usually negatively charged with roles in numerous enzymatic catalytic mechanisms and in protein structure and stability, a pH-induced change in glutamate's side chain charge may have physiological consequences. The structures of the various ionized forms of glutamic acid are:

KEY CONCEPTS

- Molecules with an asymmetric or chiral carbon atom differ only in the spatial arrangement of the atoms attached to the carbon.
- The mirror-image forms of a molecule are called enantiomers.
- Most asymmetric molecules in living organisms occur in only one stereoisomeric form.



Titration of glutamic acid

TABLE 5.2 pK_a Values for the Ionizing Groups of the Amino Acids

Amino Acid	pK_1 (—COOH)	pK_2 (—NH+3)	pK_R
Glycine	2.34	9.60	
Alanine	2.34	9.69	
Valine	2.32	9.62	
Leucine	2.36	9.60	
Isoleucine	2.36	9.60	
Serine	2.21	9.15	
Threonine	2.63	10.43	
Methionine	2.28	9.21	
Phenylalanine	1.83	9.13	
Tryptophan	2.83	9.39	
Asparagine	2.02	8.80	
Glutamine	2.17	9.13	

Proline	1.99	10.60	
Cysteine	1.71	10.78	8.33
Histidine	1.82	9.17	6.00
Aspartic acid	2.09	9.82	3.86
Glutamic acid	2.19	9.67	4.25
Tyrosine	2.20	9.11	10.07
Lysine	2.18	8.95	10.79
Arginine	2.17	9.04	12.48

When molecules with net charges are placed in an electric field, they move in a direction dictated by their net charge. Molecules with net positive charges will move toward the cathode (negatively charged electrode); those with net negative charges will move toward the cathode (positively charged electrode). The pH at which a molecule carries no net electric charge is referred to as the **isoelectric point (p*I*)**. At this pH, molecules such as amino acids and proteins are least soluble and will not move in an electric field. Biochemists often use a purification technique called *isoelectric focusing*, which can separate proteins by placing them into a gel with a pH gradient in an electric field. The proteins move along the gel until they reach their respective isoelectric point. Each protein is then removed so that it can be analyzed further. Problems 5.1 to 5.3 are sample titration problems.

KEY CONCEPTS

- Titration is useful in determining the relative ionization potential of side chain acidic and basic groups in an amino acid or peptide.
- The pH at which molecules such as amino acids and proteins have no net charge is called the isoelectric point (p*I*).

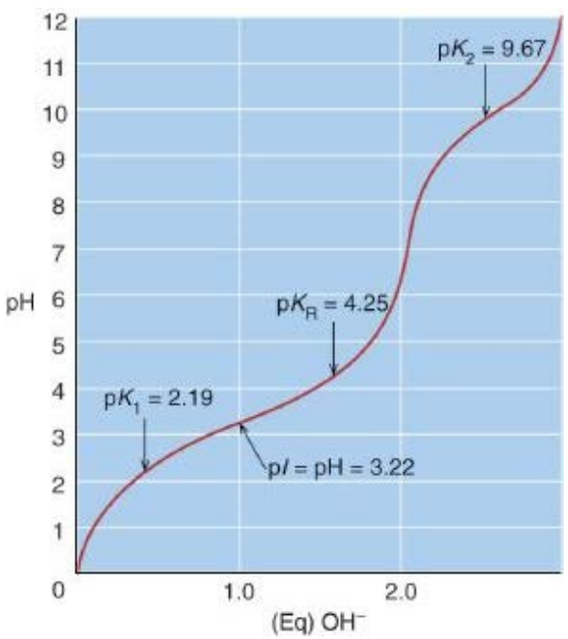


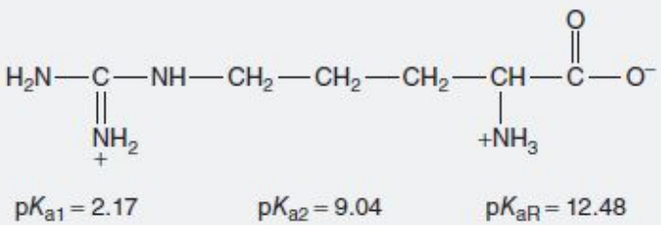
FIGURE 5.9

Titration of Glutamate

There are three stages in the titration of the free amino acid glutamate because it has three ionizable groups. Note that pK_a values are inflection points. At pK_a = 2.19, for example, half of the α-carboxylate groups are protonated and half are unprotonated. The isoelectric point (p*I*), the pH at which a molecule has no net charge, is calculated by adding the pK_a values that bracket the zwitterion and then dividing by two [(pK_{a1} + pK_{a2})/2]. In the case of glutamate, the p*I* is the pH halfway between the pK_a values for the two carboxyl groups: (2.19 + 4.25)/2 = 3.22.

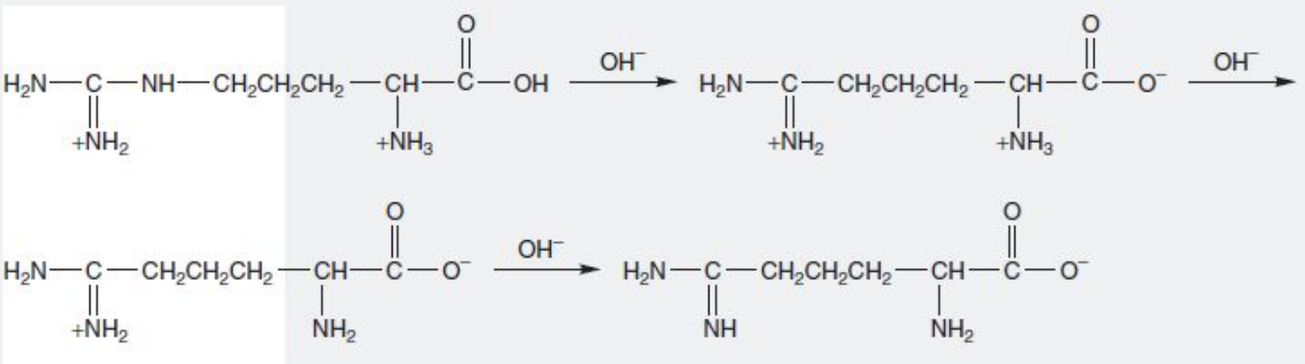
WORKED PROBLEM 5.1

Consider the following amino acid and its pK_a values:



Draw the structure of the amino acid as the pH of the solution changes from highly acidic to strongly basic.

SOLUTION



The ionizable hydrogens are lost in the order of acidity, with the most acidic ionizing first. ■

WORKED PROBLEM 5.2

a. Sketch the titration curve for the amino acid aspartic acid.

SOLUTION (A)

Plateaus appear at the pK_a and are centered about 0.5 equivalent (Eq), 1.5 Eq, and 2.5 Eq of base. There is a sharp rise at 1 Eq, 2 Eq, and 3 Eq. The isoelectric point is midway on the sharp rise between pK_{a1} and pK_{aR}. The pI value for aspartic acid is (2.09 + 3.86)/2 = 2.98.

b. In what direction does the amino acid move when placed in an electric field at the following pH values: 1, 3, 5, 7 and 9? *Choice 1*: does not move, *Choice 2*: toward the *cathode* (negative

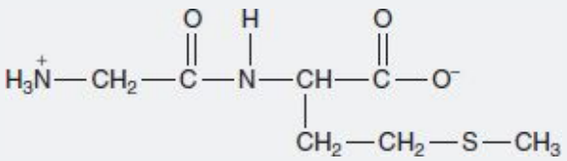
electrode), *Choice 3*: toward the *anode* (positive electrode).

SOLUTION (B)

At pH values below the p*I* (in this case 2.98), aspartic acid is positively charged and moves to the cathode. Therefore, the amino acid in this problem will move to the cathode at the pH values of 1. At a pH of 3, the amino acid does not move. The amino acid will be negatively charged at pH values above 3. Under this condition, aspartic acid will move to the anode at pH values of 5, 7, and 9. ■

WORKED PROBLEM 5.3

Consider the following dipeptide:



a. What is its isoelectric point?

SOLUTION (A)

The isoelectric point is the average of the pK_as of the amino group of glycine and the carboxyl group of methionine (obtained from Table 5.2).

$$pI = (9.60 + 2.28)/2 = 5.94$$

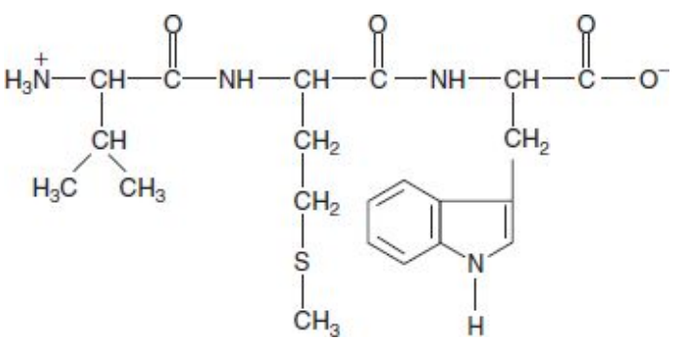
b. In which direction will the dipeptide move at pH 1, 3, 5, 7, 9, and 12?

SOLUTION (B)

At pH values below that of the p*I*, the dipeptide will move to the cathode (i.e., 1, 3, and 5). At pH values above the p*I*, the dipeptide will move to the anode. These are 7, 9, and 12. ■

QUESTION 5.3

Calculate the isoelectric point of the following tripeptide:

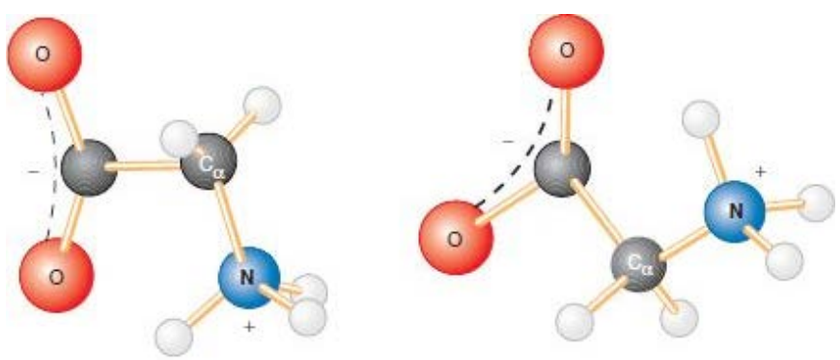


Assume that the pK_a values listed for the amino acids in Table 5.2 are applicable to this problem.

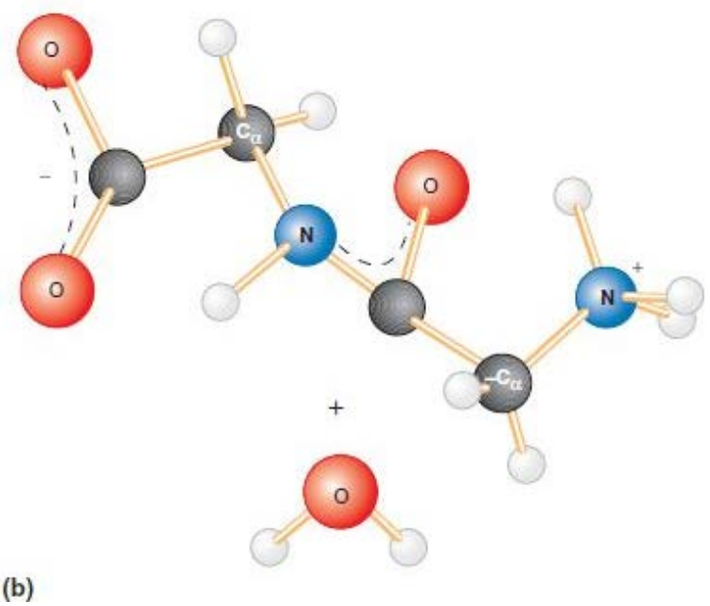
Amino Acid Reactions

The functional groups of organic molecules determine which reactions they may undergo. Amino acids with their carboxyl groups, amino groups, and various R groups can undergo numerous chemical reactions. Peptide bond and disulfide bridge formation, however, are of special interest because of their effect on protein structure. Schiff base formation is another important reaction.

PEPTIDE BOND FORMATION Polypeptides are linear polymers composed of amino acids linked together by peptide bonds. **Peptide bonds** (**Figure 5.10**) are amide linkages formed when the unshared electron pair of the α-amino nitrogen atom of one amino acid attacks the α-carboxyl carbon of another in a nucleophilic acyl substitution reaction (see pp. P-28). A generalized acyl substitution reaction is shown:



(a)



(b)

FIGURE 5.10

Formation of a Dipeptide

(a) A peptide bond forms when the α-carboxyl group of one amino acid reacts with the amino group of another. (b) A water molecule is formed in the reaction.

The linked amino acids in a polypeptide are referred to as *amino acid residues* because peptide bond formation is a dehydration reaction (i.e., a water molecule is removed). When two amino acid molecules are linked, the product is called a dipeptide. For example, glycine and serine can form the dipeptides glycylserine or serylglycine. As amino acids are added and the chain lengthens, the prefix reflects the number of residues: a tripeptide contains three amino acid residues, a tetrapeptide four, and so on. By convention, the amino acid residue with the free amino group is called the *N-terminal* residue and is written to the left. The free carboxyl group on the *C-terminal* residue appears on the right. Peptides are named using their amino acid sequences, beginning from their N-terminal residue. For example,

$$H_3\overset{+}{N}\text{— Tyr — Ala — Cys — Gly — } COO^-$$

is a tetrapeptide identified as tyrosylalanylcysteinylglycine.

Large polypeptides often have well-defined, three-dimensional structures. This structure, referred to as the molecule's native conformation, is a direct consequence of its *amino acid sequence* (the order in which the amino acids are linked together). Because all the linkages connecting the amino acid residues consist of single bonds, each polypeptide might be expected to undergo constant conformational changes caused by rotation around the single bonds. However, many polypeptides spontaneously fold into a single biologically active form. In the early 1950s, Linus Pauling (1901–1994, 1954 Nobel Prize in Chemistry) and his colleagues proposed an explanation. Using X-ray diffraction studies, they characterized the peptide bond (1.33 Å) as rigid and planar (flat) (**Figure 5.11**). Having discovered that the C—N bonds joining each two amino acids are shorter than other types of C—N bonds (1.45 Å), Pauling deduced that peptide bonds have a partial double-bond character. (This indicates that peptide bonds are resonance hybrids.) Because of the rigidity of the peptide bond, fully one-third of the bonds in a polypeptide backbone chain cannot rotate freely. Consequently, there are limits to the number of conformational possibilities.

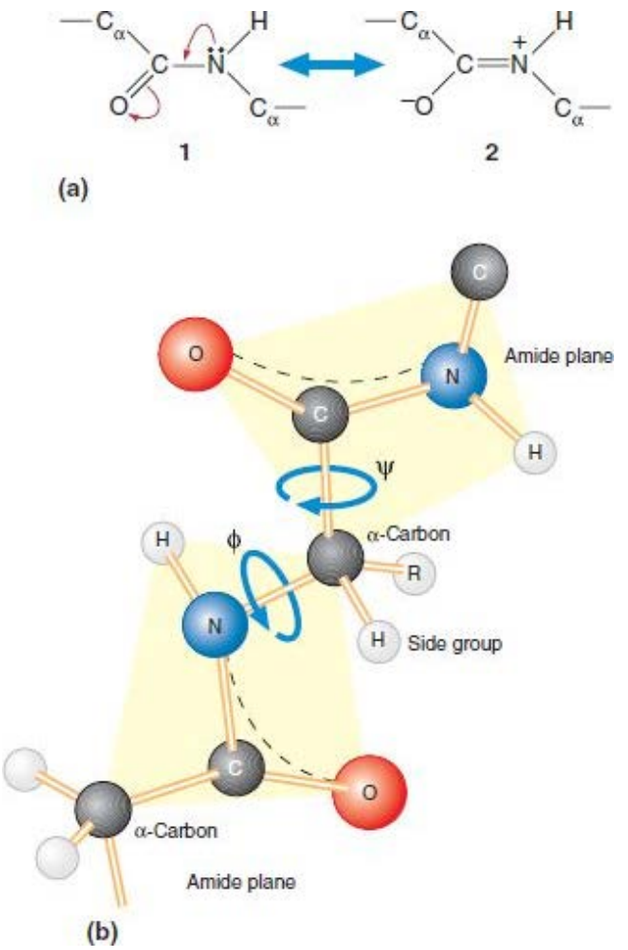


FIGURE 5.11

The Peptide Bond

(a) Resonance forms of the peptide bond. (b) Dimensions of a dipeptide. Because peptide bonds are rigid, the conformational degrees of freedom of a polypeptide chain are limited to rotations around the Cα—C and Cα—N bonds. The corresponding rotations are represented by Ψ and φ, respectively.

CYSTEINE OXIDATION The sulfhydryl group of cysteine is highly reactive. The most common reaction of this group is a reversible oxidation that forms a disulfide. Oxidation of two molecules of cysteine forms cystine, a molecule that contains a disulfide bond (**Figure 5.12**). When two cysteine residues form such a bond in peptides or polypeptides, it is referred to as a **disulfide bridge**. This bond can occur in a single chain to form a ring or between two separate chains to form an intermolecular covalent bond. Disulfide bridges help stabilize many polypeptides and proteins.

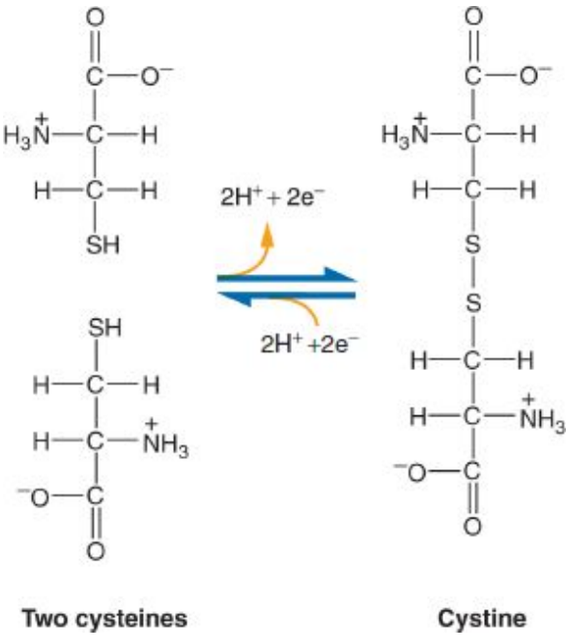


FIGURE 5.12

Oxidation of Two Cysteine Molecules to Form Cystine

The disulfide bond in a polypeptide is called a disulfide bridge.

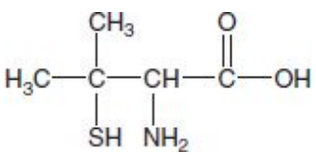


FIGURE 5.13

Structure of Penicillamine

Cystinuria

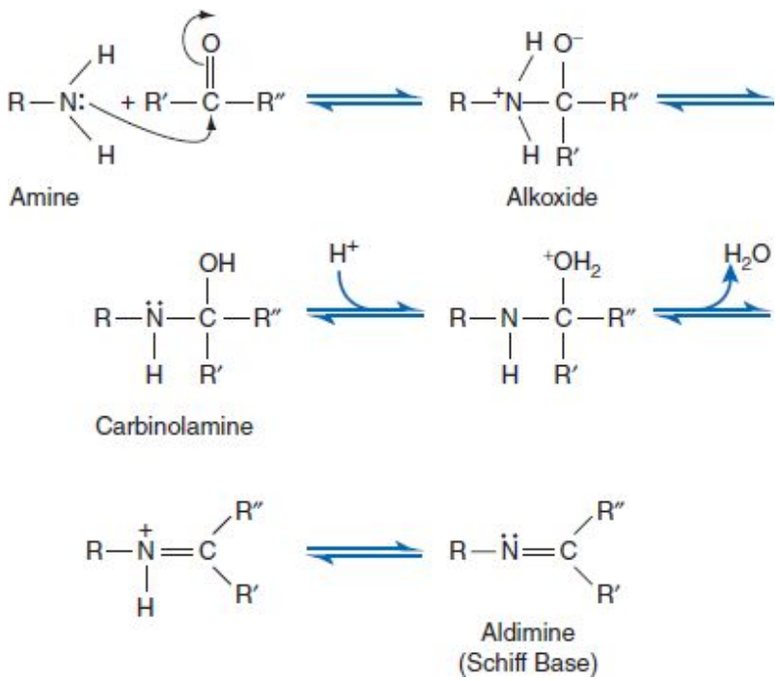
QUESTION 5.4

In extracellular fluids such as blood (pH 7.2–7.4) and urine (pH 6.5), the sulfhydryl groups of cysteine (pK_a 8.1) are subject to oxidation to form cystine. In peptides and proteins, thiol groups are used to advantage in stabilizing protein structure and in thiol transfer reactions, but free cystine in tissue fluids can be problematic because of its low solubility. In a genetic disorder known as *cystinuria*, defective membrane transport of cystine results in excessive excretion of cystine into the urine. Crystallization of the amino acid results in formation of calculi (stones) in the kidney, ureter, or urinary bladder. The stones may cause pain, infection, and blood in the urine. Cystine concentration in the kidney is reduced by massively increasing fluid intake and administering D-penicillamine. It is believed that penicillamine (**Figure 5.13**) is effective because penicillamine–cysteine disulfide, which is substantially more soluble than cystine, is formed. What is the structure of the penicillamine–cysteine disulfide?

SCHIFF BASE FORMATION Molecules such as amino acids that possess primary amine groups can reversibly react with carbonyl groups. The imine products of this reaction are often referred to as **Schiff bases**. In a *nucleophilic addition reaction*, an amine nitrogen attacks the electrophilic carbon of a carbonyl group to form an alkoxide product. The transfer of a proton from the amine group to the oxygen to form a carbinolamine, followed by the transfer of another proton from an acid catalyst, converts the oxygen into a good leaving group (OH_2^+). The subsequent elimination of a water molecule followed by loss of a proton from the nitrogen yields the imine product. The most important examples of Schiff base formation in biochemistry occur in amino acid metabolism. Schiff bases, referred to as **aldimines**, formed by the reversible reaction of an amino group with an aldehyde group, are *intermediates* (species formed during a reaction) in transamination reactions (pp. 533–36). Schiff bases are also intermediates in glycation (p. 256), nonenzymatic reactions that contribute to atherosclerosis, neurodegenerative diseases, and arthritis.

KEY CONCEPTS

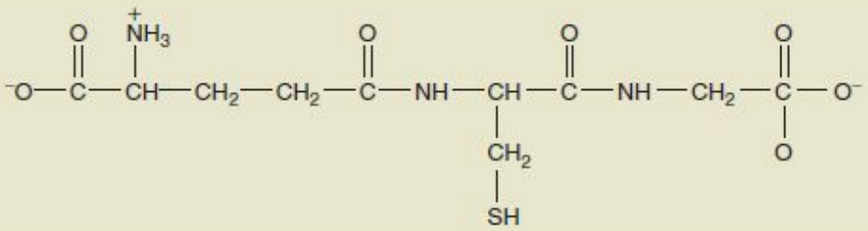
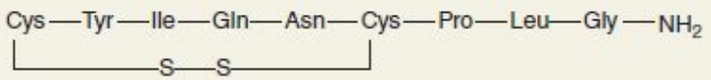
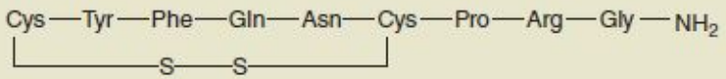
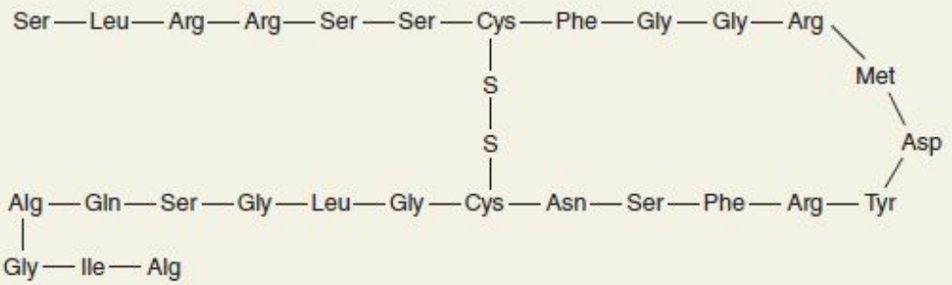
- Polypeptides are polymers composed of amino acids linked by peptide bonds. The order of the amino acids in a polypeptide is called the amino acid sequence.
- Disulfide bridges, formed by the oxidation of cysteine residues, are an important structural element in polypeptides and proteins.
- Schiff bases are imines that form when amine groups react reversibly with carbonyl groups.



5.2 PEPTIDES

Although less structurally complex than the larger protein molecules, peptides have significant biological activities. The structure and function of several interesting examples, presented in **Table 5.3**, are now discussed.

TABLE 5.3 Selected Biologically Important Peptides

Name	Amino Acid Sequence
Glutathione	
Oxytocin	
Vasopressin	
Atrial natriuretic factor	

The tripeptide *glutathione* (γ-glutamyl-l-cysteinylglycine) is a singularly important peptide that contains an unusual γ-amide bond; the γ-carboxyl group of the glutamic acid residue, not the α-carboxyl group, contributes to the peptide bond. Found in almost all organisms, glutathione (GSH; pp. 549–51) is involved in protein and DNA synthesis, drug and environmental toxin metabolism, amino acid transport, and other important biological processes. One group of glutathione's functions exploits its effectiveness as a reducing agent. Glutathione protects cells from the destructive effects of oxidation by reacting with substances such as peroxides (R–O–O–R), by-products of O_2 metabolism. For example, in red blood cells, hydrogen peroxide (H_2O_2) oxidizes the iron of hemoglobin to its ferric form (Fe^{3+}). Methemoglobin, the product of this reaction, is incapable of binding O_2. Glutathione protects against the formation of methemoglobin by reducing H_2O_2 in a reaction catalyzed by the enzyme glutathione peroxidase. In the oxidized product GSSG, two tripeptides are linked by a disulfide bond:

$$2\,GSH + H_2O_2 \rightarrow GSSG + 2H_2O$$

Because of the high GSH:GSSG ratio normally present in cells, glutathione is an important intracellular antioxidant. The abbreviation GSH is used because the reducing component of the molecule is the —SH group of the cysteine residue.

Vasopressin and Oxytonic

Multicellular organisms maintain a stable internal environment, a condition called *homeostasis*, via the complex and dynamic interplay between opposing physiological processes. They do so by utilizing signal molecules with opposing functions that together regulate numerous functions (e.g., blood pressure regulation). The roles of selected peptides that promote homeostasis are briefly described.

Blood pressure, the force exerted by blood against the walls of blood vessels, is influenced by two peptides called vasopressin and atrial natriuretic factor. *Vasopressin*, also called antidiuretic hormone (ADH), is a nonapeptide (it contains nine amino acid residues) in which the two cysteine residues form a disulfide bridge. It is synthesized in the hypothalamus, a small structure located at the base of the brain that regulates a wide variety of functions, including water balance, appetite, body temperature, and sleep. In most mammals, ADH contains arginine and is also referred to as arginine vasopressin (AVP). In response to low blood pressure or a high blood Na^+ concentration, osmoreceptors in the hypothalamus trigger vasopressin secretion. Vasopressin stimulates water reabsorption in the kidneys by initiating a signal transduction mechanism that inserts aquaporins (water channels, p. 430) into kidney tubule membrane. Blood pressure rises as water then flows down its concentration gradient through the tubule cells and back into the blood. Vasopressin also has roles in parenting behavior and social bonding. High levels in males are associated with increased aggressive behavior.

Atrial natriuretic factor (*ANF*), a peptide produced by specialized cells in the upper chambers (atria) of the heart in response to stretching of the atrial wall (high blood volume), high serum Na^+ concentrations, and exercise, stimulates the production of a dilute urine, an effect opposite to that of vasopressin. ANF exerts its effect, in part, by increasing the excretion of Na^+, a process that causes increased excretion of water. ANF also inhibits the release of molecules that promote increased blood pressure. Examples include renin, aldosterone, and vasopressin. (Renin is an enzyme that catalyzes the formation of angiotensin, a hormone that constricts blood vessels. Aldosterone is a steroid that promotes Na^+ and water retention by the kidney.)

The structures of vasopressin and oxytocin, another peptide produced by the hypothalamus, illustrate an interesting feature of biomolecules: similar structures may result in overlapping functions. *Oxytocin*, the signal molecule that stimulates the ejection of milk by mammary glands during lactation and stimulates the contraction of uterine muscle during childbirth, is also associated with trust and empathy and has modulating roles in certain social behaviors. Examples of these behaviors include bonding between mating pairs, mothers and babies, and members of an in-group. As illustrated in Table 5.3, vasopressin and oxytocin have similar structures (the phenylalanine and arginine residues in vasopressin are replaced by isoleucine and leucine residues, respectively, in oxytocin). It is not surprising that the functions of the two molecules overlap. Oxytocin has mild antidiuretic activity, and vasopressin has some oxytocin-like activity. Both oxytocin and vasopressin signaling deficits have been linked to social impairment in children with autism spectrum disorder.

QUESTION 5.5

Write out the complete structure of oxytocin. What would be the net charge on this molecule at the average physiological pH of 7.3? At pH 4? At pH 9? Indicate which atoms in oxytocin can potentially form hydrogen bonds with water molecules.

QUESTION 5.6

The structural features of vasopressin that allow binding to vasopressin receptors are the rigid hexapeptide ring and the amino acid residues at positions 3 (Phe) and 8 (Arg). The aromatic phenylalanine side chain, which fits into a hydrophobic pocket in the receptor, and the large positively charged arginine side chain are especially important structural features. Compare the structures of vasopressin and oxytocin and explain why their functions overlap. Can you suggest what will happen to the binding properties of vasopressin if the arginine at position 8 is replaced by lysine?

KEY CONCEPTS

Although small in comparison to larger protein molecules, peptides have significant biological activity. They are involved in a variety of signal transduction processes.

5.3 PROTEINS

With the exception of water molecules, proteins are the most abundant biomolecules in living organisms. The number of proteins per cell varies widely: bacteria such as *E. coli* have about 2 million proteins per cell, whereas the larger eukaryotic cells may have as many as 3 billion. The number of individual copies of proteins in cells also differs widely, ranging from 10 molecules of β-galactosidase (the bacterial enzyme that converts the energy source lactose into galactose and glucose) in the absence of lactose to 250 million molecules of hemoglobin in human red blood cells. Proteins can be classified according to their functions, amino acid sequence similarities, and overall-three-dimensional form and composition.

Functions

Of all the molecules encountered in living organisms, proteins have the most diverse functions, as the following list suggests.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay on protein poisons.

1. **Catalysis.** Catalytic proteins called the *enzymes* accelerate thousands of biochemical reactions in such processes as digestion, energy capture, and biosynthesis. These molecules have remarkable properties. Enzymes can increase reaction rates by factors of between 10^6 and 10^{12}. They can perform this feat under mild conditions of pH and temperature because they can induce or stabilize strained reaction intermediates. Ribulose bisphosphate carboxylase is an important enzyme in photosynthesis. Nitrogenase is responsible for nitrogen fixation.
2. **Structure.** Structural proteins often have specialized properties. For example, collagen (the major component of connective tissues) and fibroin (silkworm protein) have significant mechanical strength. Elastin, the rubber-like protein found in elastic fibers, is found in blood vessels and skin that must be elastic to function properly.
3. **Movement.** Proteins are involved in all cell movements. Cytoskeletal proteins such as actin, tubulin, and their associated proteins, for example, are active in cell division, endocytosis, exocytosis, and the ameboid movement of white blood cells.

4. **Defense.** A wide variety of proteins are protective. In vertebrates, keratin, a protein found in skin cells, aids in protecting the organism against mechanical and chemical injury. The blood-clotting proteins fibrinogen and thrombin prevent blood loss when blood vessels are damaged. The immunoglobulins (or antibodies), produced by lymphocytes, protect against invasion of the body by foreign organisms such as bacteria.
5. **Regulation.** The binding of a hormone molecule or a growth factor to cognate receptors on its target cell changes cellular function. For example, insulin and glucagon are peptide hormones that regulate blood glucose levels. Growth hormone stimulates cell growth and division.
6. **Transport.** Many proteins function as carriers of molecules or ions across membranes or between cells. Examples of membrane transport proteins include the enzyme Na^+-K^+ ATPase and the glucose transporter. Other transport proteins include hemoglobin, which carries O_2 to the tissues from the lungs, and the lipoproteins LDL and HDL, which transport water-insoluble lipids in blood.
7. **Storage.** Certain proteins serve as a reservoir of essential nutrients. For example, ovalbumin in bird eggs and casein in mammalian milk are rich sources of organic nitrogen during development. Plant proteins such as zein perform a similar role in germinating seeds.
8. **Stress response.** The capacity of living organisms to survive abiotic stresses is mediated by a variety of proteins. For example, cytochrome P_{450} is a diverse group of enzymes found in animals and plants that usually convert a variety of toxic organic contaminants into less toxic derivatives. Excessively high temperatures and other stresses result in the synthesis of a class of proteins called the **heat shock proteins** (hsps), which promote the correct refolding of damaged proteins. When proteins are severely damaged, hsps promote their degradation. (Certain hsps function in the normal process of protein folding.) Cells are protected from radiation by DNA repair enzymes.
9. **Toxins.** Many organisms produce protein toxins, which in general are used in predation or in defense. For example, snakes, scorpions, and some spiders use neurotoxins to subdue prey animals. The bacterium *Clostridium botulinum* secretes botulinum toxin, which causes muscle paralysis.

MULTIFUNCTIONAL PROTEINS As a result of early research on protein structure and genetics, biochemists have long assumed that each gene encodes one protein, which in turn has a single function. In recent years it has become apparent that numerous polypeptides, referred to as *multifunctional proteins*, have multiple and autonomous functions. *Moonlighting proteins* are a subset of multifunctional proteins that are not the result of gene fusion or alternate splicing (a process whereby several proteins can be synthesized from a single gene by removing certain segments from an mRNA; see pp. 727–30). Most moonlighting proteins were first identified as enzymes or receptors. Subsequent research demonstrated that these polypeptides have one or more unexpected roles.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a well-researched example of a moonlighting protein. As its name suggests, GAPDH is an enzyme that catalyzes the oxidation of glyceraldehyde-3-phosphate, an intermediate in glucose catabolism (p. 288). The GAPDH polypeptide in eukaryotes has an extraordinary array of functions. Examples include roles in DNA synthesis and repair, microtubule bundling, apoptosis initiation, vesicular transport, and nuclear tRNA export. Several pathogenic bacteria use GAPDH as a virulence factor. After its export to the bacterial cell surface, GAPDH, acting as a proteolytic enzyme, facilitates the invasion of host tissue.

PROTEIN FAMILIES **Protein families** are composed of protein molecules that are related by amino acid sequence similarity. Such proteins share an obvious common ancestry. The hemoglobins (blood

oxygen transport proteins; see pp. 176–80) are a classic protein family. Proteins more distantly related are often classified into **superfamilies**. For example, the globin superfamily includes a variety of heme-containing proteins that serve in the binding and/or transport of oxygen. In addition to the hemoglobins and myoglobins (oxygen-binding proteins in muscle cells), the globin superfamily includes neuroglobin and cytoglobin (oxygen-binding proteins in brain and other tissues, respectively) and the leghemoglobins (oxygen-sequestering proteins in the root nodules of leguminous plants).

Shape and Composition

Proteins are often classified according to their shape and composition. There are two major protein shapes. As the name suggests, **fibrous proteins** are long, rod-shaped molecules that are insoluble in water and physically tough. Fibrous proteins, such as the keratins found in skin, hair, and nails, have structural and protective functions. **Globular proteins** are compact spherical molecules that are usually water-soluble. Typically, globular proteins have dynamic functions. For example, nearly all enzymes have globular structures. Other examples include the immunoglobulins (antibodies) and the transport proteins hemoglobin and albumin (a carrier of fatty acids in blood).

On the basis of composition, proteins are classified as simple or conjugated. Simple proteins, such as serum albumin and keratin, contain only amino acids. Each **conjugated protein** consists of a simple protein combined with a nonprotein component called a **prosthetic group**. (A protein without its prosthetic group is called an **apoprotein**. A protein molecule combined with its prosthetic group is referred to as a **holoprotein**.) Prosthetic groups typically play an important, even crucial, role in the function of proteins. Conjugated proteins are classified according to the nature of their prosthetic groups. For example, **glycoproteins** contain a carbohydrate component, **hemoproteins** contain heme groups (p. 177), and **lipoproteins** contain lipid molecules. Examples of proteins with inorganic cofactors include **metalloproteins**, which contain metal ions, and **phosphoproteins**, which possess phosphate groups.

Protein Structure

Protein research efforts through most of the twentieth century strongly supported the protein structure–function paradigm: each nascent polypeptide folds into a highly ordered structure, a three-dimensional conformation that determines its function. For many proteins, this structure–function relationship holds true. However, over the past two decades, it has become increasingly apparent that some proteins lack a unique three-dimensional structure, either in whole or in part. It is now estimated that at least 30% of eukaryotic proteins are completely or mostly disordered and about 50% have at least one disordered region. Estimates of prokaryotic disordered proteins are considerably lower at approximately 2% (Archaea) and 4 % (Bacteria).

This section of the chapter begins with descriptions of the structural features of classic ordered proteins and with brief overviews of disordered proteins, protein denaturation, and protein folding. This section ends with descriptions of the structural and functional features of examples of the two major categories of ordered proteins: fibrous and globular.

Proteins are extraordinarily complex molecules. Complete models depicting even the smallest of the polypeptide chains are almost impossible to comprehend. Simpler images that highlight specific features of a molecule are useful. Two methods of conveying structural information about proteins are presented in **Figure 5.14**. Another structural representation, referred to as a ball-and-stick model, is presented later (**Figures 5.36** and **5.39**).

CLASSIC PROTEIN STRUCTURE Biochemists have distinguished several levels of the structural

organization of proteins. **Primary structure**, the amino acid sequence, is specified by genetic information. As the *nascent* (newly synthesized) polypeptide chain folds, it forms certain localized arrangements of adjacent (but not necessarily contiguous) amino acids that constitute **secondary structure**. The overall three-dimensional shape that a polypeptide assumes is called the **tertiary structure**. Proteins that consist of two or more polypeptide chains (or subunits) are said to have a **quaternary structure**.

PRIMARY STRUCTURE Every polypeptide has a specific linear sequence of amino acids that are linked by peptide bonds. Polypeptides that have similar amino acid sequences and have arisen from the same ancestral gene are said to be **homologous**. The amino acid residues that are identical in all homologues of a protein, referred to as *invariant*, are presumed to be essential for the protein's function. (In cytochrome c, the invariant residues interact with heme, a prosthetic group, or certain other proteins involved in energy generation.) For conventional polypeptides, the interactions between amino acid residues determine the protein's three-dimensional structure and its functional role and relationship to other proteins.

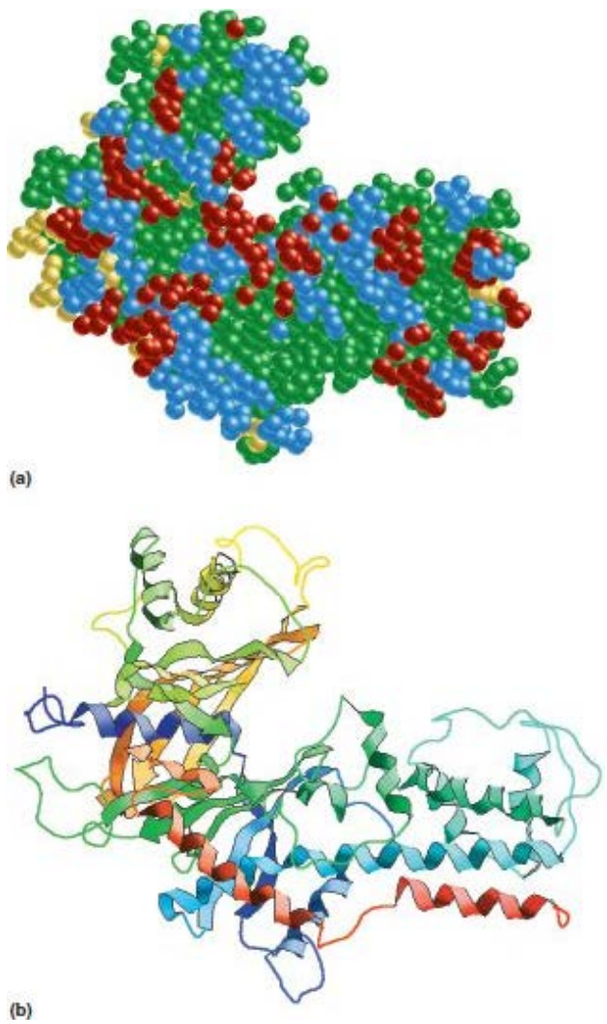


FIGURE 5.14

The Enzyme Adenylate Kinase

(a) This space-filling model illustrates the volume occupied by molecular components and overall shape. (b) In a ribbon model, α-helices appear as spiral ribbons, and flat arrows represent β-strands. The α-helices and β-strands are described on p. 156.

PRIMARY STRUCTURE, EVOLUTION, AND MOLECULAR DISEASES Over time, as the result of evolutionary processes, the amino acid sequences of polypeptides change. These modifications are caused by random and spontaneous alterations in DNA sequences called *mutations*. A significant number of primary sequence changes do not affect a polypeptide's function. Some of these substitutions are said to be *conservative* because an amino acid with a chemically similar side chain is substituted. For example, at certain sequence positions leucine and isoleucine, both of which contain hydrophobic side chains, may be substituted for each other without affecting function. Some sequence positions are significantly less stringent. These residues, referred to as *variable*, apparently perform nonspecific roles in the polypeptide's function.

KEY CONCEPTS

- The primary structure of a polypeptide is its amino acid sequence. The amino acids are connected by peptide bonds.
- Amino acid residues that are essential for the molecule's function are referred to as invariant.
- Proteins with similar amino acid sequences and functions and a common origin are said to be homologous.

Substitutions at conservative and variable sites have been used to trace evolutionary relationships. For example, the sequence homologies of the mitochondrial redox protein cytochrome c have been used to trace the evolution of species. These studies assume that the longer the time since two species diverged from each other, the larger the number of differences in a polypeptide's primary structure. For example, humans and chimpanzees are believed to have diverged perhaps only 4 million years ago. This presumption, based principally on fossil and anatomical evidence, is supported by cytochrome c primary sequence data indicating that this protein is identical in both species. Kangaroos, whales, and sheep, whose cytochrome c molecules each differ by 10 residues from the human protein, are believed to have evolved from a common ancestor that lived more than 50 million years ago. It is interesting to note that quite often a polypeptide's overall three-dimensional structure does not change despite numerous amino acid sequence changes.

Molecular Diseases

Mutations, however, can be deleterious. Changes in gene sequence can have moderate to severe effects on health. Individual organisms with amino acid substitutions at the conservative, invariant residues of cytochrome c, for example, are not viable. Mutations can also have a profound effect without being immediately lethal. Sickle-cell anemia, which is caused by mutant hemoglobin, is a classic example of a group of maladies that Linus Pauling and his colleagues referred to as **molecular diseases**. (Dr. Pauling first demonstrated that sickle-cell patients have a mutant hemoglobin through the use of electrophoresis, a technique described on p. 186.) Human adult hemoglobin (HbA) is composed of two identical α-globin chains and two identical β-globin chains. Sickle-cell anemia results from a single amino acid substitution in the β-globin chain of HbA. Analysis of the hemoglobin molecules of sickle-cell patients reveals that the only difference between HbA and sickle-cell hemoglobin (HbS) is at amino acid residue 6 in the β-chain (**Figure 5.15**). Because of the substitution of a hydrophobic valine for a negatively charged glutamate, HbS molecules aggregate to form rigid rod-like structures in the oxygen-free state (**Figure 5.16**). The patient's red blood cells become sickle shaped and are susceptible to hemolysis, resulting in severe anemia. These red blood cells have an abnormally low oxygen-binding capacity. Intermittent

clogging of capillaries by rigid sickled cells also causes tissues to be deprived of oxygen. Sickle-cell anemia is characterized by excruciating pain, organ damage, and eventually death.

Hb A	Val	His	Leu	Thr	Pro	Glu	Glu	Lys
Hb S	Val	His	Leu	Thr	Pro	Val	Glu	Lys
	1	2	3	4	5	6	7	8

FIGURE 5.15

Segments of the *β*-Globin Chain in HbA and HbS

Individuals possessing the gene for sickle-cell hemoglobin produce *β*-chains with valine instead of glutamic acid at residue 6.

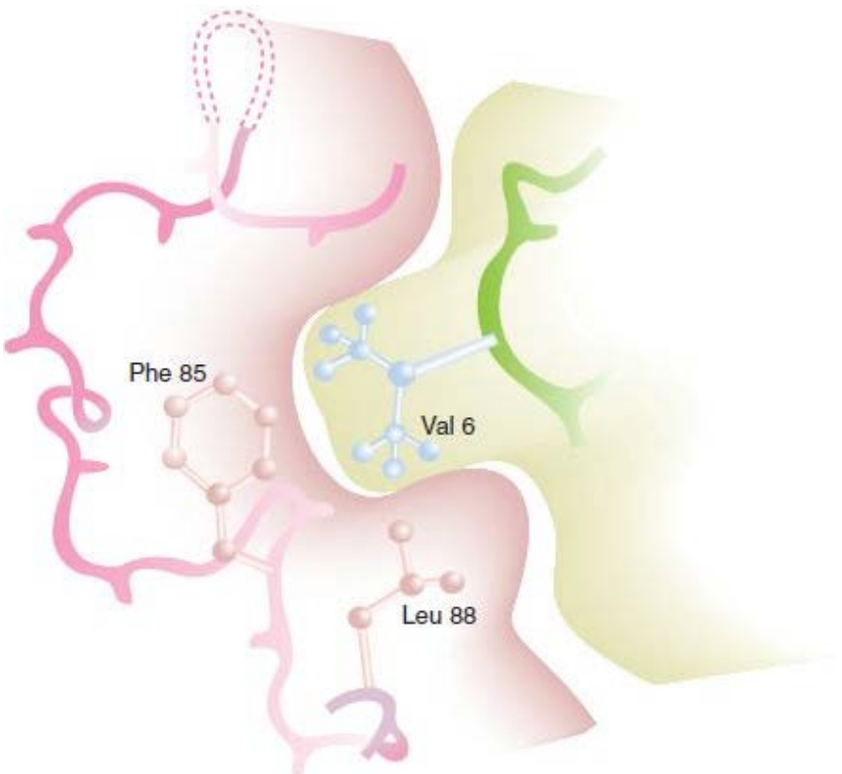


FIGURE 5.16

Sickle-Cell Hemoglobin

HbS molecules aggregate into rod-like filaments because the hydrophobic side chain of valine, the substituted amino acid in the *β*-chain, interacts with a hydrophobic pocket in a second hemoglobin molecule.

Until recently, because of the debilitating nature of sickle-cell disease, affected individuals rarely survived beyond childhood. Thus, one might predict that the deleterious mutational change that causes this affliction would be rapidly eliminated from human populations. However, the sickle-cell gene is not as rare as would be expected. Sickle-cell disease is a homozygous recessive illness; that is, it occurs only in individuals who have inherited two copies of the sickle-cell gene. The term *homozygous* indicates that the affected individual has inherited one copy of the defective gene from each parent. Each of the parents is said to have the *sickle-cell trait*. Such people are referred to as *heterozygous* because they have one normal HbA *β*-chain gene and one defective HbS *β*-chain gene. Except for athletes enduring heat and exercise-induced dehydration or exhaustion, most individuals with the sickle-cell trait are relatively symptom-free, even though about 40% of their hemoglobin is

HbS. The incidence of sickle-cell trait is especially high in some regions of Africa. In these areas malaria, caused by the *Anopheles* mosquito-borne parasite *Plasmodium*, is a serious health problem. Individuals with the sickle-cell trait are less vulnerable to malaria because their red blood cells are a less favorable environment for the growth of the parasite than are normal cells. Infected red blood cells, for example, produce levels of oxygen radicals that are toxic to parasites. Individuals with the sickle-cell trait have levels of sickled cells that are sufficiently low that cells and tissues of the reticuloendothelial system (e.g., phagocytes, spleen, and bone marrow) can selectively destroy the infected cells. Because sickle-cell trait carriers are more likely to survive malaria than normal individuals, the incidence of the sickle-cell gene has remained high. (In some areas, where malaria is endemic, the sickle-cell trait is present in as much as 40% of the native population.)

QUESTION 5.7

A genetic disease called *glucose-6-phosphate dehydrogenase deficiency* is inherited in a manner similar to that of sickle-cell anemia except that it occurs most frequently in males. The defective enzyme cannot keep erythrocytes supplied with sufficient amounts of the antioxidant molecule NADPH (Chapter 8). NADPH protects cell membranes and other cellular structures from oxidation. Describe in general terms the inheritance pattern of this molecular disease. Why do you think that the antimalarial drug primaquine, which stimulates peroxide formation, results in devastating cases of hemolytic anemia in carriers of the defective gene? Does it surprise you that this genetic anomaly is commonly found in African and Mediterranean populations?

SECONDARY STRUCTURE The secondary structure of polypeptides consists of several repeating patterns. The most commonly observed types of secondary structure are the α-helix and the β-pleated sheet. Both α-helix and β-pleated sheet patterns are stabilized by localized hydrogen bonding between the carbonyl and N—H groups in the polypeptide's backbone. Because peptide bonds are rigid, the α-carbons are swivel points for the polypeptide chain. Several properties of the R groups (e.g., size and charge, if any) attached to the α-carbon influence the φ and Ψ angles. Certain amino acids foster or inhibit specific secondary structural patterns. Many fibrous proteins are composed almost entirely of secondary structural patterns.

α-Helix. The *α-helix* is a rigid, rod-like structure that forms when a polypeptide chain twists into a right-handed helical conformation (**Figure 5.17**). Hydrogen bonds form between the N—H group of each amino acid and the carbonyl oxygen of the amino acid four residues away. There are 3.6 amino acid residues per turn of the helix, and the pitch (the distance between corresponding points per turn) is 0.54 nm. Amino acid R groups extend outward from the helix. Because of several structural constraints (i.e., the rigidity of peptide bonds and the allowed limits on the values of the φ and Ψ angles), certain amino acids do not foster α-helical formation. For example, glycine's R group (a hydrogen atom) is so small that the polypeptide chain may be too flexible. Proline, on the other hand, contains a rigid ring that prevents the N—*C*α bond from rotating. In addition, proline has no N—H group available to form the intrachain hydrogen bonds that are crucial in α-helix structure. Large numbers of charged amino acids (e.g., glutamate and aspartate) and bulky R groups (e.g., tryptophan) are also incompatible with α-helix structures.

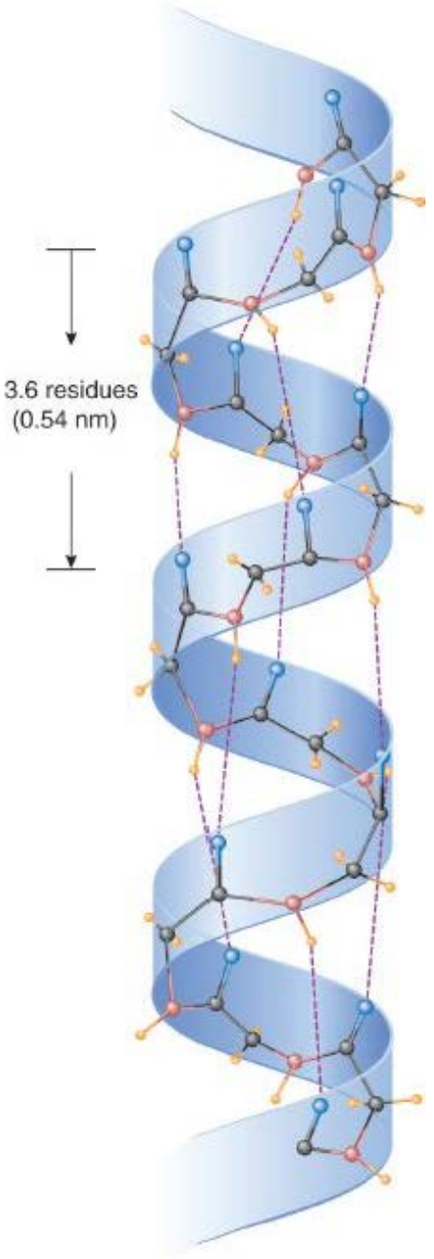


FIGURE 5.17

The *α*-Helix

Hydrogen bonds form between carbonyl and N—H groups along the long axis of the *α*-helix. Note that there are 3.6 residues per turn of the helix, which has a pitch of 0.54 nm. The hyphenated lines represent hydrogen bonds.

***β*-Strands.** *β-Strands* are a second type of secondary structure in which a peptide segment is fully extended into a zigzag-like structure. Solitary *β*-strands are rarely found in proteins because they are unstable. However, when two or more *β*-strands line up side by side, they form very stable structures called *β-pleated sheets* (**Figure 5.18**). *β*-Pleated sheets are stabilized by hydrogen bonds that form between the polypeptide backbone N—H and carbonyl groups of adjacent chains or chain segments. *β*-Pleated sheets are either parallel or antiparallel. In parallel *β*-pleated sheet structures, the hydrogen bonds in the polypeptide chains are arranged in the same direction; in antiparallel chains, these bonds are arranged in opposite directions. Occasionally, mixed parallel–antiparallel *β*-sheets are observed.

Supersecondary Motifs. Many globular proteins contain combinations of *α*-helix and *β*-strand secondary structures (**Figure 5.19**). These patterns are called **supersecondary structures** or **motifs**.

In the *βαβ unit*, two parallel *β*-strand segments are connected by an *α*-helix segment. The structure of *βαβ* units, typically found in ion channels, is stabilized by hydrophobic interactions between nonpolar side chains projecting from the interacting surfaces of the *β*-strands and the *α*-helix. Abrupt changes in direction of a polypeptide involve structural elements called loops. The *β*-hairpin loop, a commonly observed type of loop, is a 180° turn commonly involving two to seven residues between antiparallel hydrogen-bonded *β*-strands. In the frequently observed four-residue *β*-hairpin, for example, the carbonyl oxygen of the first residue in the loop forms a hydrogen bond with the amide hydrogen of the fourth residue.

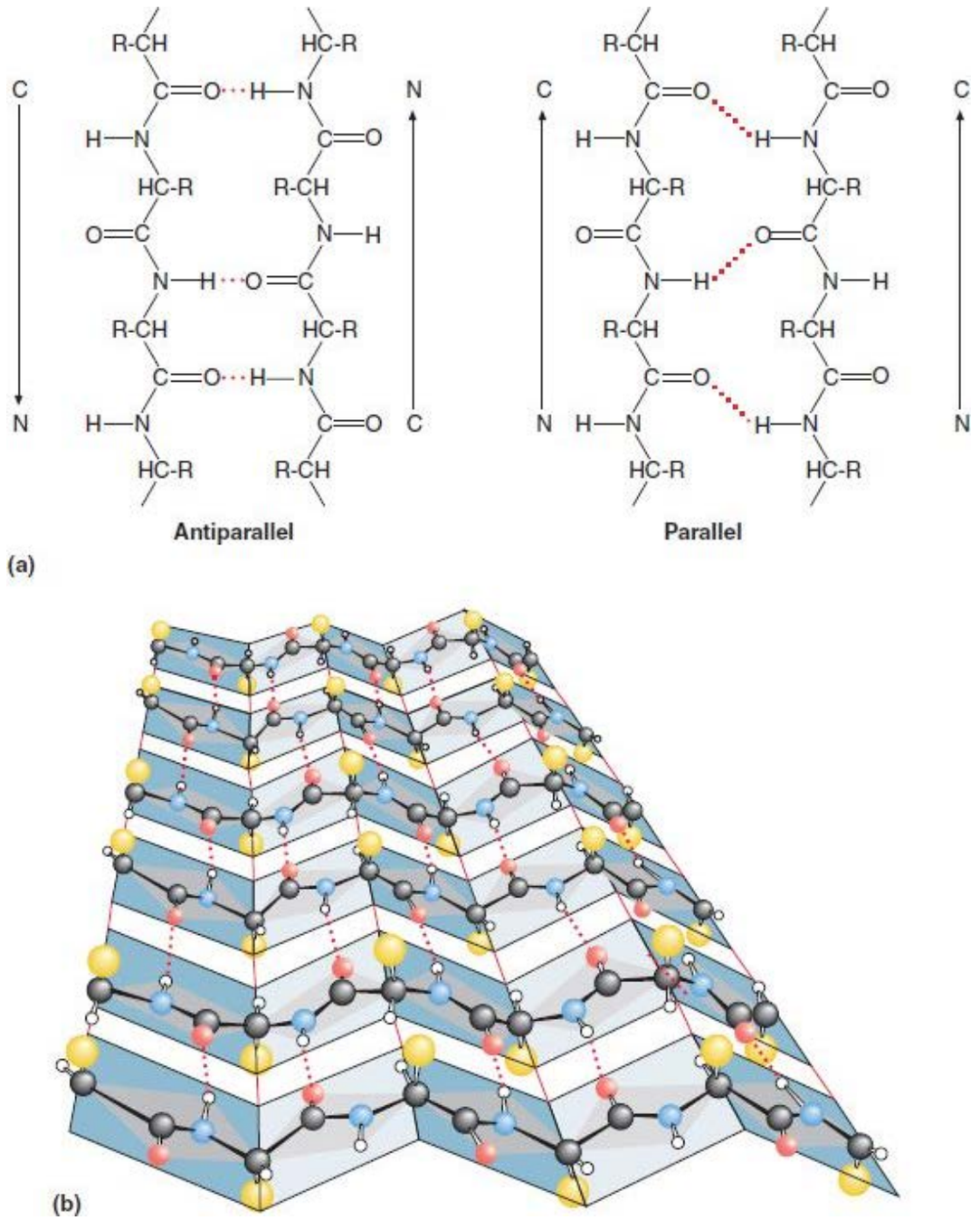


FIGURE 5.18

***β*-Pleated Sheet**

(a) Two forms of *β*-pleated sheet: antiparallel and parallel. Hydrogen bonds are represented by dotted lines. (b) A more detailed view of an antiparallel *β*-pleated sheet. Note that the side chains of adjacent residues point in opposite directions. The hydrogen bonds in antiparallel *β*-pleated sheets are perpendicular to the *β*-strands, and

those in parallel β-pleated sheets are evenly spaced but slanted.

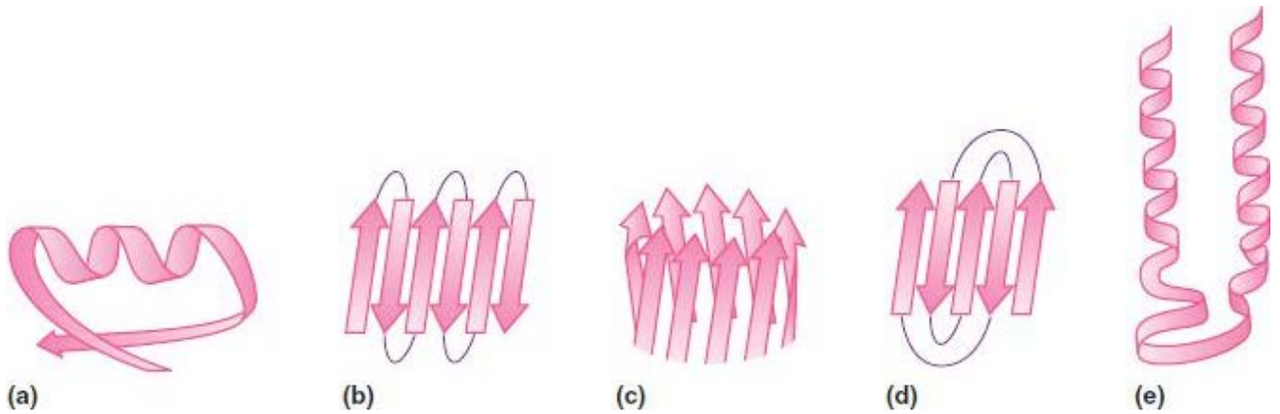


FIGURE 5.19

Selected Supersecondary Structures

(a) $\beta\alpha\beta$ units, (b) β-meander, (c) β-barrel, (d) Greek key, and (e) $\alpha\alpha$ unit. Note that β-strands are depicted as arrows. Arrow tips point toward the C-terminus.

Glycine and proline residues often occur in β-turns. Glycine's lack of an organic side group provides flexibility that permits a contiguous proline to assume a cis orientation (the same side of the peptide plane), allowing a tight turn to form in a polypeptide strand. Proline is a helix-breaking residue that alters the direction of the polypeptide chain because rotation around its α-carbon is not possible.

In the *β-meander* pattern, two or more consecutive antiparallel β-strands are connected by short loops with polar amino acids and glycines that permit a more abrupt change in direction involving a one- or two-residue *hairpin turn*. Hairpin turns allow flanking β-strands to interact. Several *β-barrel* arrangements are formed when β-sheets fold back on themselves. When an antiparallel β-sheet doubles back on itself in a pattern that resembles a common Greek pottery design, the motif is called the *Greek key*. In *$\alpha\alpha$ units* (or helix-loop-helix units), two α-helical regions separated by a nonhelical loop become aligned in a defined way because of interacting side chains. Helix-loop-helix motifs are often found in in DNA- and calcium-binding proteins.

TERTIARY STRUCTURE Although globular proteins often contain significant numbers of secondary structural elements, several other factors contribute to their structure. The term *tertiary structure* refers to the unique three-dimensional conformations that globular proteins assume as they fold into their native (biologically active) structures and prosthetic groups, if any, are inserted. **Protein folding**, a process in which an unorganized, nascent molecule acquires a highly organized structure, occurs as a consequence of the interactions between the molecule's side chains. Tertiary structure has several important features:

1. Many polypeptides fold in such a fashion that amino acid residues distant from each other in the primary structure come into close proximity.
2. Globular proteins are compact because of efficient packing as the polypeptide folds. During this process, most water molecules are excluded from the protein's interior, making interactions between both polar and nonpolar groups possible.
3. Large globular proteins (i.e., those with more than 200 amino acid residues) often contain several compact units called structural domains. Domains (**Figure 5.20**) are structurally independent segments that have specific functions (e.g., binding an ion or small molecule). The core three-dimensional structure of a domain is called a **fold**. Well-known examples of

folds include the nucleotide-binding Rossman fold, often found in nucleotide-binding proteins, and the globin fold, found in the oxygen-binding globins. Domains are classified on the basis of their core motif structure. Examples include α, β, α/β, and $\alpha + \beta$. α-Domains are composed exclusively of α-helices, and β-domains consist of antiparallel β-**strands.** α/β-Domains contain various combinations of an α-helix alternating with β-strands ($\beta\alpha\beta$ motifs). $\alpha + \beta$-Domains are primarily β-sheets with one or more outlying α-helices. Many globular proteins contain two or more domains.

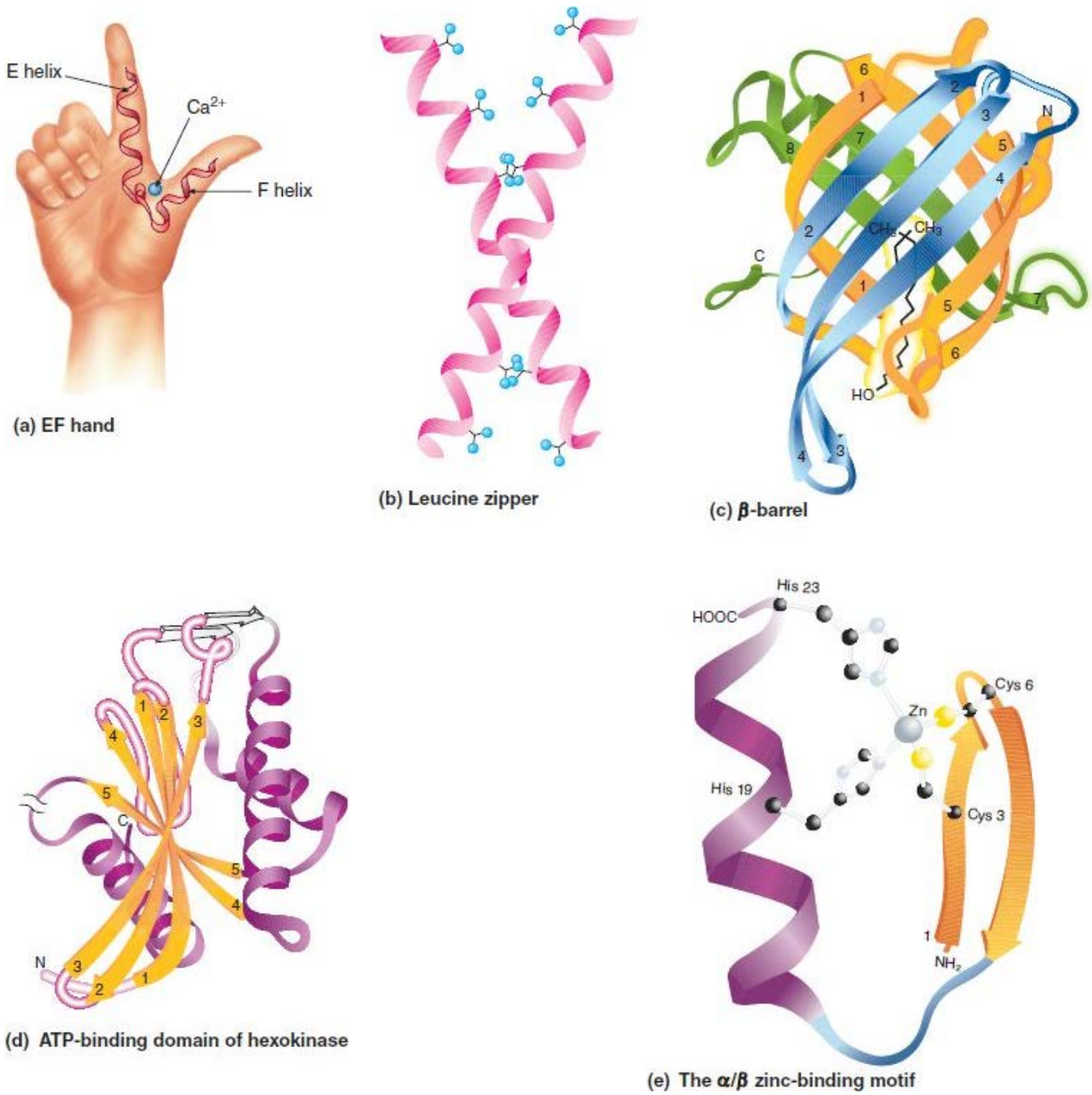


FIGURE 5.20

Selected Domains Found in Large Numbers of Proteins

(a) The EF hand, a helix-loop-helix that binds specifically to Ca^{2+}, and (b) the leucine zipper, a DNA-binding domain, are two examples of α-domains. (c) Human retinol-binding protein, a type of β-barrel domain (retinol, a visual pigment molecule is shown in yellow). (d) The ATP-binding domain of hexokinase, a type of α/β-domain. (e) The α/β zinc-binding motif, a core feature of numerous DNA-binding domains.

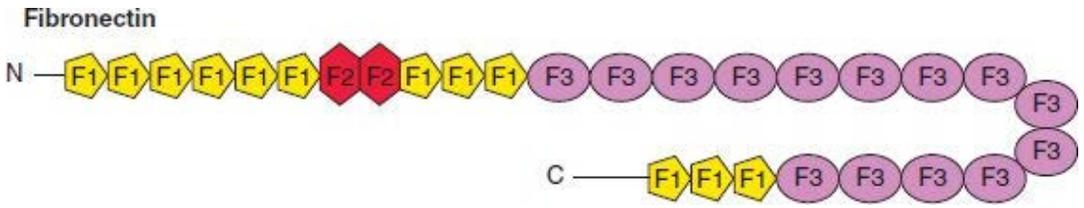


FIGURE 5.21

Fibronectin Structure

Fibronectin is a mosaic protein that is composed of multiple copies of F1, F2, and F3 modules. Fibronectin, a glycoprotein in the extracellular matrix, binds to components such as collagen and proteoglycans (p. 266) and links them to membrane-spanning proteins called integrins (p. 269).

4. A number of eukaryotic proteins, referred to as **modular** or **mosaic proteins**, contain numerous duplicate or imperfect copies of one or more domains that are linked in series. Fibronectin **(Figure 5.21)** contains three repeating domains: Fl, F2, and F3. All three domains, which are also found in a variety of extracellular matrix (ECM) proteins, contain binding sites for other ECM molecules such as collagen (p. 174) and heparan sulfate (p. 266), as well as certain cell-surface receptors. Domain modules are coded for by genetic sequences created by gene duplications (extra gene copies that arise from errors in DNA replication). Such sequences are used by living organisms to construct new proteins. For example, the immunoglobulin structural domain is found not only in antibodies, but also in a variety of cell-surface proteins.

Several types of interactions stabilize tertiary structure (**Figure 5.22**):

1. **Hydrophobic interactions.** As a polypeptide folds, hydrophobic R groups are brought into close proximity because they are excluded from water. Then the highly ordered water molecules in solvation shells are released from the interior, increasing the disorder (entropy) of the water molecules. The favorable entropy change is a major driving force in protein folding. It should be noted that a few water molecules remain within the core of folded proteins, where each forms as many as four hydrogen bonds with the polypeptide backbone. The stabilization contributed by small “structural” water molecules may free the polypeptide from some of its internal interactions. The resulting increased flexibility of the polypeptide chain is believed to play a critical role in the binding of molecules called **ligands** to specific sites. Ligand-binding pockets are water-depleted regions of the protein. Ligand binding is an important protein function.
2. **Electrostatic interactions.** The strongest electrostatic interaction in proteins occurs between ionic groups of opposite charge. Referred to as **salt bridges**, these noncovalent bonds are significant only in regions of the protein where water is excluded because of the energy required to remove water molecules from ionic groups. Salt bridges have been observed to contribute to the interactions between adjacent subunits in complex proteins. The same is true for the weaker electrostatic interactions (ion–dipole, dipole–dipole, van der Waals). They are significant in the interior of the folded protein and between subunits or in protein-ligand interactions. (In proteins that consist of more than one polypeptide chain, each polypeptide is called a subunit.)
3. **Hydrogen bonds.** A significant number of hydrogen bonds form within a protein’s interior and on its surface. In addition to forming hydrogen bonds with one another, the polar amino acid side chains may interact with water or with the polypeptide backbone. Again, the presence of water precludes the formation of hydrogen bonds with other species.
4. **Covalent bonds.** Covalent linkages are created by chemical reactions that alter a

polypeptide's structure during or after its synthesis. (Examples of these reactions, referred to as posttranslational modifications, are described in Section 19.2.) The most prominent covalent bonds in tertiary structure are the disulfide bridges found in many extracellular proteins. In extracellular environments, these strong linkages partly protect protein structure from adverse changes in pH or salt concentrations. Intracellular proteins do not contain disulfide bridges because of high cytoplasmic concentrations of reducing agents such as GSH.

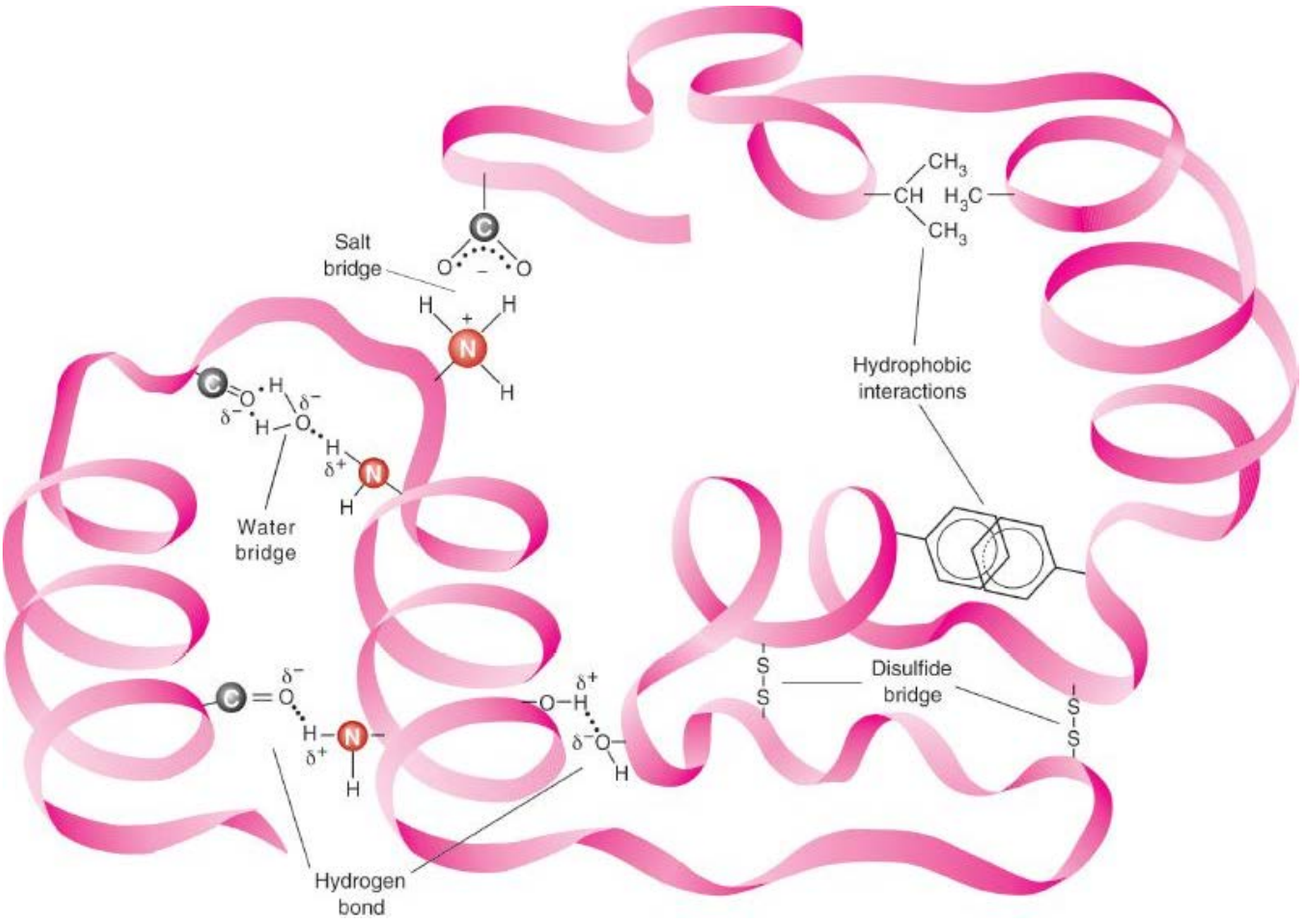


FIGURE 5.22

Interactions That Maintain Tertiary Structure

5. **Hydration.** As described previously (p. 86), structured water is an important stabilizing feature of protein structure. The dynamic hydration shell that forms around a protein (**Figure 5.23**) also contributes to the flexibility required for biological activity.

The precise nature of the forces that promote the folding of proteins (described on pp. 168–72) has not been completely resolved. It is clear, however, that protein folding is a thermodynamically favorable process with an overall negative free energy change. According to the free energy equation

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$$

a negative free energy change in a process is the result of a balance between favorable and unfavorable enthalpy and entropy changes (pp. 119–20). As a polypeptide folds, favorable (negative) ΔH values are the result in part of the sequestration of hydrophobic side chains within the interior of the molecule and the optimization of other noncovalent interactions. Opposing these factors is the unfavorable decrease in entropy that occurs as the disorganized polypeptide folds into its highly

organized native state. The change in entropy of the water that surrounds the protein is positive because of the decreased organization of the water in going from the unfolded to the folded state of the protein. For most polypeptide molecules, the net free energy change between the folded and unfolded state is relatively modest (the energy equivalent of several hydrogen bonds). The precarious balance between favorable and unfavorable forces allows proteins the flexibility they require for biological function.

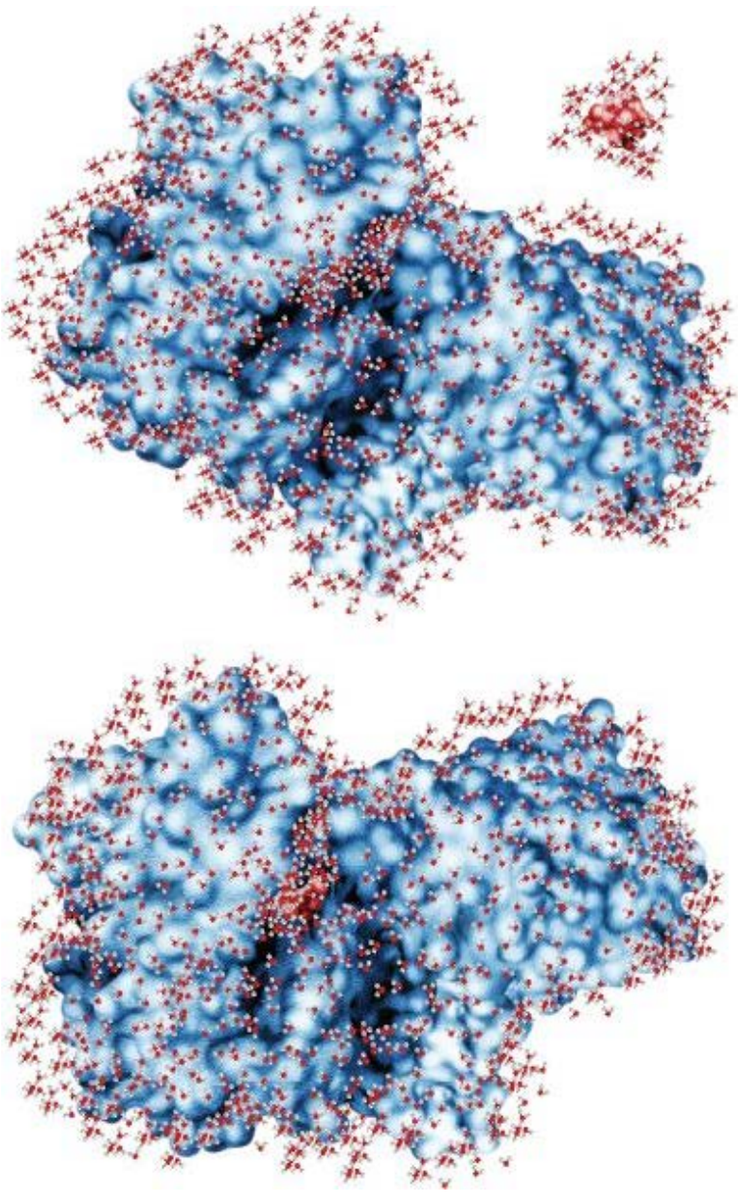


FIGURE 5.23

Hydration of a Protein

Three layers of structured water molecules surround a space-filling model of the enzyme hexokinase before and after binding the sugar glucose. Hexokinase (p. 286) is an enzyme that catalyzes the nucleophilic attack of the carbon–6 hydroxyl group of glucose on the phosphorus in the terminal phosphate of ATP. As the hydrated glucose molecule enters its binding site in a cleft in the enzyme, it displaces water molecules occupying the binding site. The water exclusion process promotes the conformation change that moves the domains together to create the catalytic site. Water exclusion from this site also prevents the unproductive hydrolysis of ATP.

QUATERNARY STRUCTURE Many proteins, especially those with high molecular weights, are composed of several polypeptide chains. Each polypeptide component in these molecular complexes is called a **subunit**. Protein subunits in a multisubunit complex may be identical or quite different.

Multisubunit proteins in which individual polypeptides are different or identical are referred to as **oligomers**. Oligomers are composed of **protomers**, which may consist of one or more subunits. A large number of oligomeric proteins contain two or four subunit protomers, referred to as dimers and tetramers, respectively. There appear to be several reasons for the common occurrence of multisubunit proteins:

1. Synthesis of separate subunits may be more efficient than substantially increasing the length of a single polypeptide chain.
2. In supramolecular complexes such as collagen fibers, replacement of smaller worn-out or damaged components can be managed more effectively.
3. The complex interactions among multiple subunits help regulate a protein's biological function.

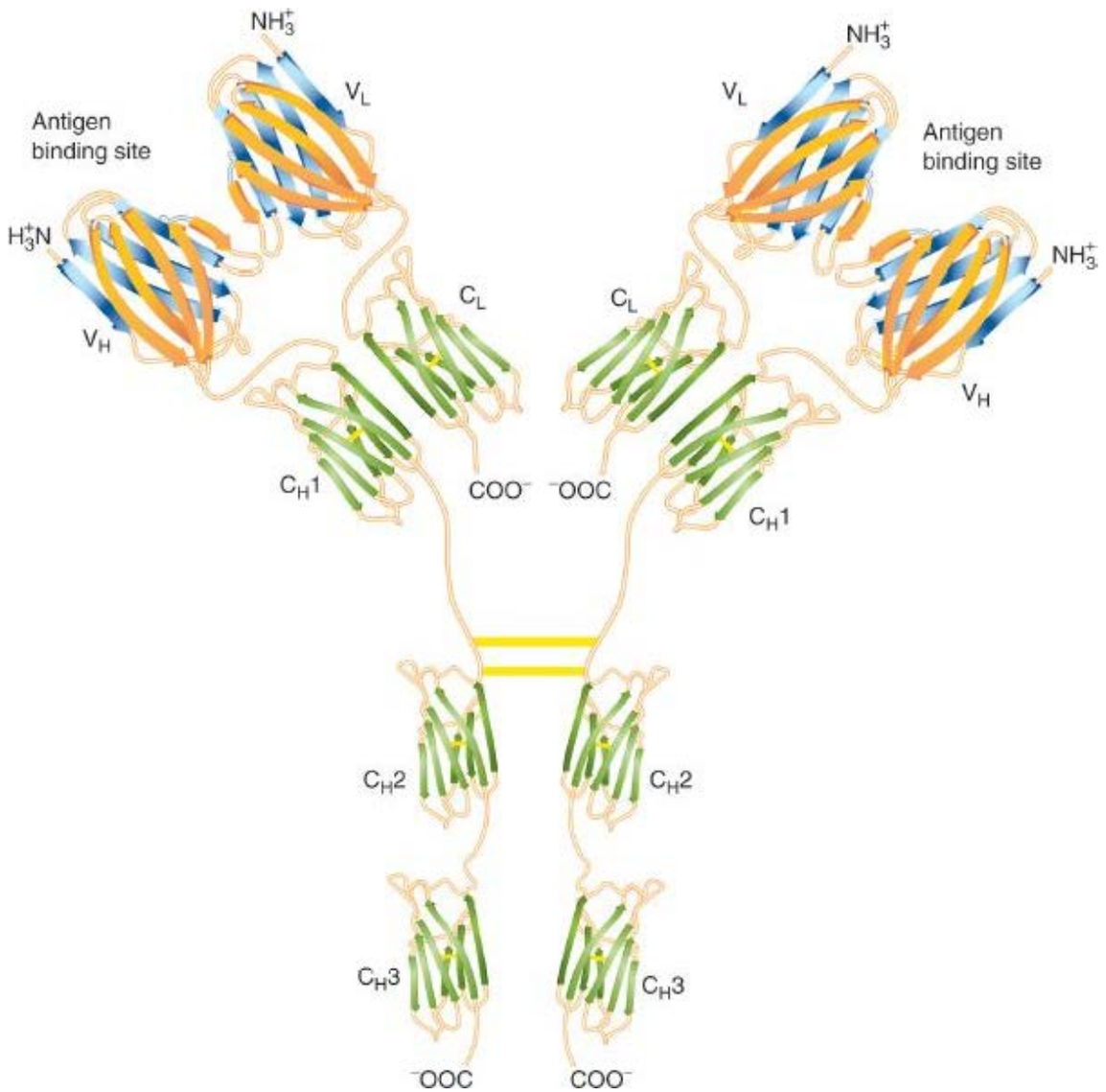


FIGURE 5.24

Structure of Immunoglobulin G (IgG)

IgG is an antibody molecule composed of two heavy chains (H) and two light chains (L) that together form a Y-shaped molecule. Each of the heavy and light chains contains constant (C) and variable (V) β-barrel domains (the classic immunoglobulin fold). The chains are held together by disulfide bridges (yellow lines) and noncovalent interactions. The variable domains of the H and L chains form the site that binds to antigens (foreign molecules). Many antigenic proteins bind to the external surface of these sites. Note that disulfide bridges are also a structural

feature within each constant domain.

Polypeptide subunits assemble and are held together by noncovalent interactions such as hydrophobic and electrostatic interactions and hydrogen bonds, as well as covalent cross-links. As with protein folding, the hydrophobic effect is clearly the most important because the structures of the complementary interfacing surfaces between subunits are similar to those observed in the interior of globular protein domains. Covalent cross-links significantly stabilize certain multisubunit proteins. Prominent examples include the disulfide bridges in the immunoglobulins (**Figure 5.24**) and the desmosine and lysinonorleucine linkages in certain connective tissue proteins. *Desmosine* (**Figure 5.25**) cross-links connect four polypeptide chains in the rubber-like connective tissue protein elastin. They are formed as a result of a series of reactions involving the oxidation and condensation of lysine side chains. A similar process results in the formation of *lysinonorleucine*, a cross-linking structure that is found in elastin and collagen fibers.

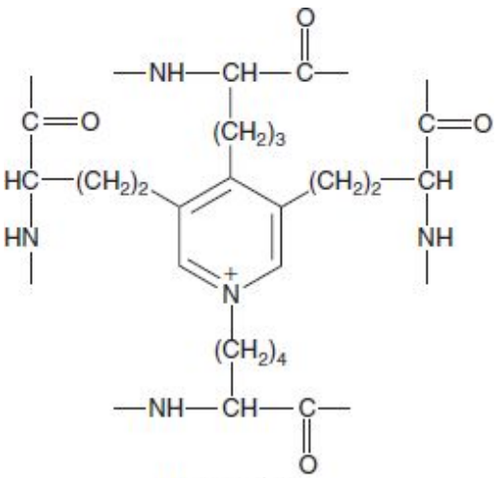


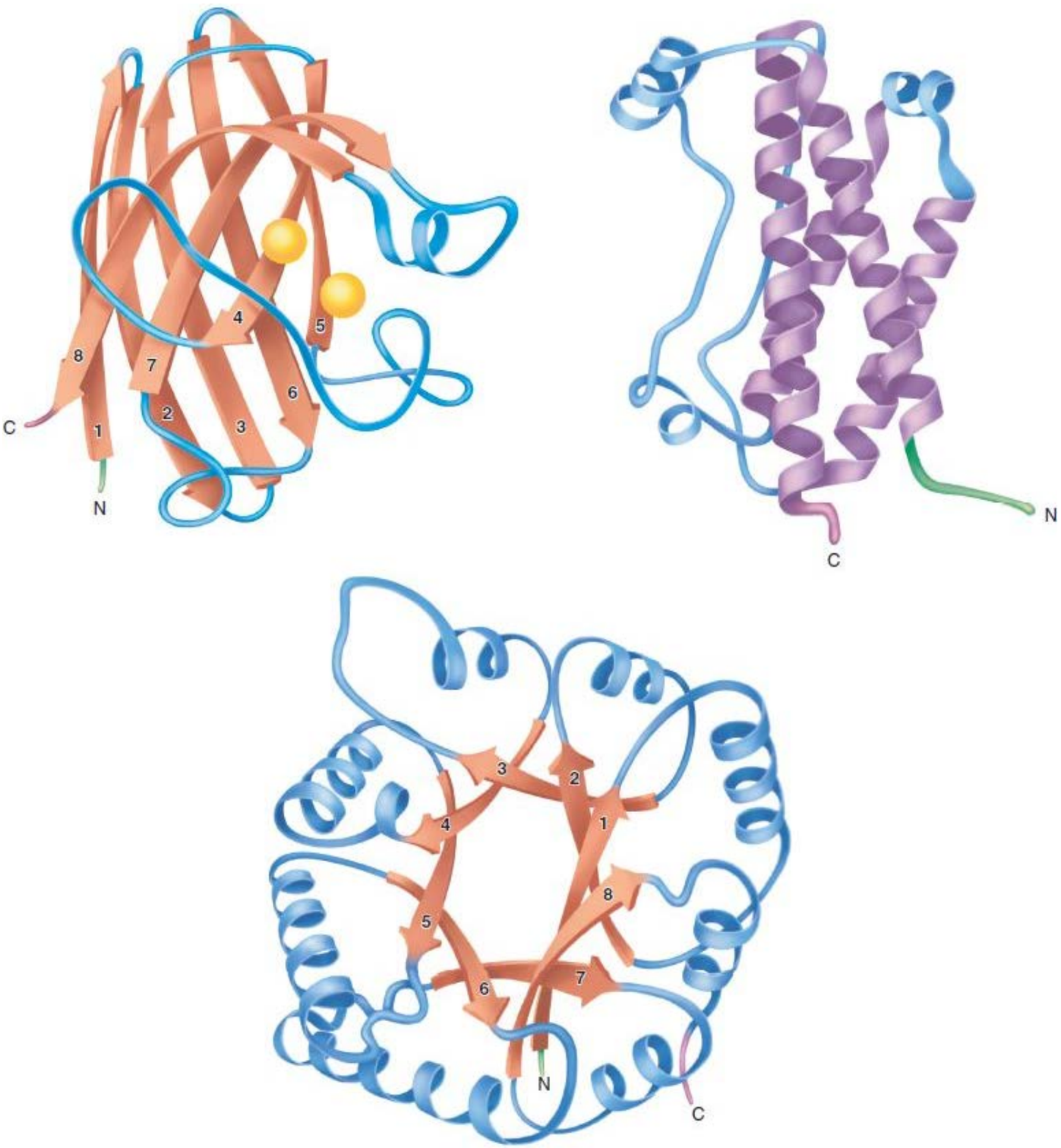
FIGURE 5.25

Desmosine and Lysinonorleucine Linkages

Quite often the interactions between subunits are affected by the binding of ligands. In **allostery**, which is the control of protein function through ligand binding, binding a ligand to a specific site in a protein triggers a conformational change that alters its affinity for other ligands. Ligand-induced conformational changes in such proteins are called **allosteric transitions**, and the ligands that trigger them are called **effectors** or **modulators**. Allosteric effects can be positive or negative, depending on whether effector binding increases or decreases the protein's affinity for other ligands. One of the best-understood examples of allosteric effects, the reversible binding of O_2 and other ligands to hemoglobin, is described on pp. 178–80. (Because allosteric enzymes play a key role in the control of metabolic processes, allostery is discussed further in Sections 6.3 and 6.5.)

QUESTION 5.8

Review the following illustrations of globular proteins. Identify examples of secondary and supersecondary structure.



QUESTION 5.9

Illustrate the noncovalent interactions that can occur between the following side chain groups in

folded polypeptides: (a) serine and glutamate, (b) arginine and aspartate, (c) threonine and serine, (d) glutamine and aspartate, and (e) phenylalanine and tryptophan.

UNSTRUCTURED PROTEINS Polypeptides fall on a structural continuum from highly folded containing multiple domains linked by flexible linkers, those with varying percentages of disordered segments, and finally those that are completely disordered. The term **intrinsically unstructured** (or **disordered**) **proteins** (IDPs) is used to describe molecules that are completely unstructured. Unstructured segments within an otherwise structured molecule are called *intrinsically disordered regions* (*IDRs*). The folding of IDPs into stable three-dimensional conformations is prevented by biased amino acid sequences that contain high percentages of polar and charged amino acids (e.g., Ser, Gln, Lys, and Glu) and low quantities of hydrophobic amino acids (e.g., Leu, Val, Phe, and Trp).

Unstructured proteins have a diversity of functions. IDPs and IDRs play a central role in cellular signaling and regulatory processes because, in addition to providing a larger surface area than globular proteins and accessible sequences for covalent modification (advantages in signal transduction), IDPs and IDRs have multiple short-sequence motifs (3 to 7 amino acid residues) that permit interaction with numerous other proteins. Highly extended and malleable disordered segments enable the molecule to "search" for binding partners in a manner similar to that of a hook on the end of a fly fisherman's fishing rod. cAMP response element binding protein (CREB) and p53 are examples of proteins whose functions depend on disordered segments.

CREB, a transcription regulatory protein discussed later (p. 607), binds to cAMP response elements (CREs), one type of DNA sequence called a *response element*. When the KID (*k*inase *i*nducible *d*omain) of CREB is phosphorylated by a kinase (an enzyme that attaches phosphate groups to specific amino acid side chains), it becomes completely unstructured. The unstructured phosphorylated KID (pKID) domain is then able to search out and bind to a domain of CREB-binding protein (CBP) called KIX (KID-binding domain) (**Figure 5.26**). As often happens with IDPs, the disordered pKID domain transitions into a more ordered conformation as it binds to the KIX domain of CBP. As a result of CREB-CBP binding, CREB forms a dimer that alters the expression of certain genes when it binds to its response element.

p53 (**Figure 5.27**) provides a remarkable example of the utility of unstructured protein domains. p53 is a major tumor suppressor (p. 743), as indicated by the fact that p53 mutations occur in at least 50% of all cancers. Active as a homotetramer, p53 regulates the expression of hundreds of genes, most notably those involved in cancer-suppression processes such as DNA repair, cell cycle arrest, autophagy (p. 574), apoptosis (p. 58), and energy metabolism. p53 is activated in response to a variety of stimuli such as DNA damage, ER stress, ultraviolet (UV) light, and hypoxia (oxygen deficiency). Each p53 polypeptide possesses four domains: an N-terminal disordered transactivation domain (a peptide sequence that recruits and binds to numerous transcription factors); a DNA-binding domain; a tetramerization domain; and a C-terminal disordered domain that is involved in nuclear localization. p53 can integrate information from multiple signaling pathways because covalent modifications (e.g., phosphorylation and acetylation) of its disordered N-terminal domain result in structural variations that allow interactions with a wide variety of proteins.

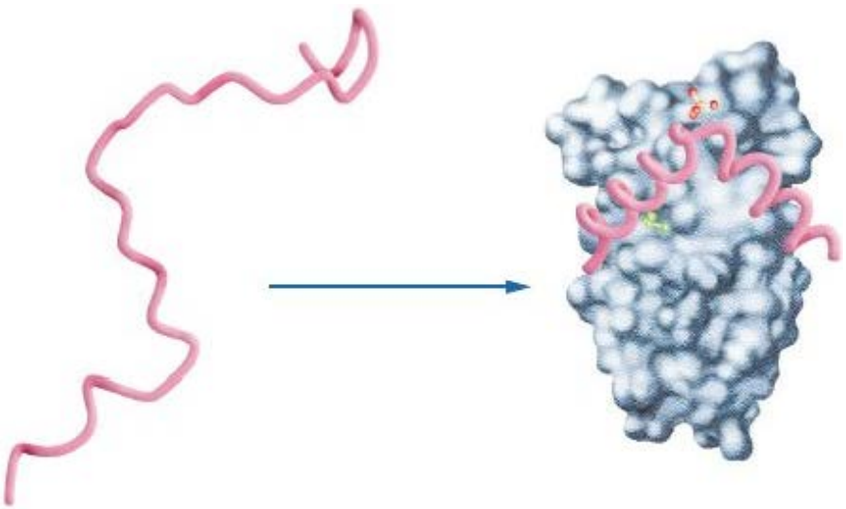


FIGURE 5.26

Disordered Protein Binding

The intrinsically disordered phosphorylated KID domain (pKID) (left) of the transcription regulatory protein CREB searches out and binds to the KIX domain of the transcription coactivator protein CBP (right). As pKID binds to KIX, it undergoes a disorder-to-order transition as it folds into a pair of helices.

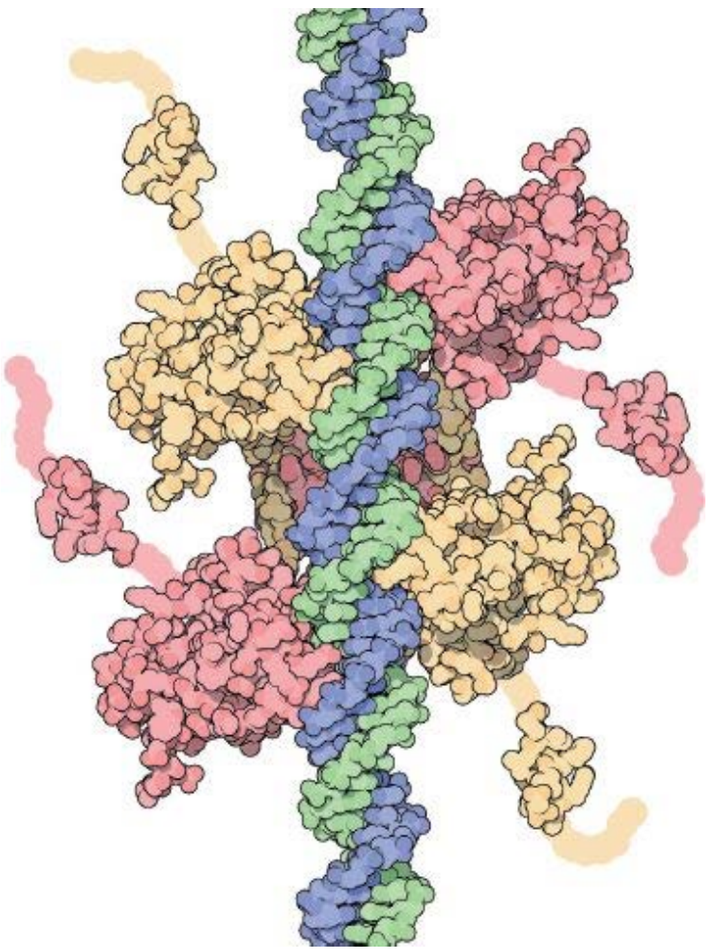


FIGURE 5.27

p53 Structure

p53 is a key regulatory tumor suppressor transcription factor that functions as a tetramer. The tetramer is formed by the interactions (behind the DNA helix, which is illustrated in blue and green) of the four tetramerization domains. The four DNA-binding domains are linked to the tetramerization domains by disordered segments (also

hidden). The four N-terminal transactivation domains with their disordered segments extend outward in search of partner proteins (transcription factors and other signaling proteins).

KEY CONCEPTS

- Biochemists distinguish four levels of the structural organization of proteins.
- In primary structure, the amino acid residues are connected by peptide bonds.
- The secondary structure of polypeptides is stabilized by hydrogen bonds. Prominent examples of secondary structure are α-helices and β-pleated sheets.
- Tertiary structure is the unique three-dimensional conformation that a protein assumes because of the interactions between amino acid side chains. Several types of interaction stabilize tertiary structure: hydrophobic and electrostatic interactions, hydrogen bonds, and certain covalent bonds.
- Proteins that consist of several separate polypeptide subunits exhibit quaternary structure. Both noncovalent and covalent bonds hold the subunits together.
- Some proteins are partially or completely unstructured.

Proteoforms and Proteome Complexity. The proteomes (p. 637) of the cells in multicellular eukaryotes can be quite diverse. For example, as a result of protein research efforts, the original estimate of human protein coding genes was 100,000, instead of the approximately 20,000 that were ultimately confirmed. This discrepancy is now explained by three factors: (1) alternative splicing (a regulated process in which specific sequences are excluded from mRNAs, pp. 736–38), (2) elimination of N-terminal segments as a result of alternative translation initiation, and (3) posttranslational modifications (enzyme-catalyzed alterations of polypeptides that occur after synthesis, pp. 776–79). As a result of these processes, several protein products, referred to collectively as *proteoforms*, can be produced from a single-polypeptide coding gene.

LOSS OF PROTEIN STRUCTURE When one considers the small differences in the free energy of folded and their unfolded conformations, it is not surprising to learn that protein structure is especially sensitive to environmental factors. Many physical and chemical agents can disrupt a protein's native conformation. The process of structure disruption, which may or may not involve protein unfolding, is called **denaturation**. (Denaturation is not usually considered to include the breaking of peptide bonds.) Depending on the degree of denaturation, the molecule may partially or completely lose its biological activity. Denaturation often results in easily observable changes in the physical properties of proteins. For example, soluble and translucent egg albumin (egg white) becomes insoluble and opaque on heating. Like many denaturations, cooking eggs is an irreversible process.

Christian Anfinsen, who shared the Nobel Prize in Chemistry in 1972, demonstrated the following example of a reversible denaturation in the 1950s. Bovine pancreatic ribonuclease (a digestive enzyme from cattle that degrades RNA) is denatured when treated with β-mercaptoethanol and 8 M urea (**Figure 5.28**). During this process, ribonuclease, composed of a single polypeptide with four disulfide bridges, completely unfolds and loses all biological activity. Careful removal of the denaturing agents with dialysis (p. 184) results in a spontaneous and correct refolding of the polypeptide and re-formation of the disulfide bonds. Anfinsen's experimental treatment, which resulted in full restoration of the enzyme's catalytic activity, provided an important early insight into the roles of different forces and primary structure in protein folding. However, most proteins treated similarly do not renature.

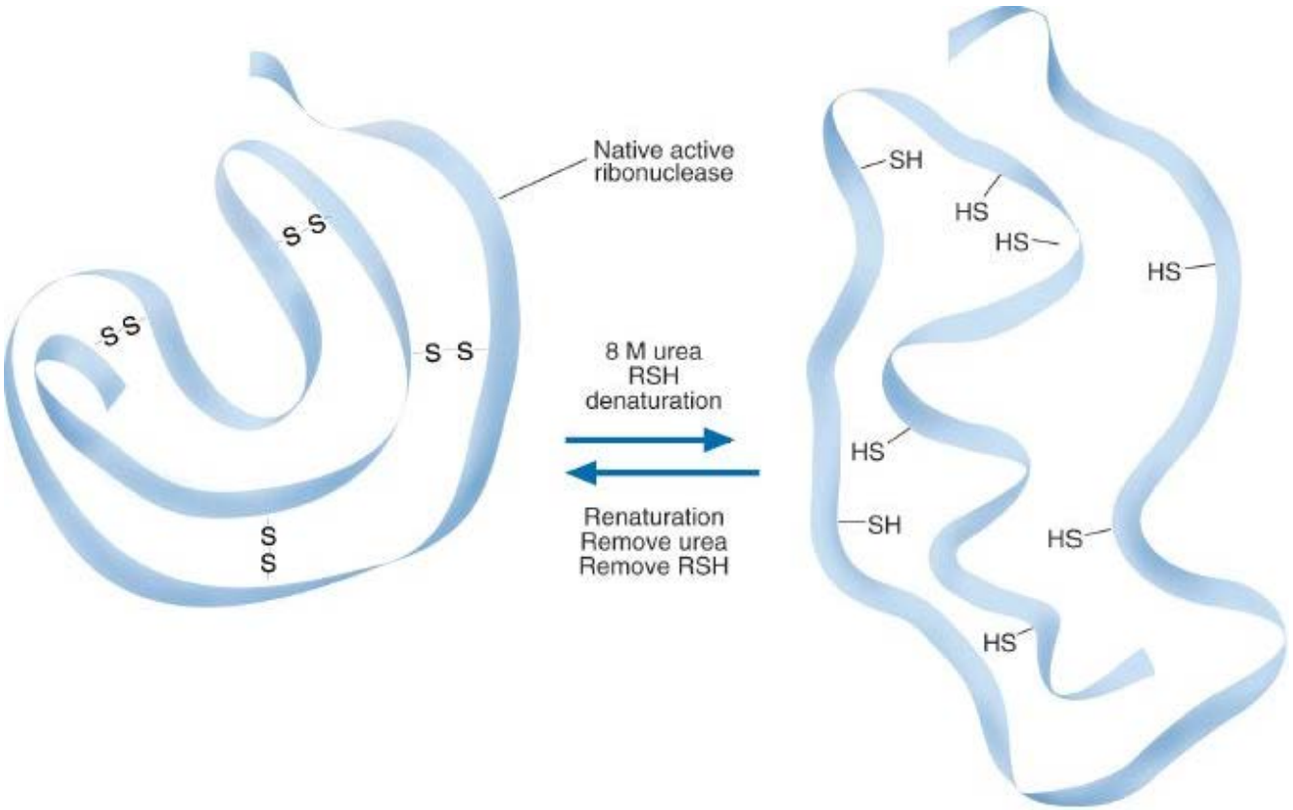


FIGURE 5.28

The Anfinsen Experiment

Ribonuclease denatured by 8 M urea and a mercaptan (RSH, a reagent that reduces disulfides to sulfhydryl groups) can be renaturated by removing the urea and RSH and air-oxidizing the reduced disulfides.

Denaturing conditions include the following:

1. **Strong acids or bases.** Changes in pH alter the protonation state of certain protein side chain groups, which in turn alters hydrogen bonding and salt bridge patterns. As a protein approaches its isoelectric point, it becomes less soluble and may precipitate from solution.
2. **Organic solvents.** Water-soluble organic solvents such as ethanol interfere with hydrophobic interactions because they interact with nonpolar R groups and form hydrogen bonds with water and polar protein groups. Nonpolar solvents also disrupt hydrophobic interactions.
3. **Detergents.** Detergents are substances that disrupt hydrophobic interactions, causing proteins to unfold into extended polypeptide chains. These molecules are called **amphipathic** because they contain both hydrophobic and hydrophilic components.
4. **Reducing agents.** In the presence of reagents such as urea, reducing agents (e.g., β-mercaptoethanol) convert disulfide bridges to sulfhydryl groups. Urea disrupts hydrogen bonds and hydrophobic interactions.
5. **Salt concentration.** When there is an increase in the salt concentration of an aqueous solution of protein, some of the water molecules that interact with the protein's ionizable groups are attracted to the salt ions. As the number of solvent molecules available to interact with these groups decreases, protein–protein interactions increase. If the salt concentration is high enough, there are so few water molecules available to interact with ionizable groups that the solvation spheres surrounding the protein's ionized groups are removed. The protein molecules aggregate and then precipitate. This process is referred to as *salting out*. Because salting out is usually reversible and different proteins salt out at different salt concentrations,

it is often used as an early step in protein purification.

6. **Heavy metal ions.** Heavy metals such as mercury (Hg^{2+}) and lead (Pb^{2+}) affect protein structure in several ways. They may disrupt salt bridges by forming ionic bonds with negatively charged groups. Heavy metals also bond with sulfhydryl groups, which may result in significant changes in protein structure and function. For example, Pb^{2+} binds to catalytically active sulfhydryl groups in two enzymes in the heme synthetic pathway. The resultant decrease in hemoglobin synthesis causes severe anemia. (In anemia, the number of red blood cells or the hemoglobin concentration is lower than normal.) Anemia is one of the most easily measured symptoms of lead poisoning.
7. **Temperature changes.** As the temperature increases, the rate of molecular vibration increases. Eventually, weak interactions such as hydrogen bonds are disrupted and the protein unfolds. Some proteins are more resistant to heat denaturation, and this fact can be used in purification procedures.
8. **Mechanical stress.** Stirring and grinding actions disrupt the delicate balance of forces that maintain protein structure. For example, the foam formed when egg white is beaten vigorously contains denatured protein.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essays on lead poisoning and heme biosynthesis (Chapter 14).

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay on protein folding and human health.

The Folding Problem

The direct relationship between an ordered protein's primary sequence and its final three-dimensional conformation, and by extension its biological activity, is among the most important assumptions of modern biochemistry. One of the principal underpinnings of this paradigm has already been mentioned: the series of experiments reported by Christian Anfinsen. Anfinsen's discovery suggested that the three-dimensional structure of any protein could be predicted if the physical and chemical properties of the amino acids and the forces that drive the folding process (e.g., bond rotations, free energy considerations, and the behavior of amino acids in aqueous environments) were understood. Although this problem is not completely resolved despite several decades of painstaking research, protein biochemists, using the most sophisticated tools available (e.g., X-ray crystallography and nuclear magnetic resonance [NMR] in combination with site-directed mutagenesis and computer-based mathematical modeling), have made significant progress. Their work has revealed that protein folding is a stepwise process in which secondary structure formation (i.e., α-helix and β-pleated sheet) is an early feature and that hydrophobic interactions are an important force in folding. In addition, amino acid substitutions experimentally introduced into proteins (via site-directed mutagenesis, a technique that creates specific changes in a DNA sequence) reveal that changes in surface amino acids rarely affect the protein's structure. In contrast, substitutions of amino acids within the hydrophobic core often lead to serious structural changes in conformation.

In recent years, protein-folding researchers have determined that the process does not consist, as

was originally thought, of a single pathway. Instead, a polypeptide can take numerous routes to fold into its native state. As illustrated in **Figure 5.29a**, an energy landscape with a funnel shape appears to best describe how an unfolded polypeptide with its own unique set of constraints (its amino acid sequence, posttranslational modifications, and environmental features within the cell such as temperature, pH, and molecular crowding) negotiates its way to a low-energy folded state. Depending largely on its size, a polypeptide may or may not form intermediates (species existing long enough to be detected) that are momentarily trapped in local energy wells (**Figure 5.29b**). Small molecules (fewer than 100 residues) often fold without intermediate formation (**Figure 5.30a**). As these molecules begin emerging from the ribosome, a rapid and cooperative folding process begins in which side chain interactions facilitate the formation and alignment of secondary structures. The folding of larger polypeptides typically involves the formation of several intermediates (**Figure 5.30b, c**). In many of these molecules or the domains within a molecule, the hydrophobically collapsed shape of the intermediate is referred to as a molten globule. A **molten globule** is a partially organized globular state of a folding polypeptide that resembles the molecule's native state. Within the interior of a molten globule, tertiary interactions among amino acid side chains are fluctuating; that is, they have not yet stabilized.

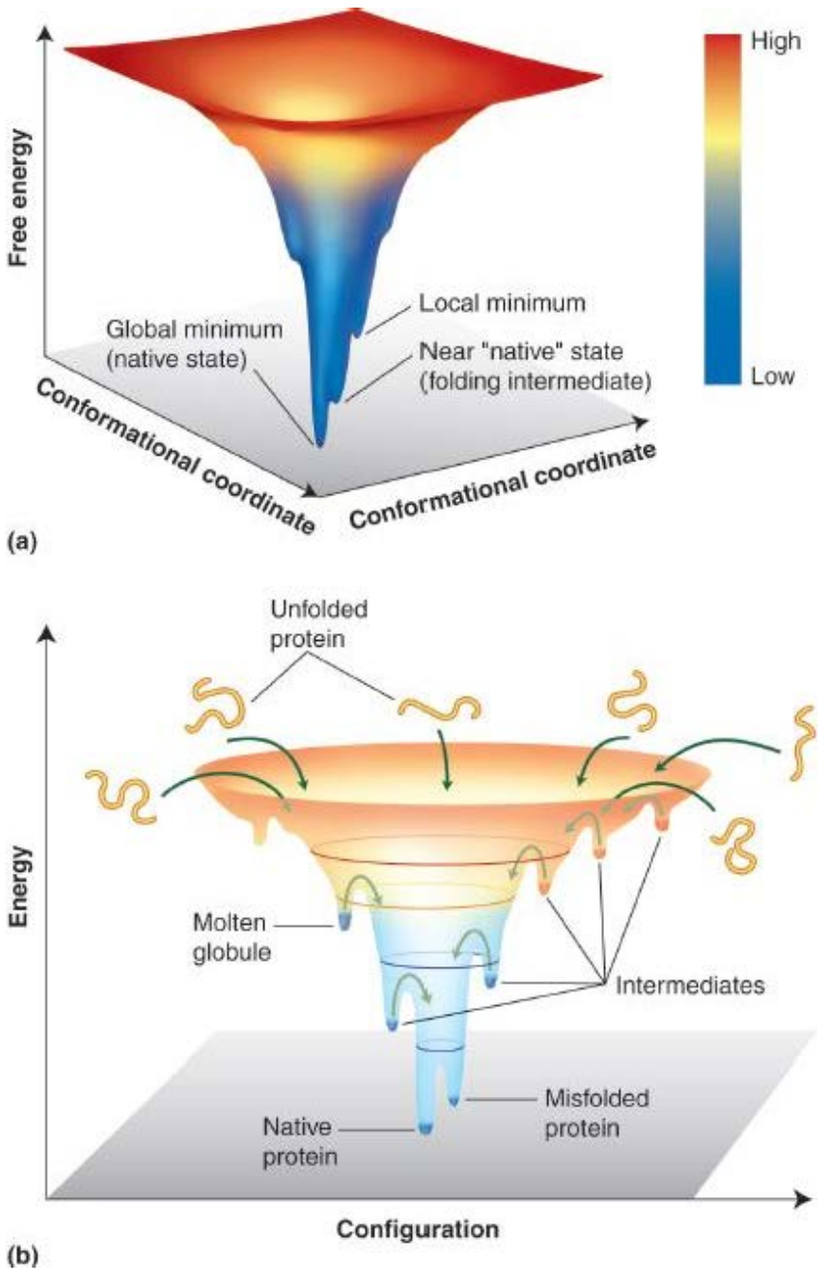


FIGURE 5.29

The Energy Landscape for Protein Folding

(a) Color is used to indicate the entropy level of the folding polypeptide. As folding progresses, the polypeptide moves from a disordered state (high entropy, red) toward a progressively more ordered conformation until its unique biologically active conformation is achieved (lower entropy, blue). (b) A depiction of the conformational state of a polypeptide during folding: polypeptides can fold into their native states by several different pathways. Many molecules form transient intermediates, which may or may not result in a trapped misfolded state.

Many proteins require assistance in folding into their native conformations. The **molecular chaperones** are a network of proteins that play a central role in cellular protein quality control by facilitating the folding of *nascent* (newly synthesized) proteins and the refolding of preexisting proteins. In addition, they also assist in the assembly of multisubunit proteins and other protein-containing structures (e.g., chromatin) and, if necessary, target misfolded proteins to the cell's degradation pathways (pp. 572–75). Many molecular chaperones are hsps (p. 152). Found in organisms ranging from bacteria to animals and plants, the molecular chaperones have a high degree of sequence homology. In eukaryotes, they are found in cytoplasm and several organelles (mitochondria, ER, and chloroplasts). The properties of these important molecules are described next.

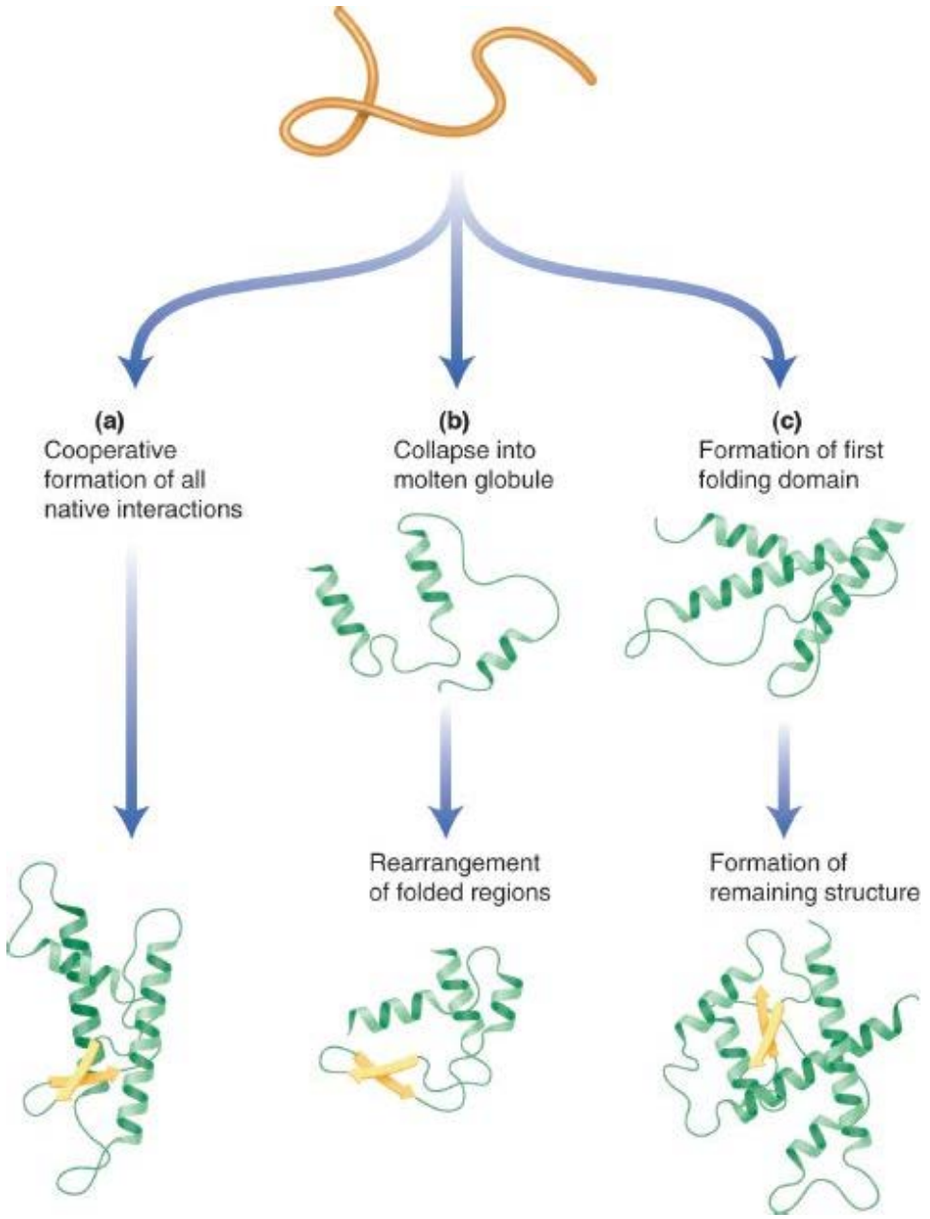


FIGURE 5.30

Protein Folding

(a) In many small proteins, folding is cooperative with no intermediates formed. (b) In some larger proteins, folding involves the initial formation of a molten globule followed by rearrangement into the native conformation. (c) Large proteins with multiple domains follow a more complex pathway, with each domain folding separately before the entire molecule progresses to its native conformation.

MOLECULAR CHAPERONES Molecular chaperones apparently assist unfolded proteins by protecting them from inappropriate hydrophobic protein–protein interactions that can result in misfolding or aggregation. There are four groups of molecular chaperones involved in de novo protein folding, that is, assistance in the folding of nascent polypeptides:

1. **Ribosome-associated chaperones**. Chaperone action begins as the nascent polypeptide begins to emerge from the ribosome's exit tunnel. Proteins such as *trigger factor* in bacteria (p. 766), RAC (ribosome-associated complex) in yeast, and NAC (*n*ascent polypeptide-*a*ssociated *c*omplex) in eukaryotes bind to the ribosome and the emerging polypeptide. This binding process prevents folding until an entire domain or polypeptide has emerged from the tunnel.
2. **Hsp70s**. The hsp70s bind to, stabilize, and promote the folding of nascent polypeptides. They

are also involved in the refolding of misfolded and aggregated proteins and in the transmembrane transfer of organelle (e.g., ER, mitochondria, and chloroplast) or secretory polypeptides. Each hsp70 possesses two domains connected by a short interdomain linker: an N-terminal ATP-binding domain and a domain containing a substrate-binding pocket with an affinity for peptides with hydrophobic amino acid residues. There is also a C-terminal α-helical structure that acts as a lid for the substrate-binding pocket. Hsp70 ATPase activity is stimulated when a peptide segment enriched in hydrophobic amino acid residues binds within the binding pocket. When ATP in the N-terminal domain of an hsp70 is hydrolyzed, the lid closes, thus trapping a peptide segment of the substrate protein in the binding pocket. Peptide segments are released from hsp70s as a result of a conformational change initiated by the exchange of ADP for ATP. Hsp70s interact with *co-chaperones*, accessory proteins that assist hsps in individual steps in chaperone activities. For example, hsp40 assists hsp70 by regulating ATP hydrolysis and mediating binding to unfolded or oxidized protein substrates. Hsp100 assists hsp70 in the disassembly of protein aggregates. On its release, a polypeptide either folds into its functional conformation or hsp70 and its associated co-chaperones transfer it to other downstream chaperones.

3. **Hsp90s**. The hsp90s, found in organisms ranging from bacteria to mammals (but not archaea), have roles in diverse cellular pathways. In eukaryotes, the hsp90s are found in cytoplasm, nucleus, mitochondria, and chloroplasts where they do not bind to nascent polypeptides. Instead, hsp90 proteins finalize the folding of a limited, but diverse, set of partially unfolded molecules, referred to as *client proteins* (e.g., protein kinases and steroid hormone receptors). Reversibly linked to hsp70, hsp90 completes the folding process. Hsp70 and hsp90 also work together to identify proteins damaged by oxidative or heat stress and either refold them or target them for proteasome-mediated destruction (p. 572). In addition, hsp90 also coordinates the assembly of protein complexes such as RNA polymerase II (p. 725), RNA-induced silencing complex (p. 738), and 26S proteasome (p. 573). Hsp90 functions as a dimer. Each hsp90 molecule is composed of three domains: the N-terminal domain, containing an ATP binding site; a middle domain that is involved in client protein binding and ATP hydrolysis regulation; and the C-terminal domain, which contains the interaction site for dimer formation. In an ATP-regulated cycle, an open hsp90 dimer (V-shaped) binds a client protein and closes in a clamp-like motion driven by ATP hydrolysis. Following ATP hydrolysis, the dimer opens, releasing the now folded client protein.

4. **Chaperonins**. The chaperonins are large, double-ring complexes that promote faster, more efficient polypeptide refolding within an internal compartment. They are classified into two groups. Group I chaperonins, found in bacteria, mitochondria, and chloroplasts, are composed of two stacked seven-membered rings. They are known as GroEL in bacteria and hsp60s in eukaryotes, and their function requires lid-shaped co-chaperones. Protein folding commences after lid closure entraps a substrate protein. ATP hydrolysis converts the cavity within GroEL into a hydrophilic microenvironment that facilitates the collapse of the hydrophobic core of the folding protein into the molten globule form. It takes 15 to 20 s for all seven ATPs to hydrolyze in the ring subunits and to complete the folding process. In the $(ADP)_7$ state, the hydrophobic character of the cavity returns, the chamber opens, and the folded protein or domain is released. A new, unfolded protein can now bind to repeat the cycle. Protein folding proceeds with two cycles occurring in an overlapping fashion, depending on the ATP/ADP binding status of the two cavities. The structure of GroEL and its lid-shaped co-chaperone GroES is illustrated in **Figure 5.31**. TRiC, a group II chaperonin in the cytoplasm of eukaryotic cells, is composed of two eight-member rings with a built-in lid. Along with its co-chaperone hsp70, TRiC facilitates the cotranslational folding of proteins with complex domain folds, most notably actin and tubulin.

In addition to promoting the folding of nascent protein, molecular chaperones direct the refolding of proteins that are partially unfolded as a consequence of stressful conditions. If refolding is not possible, molecular chaperones promote protein degradation. A diagrammatic view of protein folding involving GroEL and its lid co-chaperone, GroES, is presented in **Figure 5.32**. The effects of protein misfolding on human health can be considerable. Both Alzheimer's and Huntington's diseases are neurodegenerative diseases caused by accumulations of insoluble protein aggregates. (See the online Biochemistry in Perspective essay Protein Folding and Human Disease.)

KEY CONCEPTS

- All the information required for each newly synthesized polypeptide to fold into its biologically active conformation is encoded in the molecule's primary sequence.
- Some relatively simple polypeptides fold spontaneously into their native conformations.
- Other larger molecules require the assistance of proteins called molecular chaperones to ensure correct folding.

Alzheimer's and Huntington's Diseases

Fibrous Proteins

Fibrous proteins typically contain high proportions of regular secondary structures, such as α-helices or β-pleated sheets. As a consequence of their rod-like or sheet-like shapes, many fibrous proteins have structural rather than dynamic roles. Keratin (**Figure 5.33**) is a fibrous protein composed of bundles of α-helices, whereas the polypeptide chains of the silkworm silk protein fibroin (**Figure 5.34**) are arranged in antiparallel β-pleated sheets. The structural features of collagen, the most abundant protein in vertebrates, are described in some detail.

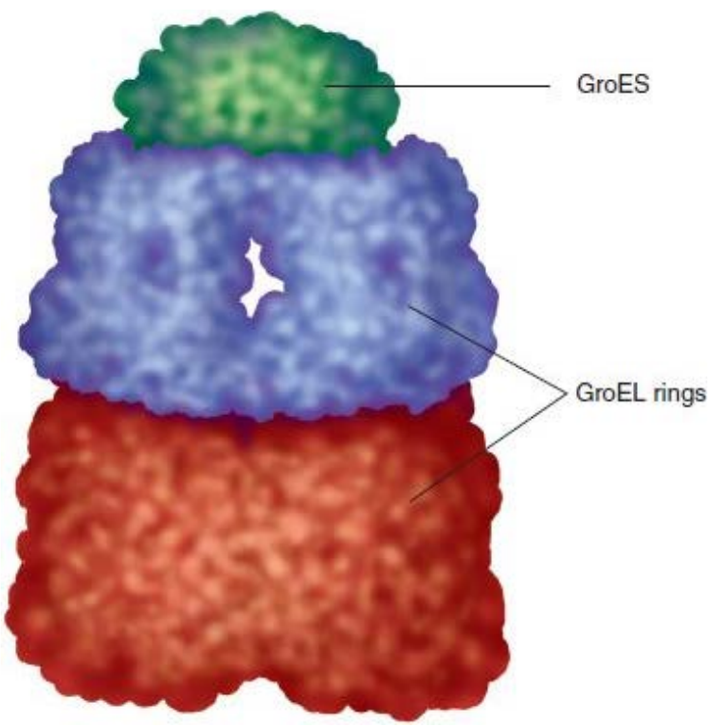


FIGURE 5.31

Space-Filling Model of the *E. coli* Chaperonin Called the GroES-GroEL Complex

GroES (*hsp10* in eukaryotes) is a seven-subunit ring that sits on top of GroEL.
GroEL (*hsp60* in eukaryotes) is composed of two stacked, seven-subunit rings with a cavity in which ATP-dependent protein folding takes place.

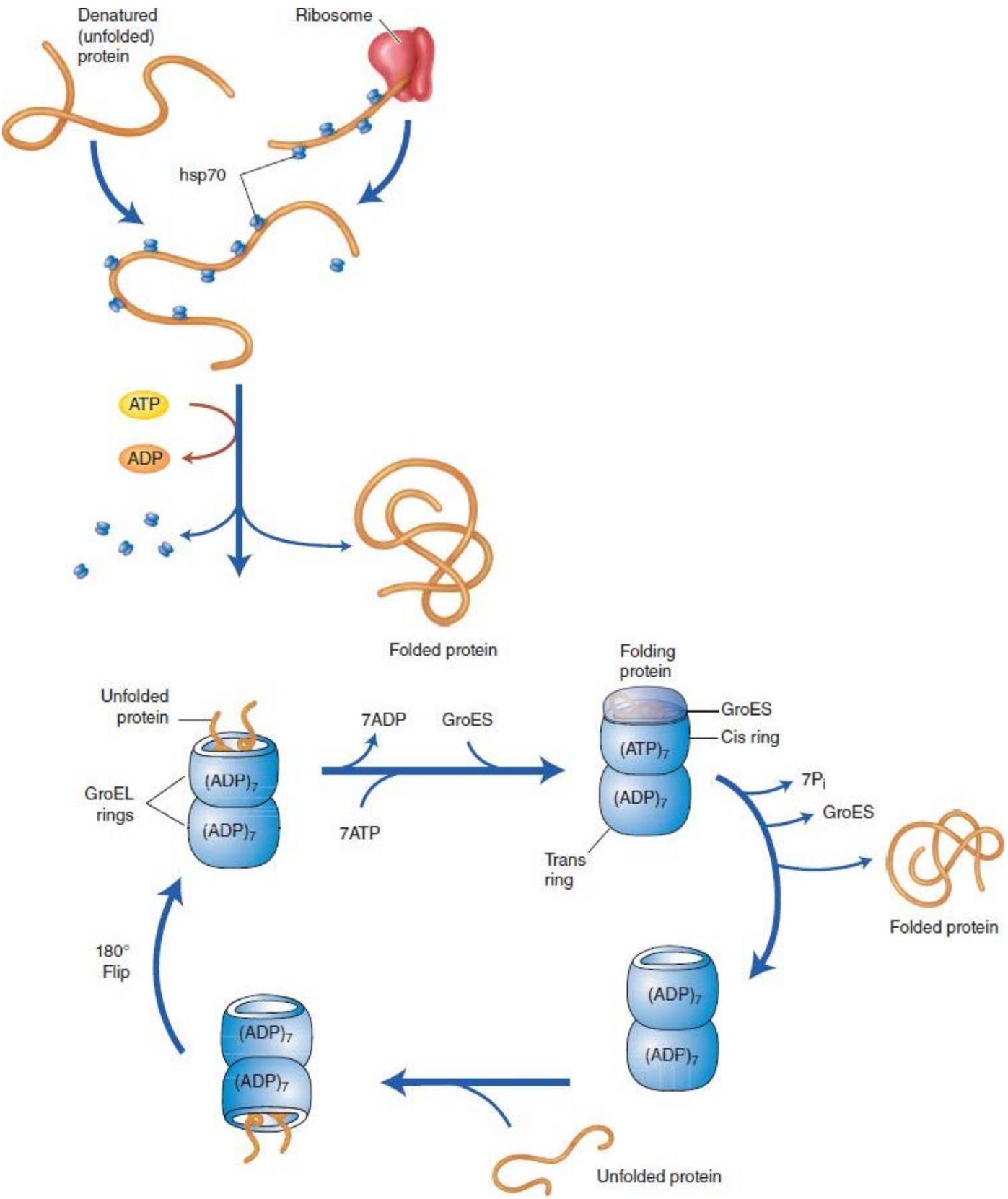


FIGURE 5.32

Molecular Chaperone-Assisted Protein Folding

Molecular chaperones bind transiently to both nascent proteins and unfolded proteins (i.e., those denatured by stressful conditions). The members of the hsp70 family stabilize nascent proteins and reactivate some denatured proteins. DnaK is the major bacterial hsp70. Many proteins also require hsp60 protein complexes to achieve their final conformations. In *E. coli*, cellular proteins that do not fold spontaneously require processing by the GroEL/GroES complex. [Note that during a folding cycle the GroEL ring capped by GroES is referred to as the

cis-ring. The other attached ring, the one that has not yet initiated a protein-folding cycle, is called the *trans*-ring.] At the beginning of a folding cycle, an unfolded protein (or protein domain) is loosely bound via hydrophobic interactions to the cavity entrance of one of the GroEL-$(ADP)_7$ rings. ADP/ATP exchange converts the cavity to a hydrophobic, expanded microenvironment that then traps the protein substrate under a GroES lid. Subsequently, sequential hydrolysis of the seven ATPs converts the cavity to a hydrophilic microenvironment, driving both the formation of the molten globule state of the protein substrate and the progression of the folding process. When all seven ATPs have been hydrolyzed, the hydrophobic surface of the cavity is reestablished, and GroES and the newly folded protein leave the GroEL ring. Meanwhile, the trans-GroEL-$(ADP)_7$ ring is already beginning the loading, trapping, and folding process for another unfolded protein or domain.

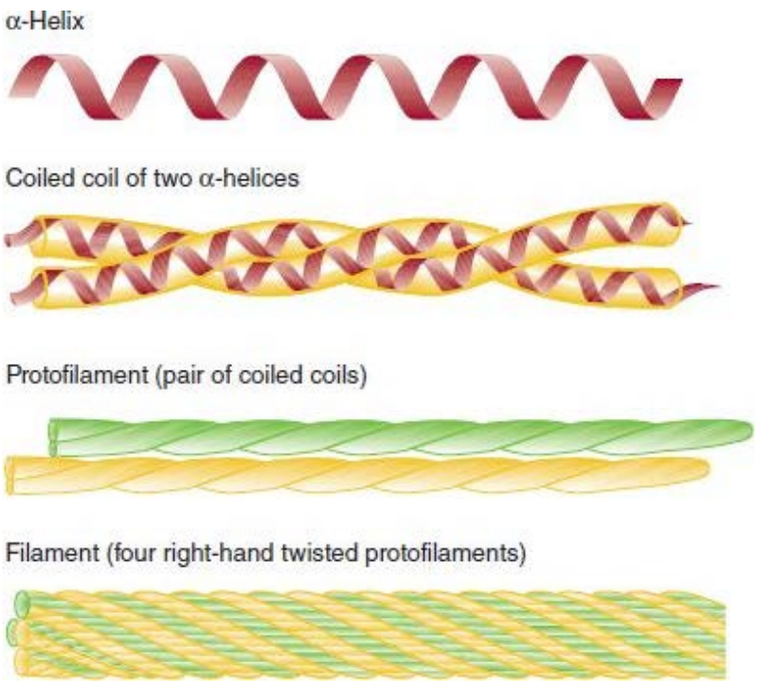


FIGURE 5.33

***α*-Keratin**

The *α*-helical rod-like domains of two keratin polypeptides form a coiled coil. Two staggered antiparallel rows of these dimers form a supercoiled protofilament. Hydrogen bonds and disulfide bridges are the principal interactions between subunits. Hundreds of filaments, each containing four protofilaments, form a macrofibril. Each hair cell, also called a fiber, contains several macrofibrils. Each strand of hair consists of numerous dead cells packed with keratin molecules. In addition to hair, the keratins are also found in wool, skin, horns, and fingernails.

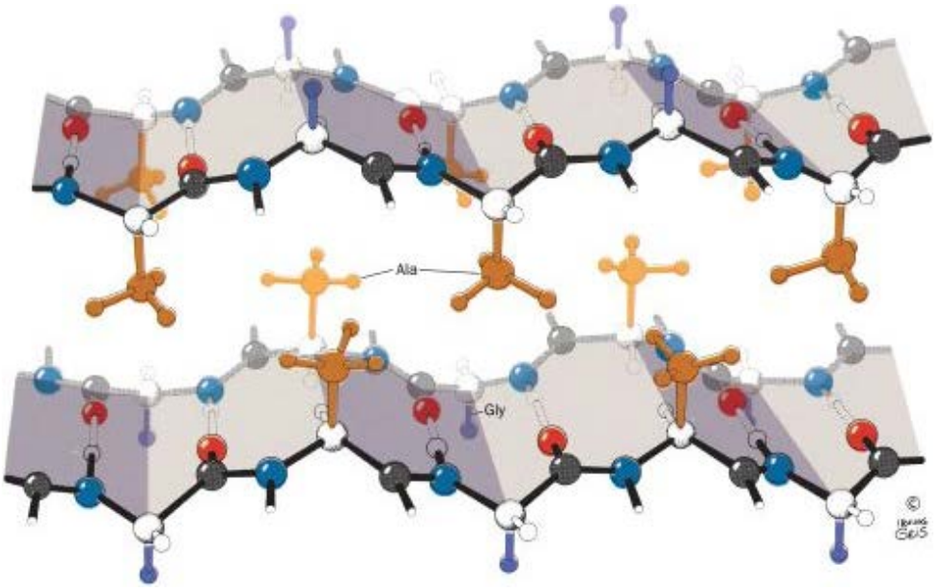


FIGURE 5.34

Molecular Model of Silk Fibroin

In fibroin, the silk fibrous protein produced by silkworms, the polypeptide chains are arranged in fully extended antiparallel β-pleated sheet conformations. Note that the R groups of alanine on one side of each β-pleated sheet interdigitate with similar residues on the adjacent sheet. Silk fibers (fibroin embedded in an amorphous matrix) are flexible because the pleated sheets are loosely bonded to each other (primarily with weak van der Waals forces) and slide over each other easily.

COLLAGEN Collagen is synthesized by connective tissue cells and then secreted into the extracellular space to become part of the connective tissue matrix. The 28 major families of collagen molecules include many closely related proteins that have diverse functions. The genetically distinct collagen molecules in skin, bones, tendons, blood vessels, and corneas impart to these structures many of their special properties (e.g., the tensile strength of tendons and the transparency of corneas).

Collagen is composed of three left-handed polypeptide helices that are twisted around each other to form a right-handed triple helix (**Figure 5.35**). Type I collagen molecules, found in teeth, bone, skin, and tendons, are about 300 nm long and 1.5 nm wide. Approximately 90% of the collagen found in humans is type I.

The amino acid composition of collagen is distinctive. Glycine constitutes approximately one-third of the amino acid residues. Proline and 4-hydroxyproline may account for as much as 30% of a collagen molecule's amino acid composition. Small amounts of 3-hydroxyproline and 5-hydroxylysine also occur. Specific proline and lysine residues in collagen's primary sequence are hydroxylated within the rough ER after the polypeptides have been synthesized. These reactions, which are discussed in Chapter 19, require ascorbic acid (p. 778).

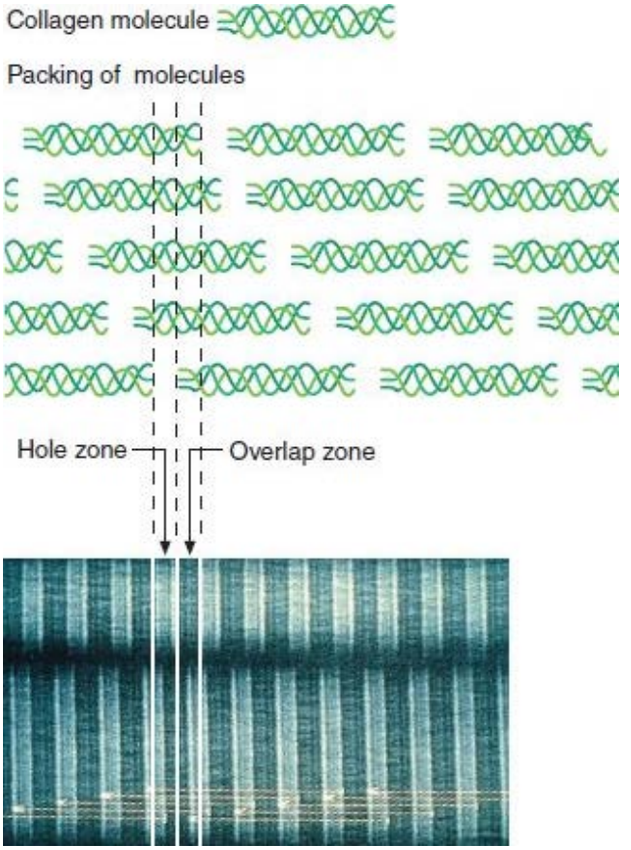


FIGURE 5.35

Collagen Fibrils

Staggered collagen molecules form the bands observed in this electron micrograph image. Cross-striations are about 680 Å apart. Each collagen molecule is about 3000 Å long.

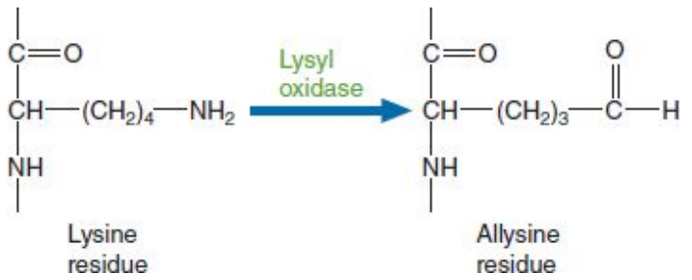
Collagen's amino acid sequence primarily consists of large numbers of repeating triplets with the sequence of Gly—X—Y, in which X and Y are often proline and hydroxyproline. Hydroxylysine is also found in the Y position. Simple carbohydrate groups are often attached to the hydroxyl group of hydroxylysine residues. It has been suggested that collagen's carbohydrate components are required for *fibrilogenesis*, the assembly of collagen fibers in their extracellular locations, such as in tendons and bone.

The enzyme lysyl oxidase converts some of the lysine and hydroxylysine side groups to aldehydes through oxidative deamination, and this facilitates the spontaneous nonenzymatic formation of strengthening aldimine and aldol cross-links. (An aldol cross-link is formed in a reaction, called an **aldol condensation**, in which two aldehydes form an α, β-unsaturated aldehyde linkage. In condensation reactions, a small molecule, in this case H_2O, is removed.) Cross-linkages also occur between hydroxylysine-linked carbohydrates and the amino group of other lysine and hydroxylysine residues on adjacent molecules. Increased cross-linking with age leads to the brittleness and breakage of the collagen fibers that occur in older individuals.

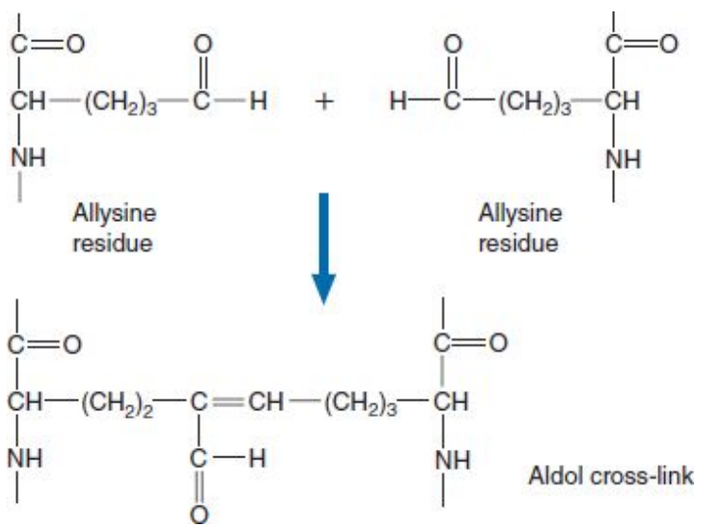
Glycine is prominent in collagen sequences because the triple helix is formed by interchain hydrogen bonding involving the glycine residues. Therefore every third residue is in close contact with the other two chains. Glycine is the only amino acid with an R group sufficiently small for the space available. Larger R groups would destabilize the superhelix structure. The triple helix is further strengthened by hydrogen bonding between the polypeptides (caused principally by the large number of hydroxyproline residues) and lysinonorleucine linkages that stabilize the orderly arrays of triple helices in the final collagen fibril.

QUESTION 5.10

Covalent cross-links contribute to the strength of collagen. The first reaction in cross-link formation is catalyzed by the copper-containing enzyme lysyl oxidase, which converts lysine residues to aldehyde allysine:



Allysine then reacts with other side chain aldehyde or amino groups to form cross-linkages. For example, two allysine residues react to form an aldol cross-linked product:



Lathyrism is a disease of humans and domestic animals caused by consumption of any of several species of the plant genus *Lathyrus*. The seeds of *Lathyrus odoratus* contain a toxin (*β*-aminopropionitrile) that inactivates lysyl oxidase. Consider the abundance of collagen in the tissues within animal bodies and suggest some likely symptoms of this malady.

Lathyrism

Globular Proteins

The biological functions of globular proteins usually involve the precise binding of small ligands or large macromolecules such as nucleic acids or other proteins. Each protein possesses one or more unique cavities or clefts whose structure is complementary to a specific ligand. After ligand binding, a conformational change occurs in the protein that is linked to a biochemical event. For example, the binding of ATP to myosin in muscle cells is a critical event in muscle contraction.

The oxygen-binding proteins myoglobin and hemoglobin are interesting and well-researched

examples of globular proteins. They are both members of the hemoproteins, a specialized group of proteins that contain the prosthetic group heme. Although the heme group (**Figure 5.36**) in both proteins is responsible for the reversible binding of molecular oxygen, the physiological roles of myoglobin and hemoglobin are significantly different. The chemical properties of heme are dependent on the Fe^{2+} ion in the center of the heme prosthetic group. Fe^{2+}, which forms six coordinate bonds, is bound to the four nitrogens in the center of the protoporphyrin ring. Two other coordinate bonds are available, one on each side of the planar heme structure. In myoglobin and hemoglobin, the fifth coordination bond is to the nitrogen atom in a histidine residue, and the sixth coordination bond is available for binding oxygen. In addition to serving as a reservoir for oxygen within muscle cells, myoglobin also facilitates the intracellular diffusion of oxygen. The role of hemoglobin, the primary protein of red blood cells, is to deliver oxygen to cells throughout the body. A comparison of the structures of these two proteins illustrates several important principles of protein structure, function, and regulation.

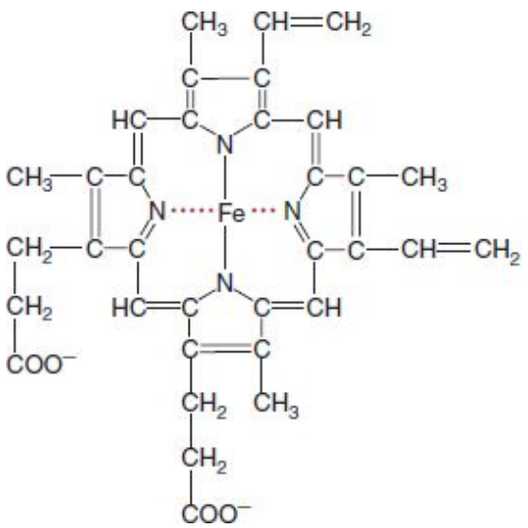


FIGURE 5.36

Heme

Heme consists of a porphyrin ring (composed of four pyrroles) with Fe^{2+} in the center.

MYOGLOBIN Myoglobin, found in high concentration in skeletal and cardiac muscle, gives these tissues their characteristic red color. The muscles of diving mammals such as whales, which remain submerged for long periods, have exceptionally high myoglobin concentrations. As a result, their muscles are typically brown. The protein component of myoglobin, called globin, is a single-polypeptide chain that contains eight segments of α-helix (**Figure 5.37**). The folded globin chain forms a crevice that almost completely encloses a heme group. Noncovalent interactions between amino acid side chains and the nonpolar porphyrin ring within myoglobin's oxygen-binding crevice decrease heme's affinity for O_2. The decreased affinity protects Fe^{2+} from oxidation and allows for the reversible binding of O_2. All of the heme-interacting amino acids are nonpolar except for two histidines, one of which (the proximal histidine) binds directly to the iron (**Figure 5.38**). The other (the distal histidine) stabilizes the oxygen-binding site. It is the globin fold within myoglobin, which is responsible for reversible oxygen binding since free heme [Fe^{2+}] has a high affinity for O_2 and is irreversibly oxidized to form hematin [Fe^{3+}], which cannot bind O_2.

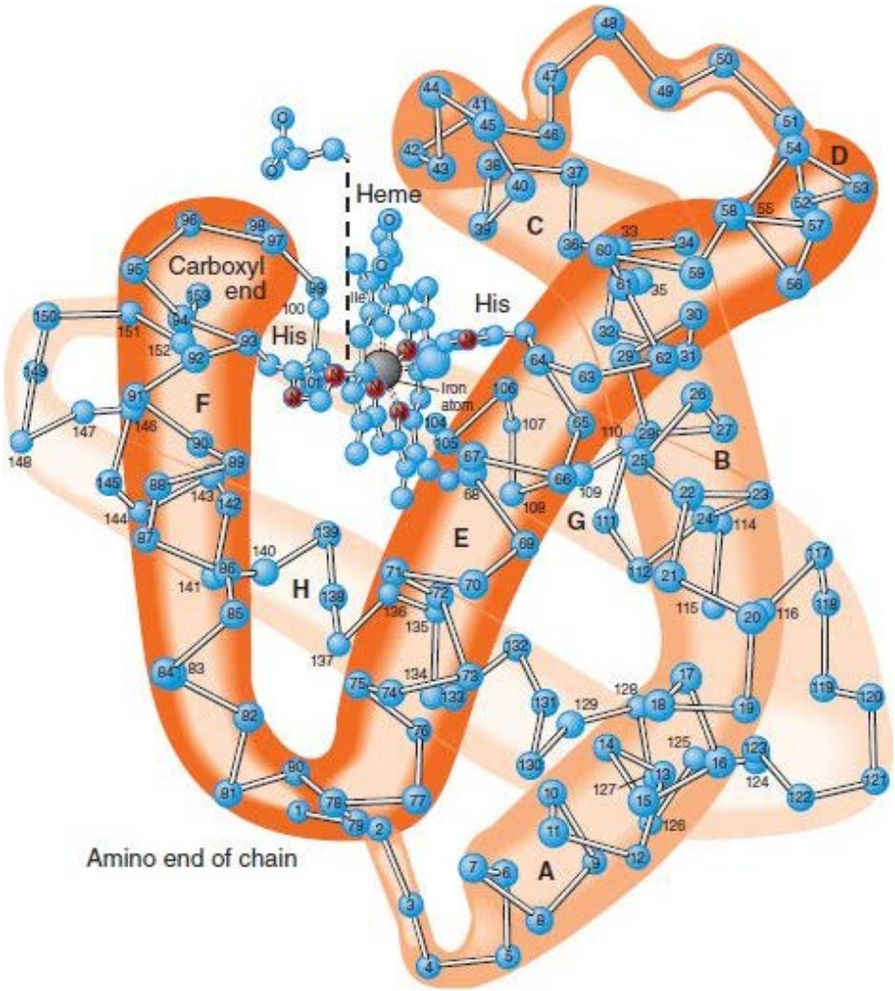


FIGURE 5.37

Myoglobin

With the exception of the side chain groups of two histidine residues, only the α-carbon atoms of the globin polypeptide are shown. Myoglobin's eight helices are designated A through H. The heme group has an iron atom that binds reversibly with oxygen. To improve clarity, one of heme's propionic acid side chains has been displaced.

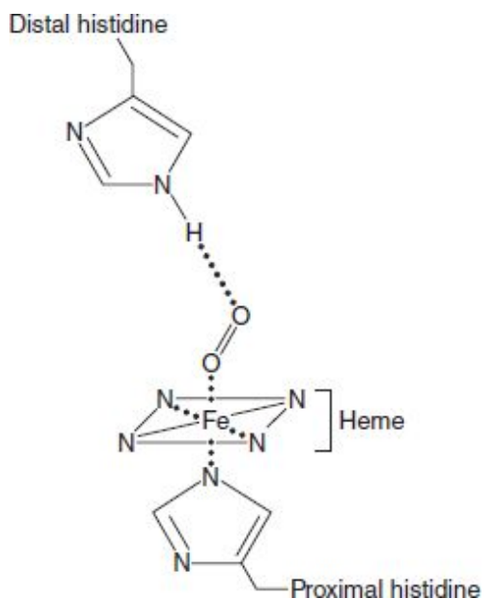


FIGURE 5.38

The Oxygen-Binding Site of Heme Created by a Folded Globin Chain

Myoglobin's oxygen-dissociation curve (**Figure 5.39**) is hyperbolic where O_2 affinity is high and O_2 saturation occurs at very low partial pressures of O_2 (pO_2). This reflects its structure as a single subunit protein and its role as an O_2 storage protein in muscle tissue. Myoglobin gives up its O_2 only when the muscle cell's oxygen concentration is very low (i.e., during strenuous exercise).

HEMOGLOBIN Hemoglobin is a roughly spherical molecule found in red blood cells, where its primary function is to transport oxygen from the lungs to every tissue in the body. The HbA molecule (**Figure 5.40**) is designated $\alpha_2\beta_2$ (HbA_2 is a variant designated $\alpha_2\delta_2$). Before birth, other variants of β-chains are produced: ε in embryonic life and γ in the fetus. Because $\alpha_2\varepsilon_2$ and $\alpha_2\gamma_2$ hemoglobins both have greater affinities for O_2 than $\alpha_2\beta_2$, the fetus can preferentially absorb oxygen from the maternal bloodstream.

HbA's four chains are arranged in two identical $\alpha\beta$-dimers. Each globin chain has a heme-binding pocket similar to that of myoglobin. Whereas myoglobin shows hyperbolic O_2-binding kinetics, hemoglobin exhibits a sigmoidal curve suggestive of **cooperative binding** (ligand binding by one subunit affects the binding behavior of other subunits) and allostery (ligand binding affected by effector molecules). (Cooperativity and allosteric regulation are described in Chapter 6.) When hemoglobin is oxygenated, specific salt bridges and hydrogen bonds between the $\alpha\beta$-dimers are disrupted, and the dimers slide past each other and rotate 15° relative to each other (**Figure 5.41**). The deoxygenated conformation (deoxyHb) is referred to as the *T(aut) state*, and the oxygenated conformation (oxyHb) is referred to as the *R(elaxed)* state. The oxygen-induced readjustment in the interdimer contacts is almost simultaneous. In other words, a conformational change in one subunit (i.e., the binding of an oxygen molecule) is rapidly propagated to the other subunits, resulting in further O_2 binding. Consequently, hemoglobin alternates between two stable conformations, the T and R states.

The oxygen dissociation curve of hemoglobin (**Figure 5.39**) has a sigmoidal shape because of T to R transitions. When an oxygen molecule binds to one subunit in the low-affinity T state, that subunit shifts to the high-affinity configuration (R state) and induces shifts in conformation in the remaining three subunits. Subsequent oxygen molecules bind with high affinity. This shift in binding properties, called cooperative binding, results in a shift from the T to R conformational forms of hemoglobin. In the lungs where pO_2 is high and the pH is high, hemoglobin is quickly saturated (converted to the R state). Hemoglobin also binds nitric oxide (NO) (p. 552), produced by vascular endothelial cells in the lungs, via thiol groups on the globin chains. NO is transported along with oxygen to the tissues in which CO_2 and H^+, the products of metabolism, are accumulating. In tissues where the pO_2 is low, CO_2 concentration is higher, and pH is lower than in the lungs, hemoglobin releases O_2 and NO. NO is transported out of the red blood cell via the anion exchanger AE-1 (p. 426) into the bloodstream, where it acts as a vasodilator, thereby increasing blood flow to the target tissue. O_2 is released as a result of proton-induced conformational changes in the Hb molecule. The α-amino groups of the globin chains are *carbamoylated* (CO_2 attached), stabilizing the T state for return to the lungs.

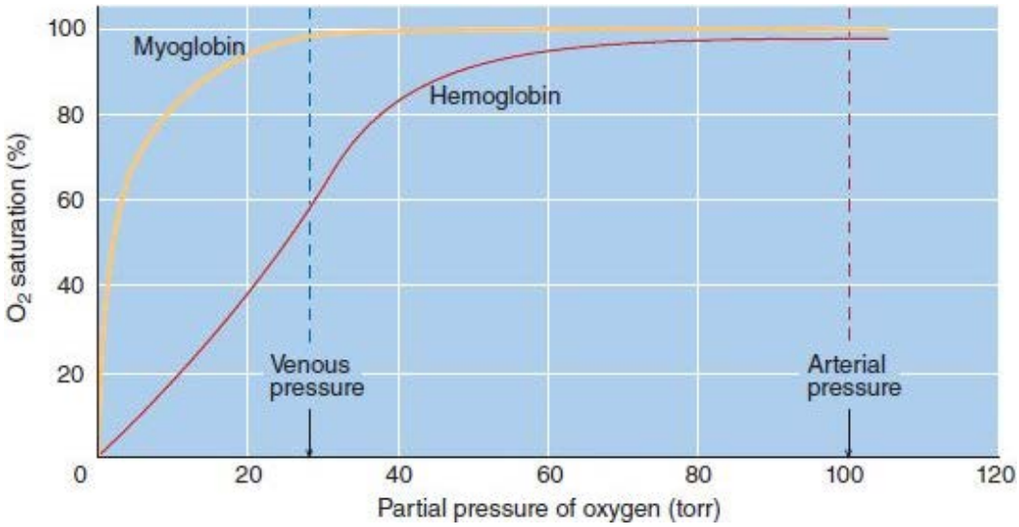


FIGURE 5.39

Dissociation Curves Measure the Affinity of Hemoglobin (red line) and Myoglobin (yellow line) for Oxygen.

Arterial blood, enriched in O_2, delivers it to the tissues. Venous blood, which drains from tissues, is O_2 depleted.

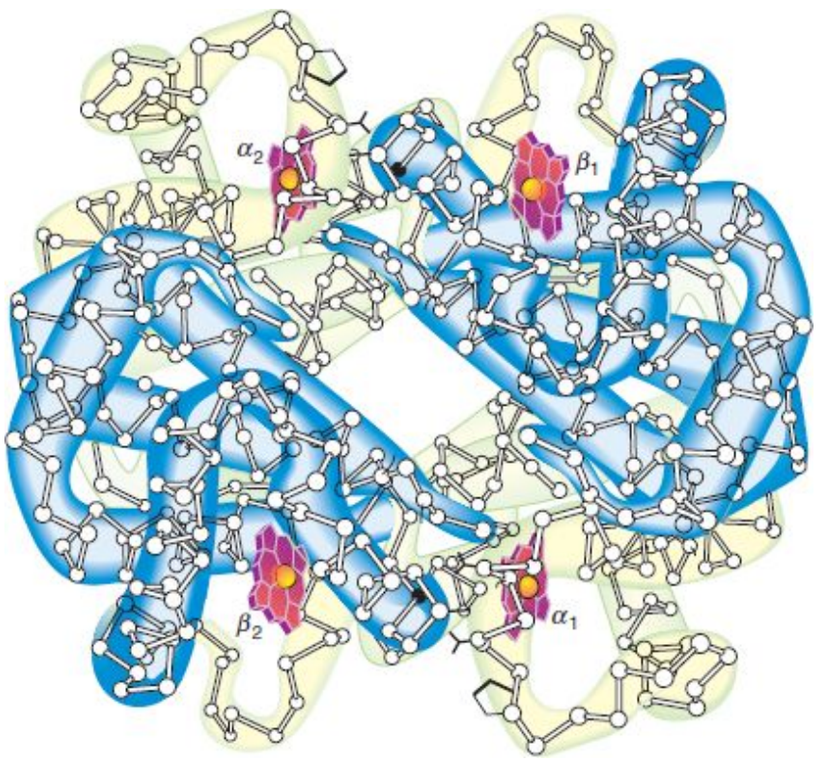


FIGURE 5.40

Hemoglobin Structure

The HbA protein contains four subunits, designated α and β. Each subunit contains a heme group that binds reversibly with oxygen.

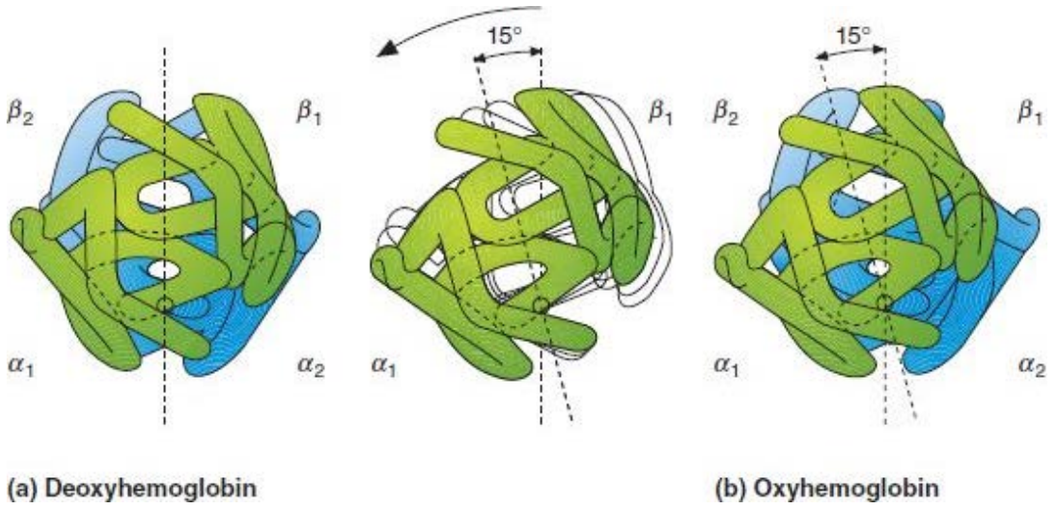


FIGURE 5.41

The Hemoglobin Allosteric Transition

When hemoglobin is oxygenated, the $\alpha_1\beta_1$ and $\alpha_2\beta_2$ dimers slide by each other and rotate 15º.

The *Bohr effect* describes the stabilization of the T state and the unloading of O_2 when pH decreases ($[H^+]$ increases). Metabolically active tissue produces large amounts of CO_2, which forms HCO_3^- and H^+ in blood. Protonation of the Hb subunits reduces salt bridge stabilization between the Hb dimers and induces an R-state to T-state transition facilitating the unloading of O_2 and NO.

2,3-Bisphosphoglycerate (BPG), an intermediate of glycolysis (glucose breakdown), binds to and stabilizes the T state, increasing the amount of O_2 that is unloaded in the tissue (**Figure 5.42**). Although most cells contain only trace amounts of BPG, red blood cells and the brain generate a considerable amount. BPG binds to the T state because a cavity lined with positively charged amino acid residue side chains is exposed in this conformation. Why is this important? BPG stabilizes deoxyHb, thus providing a mechanism for increasing O_2 unloading when energy demand is high.

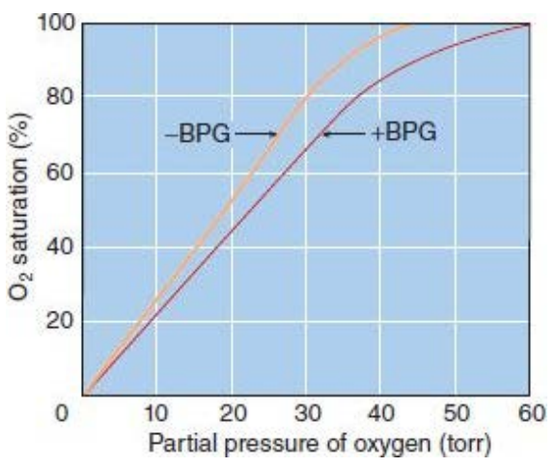


FIGURE 5.42

The Effect of 2,3-Bisphosphoglycerate (BPG) on the Affinity between Oxygen and Hemoglobin

In the absence of BPG (–BPG), hemoglobin has a high affinity for O_2; where BPG is present and binds to hemoglobin (+BPG), its affinity for O_2 decreases.

In the lungs, the process is reversed. A high oxygen concentration drives the conversion from the

deoxyHb configuration to that of oxyHb. The change in the protein's three-dimensional structure initiated by the binding of the first oxygen molecule releases bound CO_2, H^+, and BPG. The H^+ recombines with HCO_3^- to form carbonic acid, which then dissociates to form CO_2 and H_2O. Afterward, CO_2 diffuses from the blood into the alveoli and is then exhaled.

Carbon monoxide (CO) is a competitive inhibitor of hemoglobin because it binds to hemoglobin with an affinity 250 times that of O_2. Cyanide (CN^-) and NO also inhibit O_2 transport, but at higher concentration than CO. HbCO has a bright red color, so cherry-red skin is a symptom of CO poisoning. In addition, any chemical that oxidizes the Fe^{2+} results in the formation of Fe^{3+}-Hb or methemoglobin that does not bind O_2, although the oxidation state of iron in oxyHb is near 3. The extra electron in Fe^{2+} is required to coordinate with an unpaired electron in the O_2 molecule to generate oxyHb.

KEY CONCEPTS

- Globular protein function usually involves binding to small ligands or to other macromolecules.
- The oxygen-binding properties of myoglobin and hemoglobin are determined in part by the number of subunits they contain.

QUESTION 5.11

Fetal hemoglobin binds to BPG to a lesser extent than does HbA because His 143 in the BPG binding pocket in β-globin has been replaced with a serine residue in γ-globin. As a result of the loss of two positive charges (one for each of two γ-globins), the binding pocket binds BPG less avidly. What are the consequences of this phenomenon for mother and fetus?

QUESTION 5.12

Myoglobin stores O_2 in muscle tissue to be used by the mitochondria only when the cell is in oxygen debt, whereas hemoglobin can effectively transport O_2 from the lungs and deliver it discriminately to cells in need of O_2. Describe the structural features that allow these two proteins to accomplish separate functions.

Visit the companion website at www.oup.com/us/mckee to read Biochemistry in Perspective essays about the many functions of molecular machines made from proteins and details about the function of the molecular machine myosin.

Biochemistry IN PERSPECTIVE

Spider Silk and Biomimetics

What properties of spider silk have made it the subject of research worth hundreds of millions of dollars?

The female golden orb-web spider of Madagascar (*Nephila madagascariensis*) is a large spider (length = 12.7 cm or 5 in) that is so named because of its bright yellow silk. In a recent and astonishing effort, two Madagascar businessmen oversaw the creation of a hand-woven 11-foot brocaded spider silk tapestry. The large amount of silk required in this endeavor was obtained by literally harnessing thousands of orb-web spiders. (After gentle hand pulling of the dragline silk, the spiders were then released.) The silk fiber used in the weaving process was twisted into 96-ply thread. Amazingly, when the Madagascar tapestry was on display in a New York museum, the owners challenged an onlooker to break a thread in one of the tassels. Unable to do so, he compared its strength with that of a bicycle lock chain.

Biodegradable and lightweight spider silk is preferable to artificial fibers for a variety of applications, such as artificial tendon and ligament components, surgical thread (e.g., eye sutures), and lightweight armor (e.g., bulletproof vests and helmets). This fiber is desirable not only because of its remarkable mechanical properties, but also because spiders produce it at ambient temperature and pressure with water as the solvent. In contrast, Kevlar, an aramid (aromatic amide) polymer derived from petroleum, is synthesized by forcing an almost boiling mixture of monomers dissolved in sulfuric acid through the small holes of an industrial spinneret (a multipored device used to convert a polymer into individual fibers). However, despite enormous effort and millions of dollars of investment, most applications of spider silk have not materialized for the simple reason that there is no adequate source. The obvious solution would be spider farming, similar to the more than 5000-year-old practice (originating in China) of cultivating the domesticated silkworm moth (*Bombyx mori*). Unfortunately, silkworm silk is less tough and elastic than spider silk spider, and spider silk farming is not commercially feasible because spiders are cannibals. (In close quarters, these aggressive organisms proceed to eat one another.) Handling each of thousands of spiders separately is also unworkable because of its expense. (The creation of the Madagascar tapestry cost $500,000.)

An alternative strategy for producing artificial spider silk is biomimetic industrial synthesis. *Biomimetics* solves engineering problems by emulating biological processes such as the spinning of spider silk. Unfortunately, however, success has been limited. For example, attempts to use recombinant DNA technology to synthesize spider silk by inserting silk genes into bacteria and then recovering silk proteins have been disappointing for a variety of reasons (e.g., high costs and use of toxic solvents). There has been limited success with transgenic goats, animals into which a spider silk protein gene has been inserted. Efforts to spin the goat silk protein, purified from the animal's milk, resulted in fibers that had inferior mechanical properties with diameters (10–60 μm) that were considerably thicker than those of natural spider silk (2.5–5 μm). However, research efforts continue because achieving the goal of industrial engineers will certainly be worth the investment: biodegradable, environmentally safe artificial spider silk will offer an attractive alternative to petroleum-based fibers. Throughout this effort, scientists will further probe both spider silk structure and the biological spinning process.

Spider Silk Structure

Dragline silk is composed of two proteins, spidroin 1 and spidroin 2, which have molecular masses that range from 200 to 350 kDa. Spidroin amino acid composition is distinctive because the majority of residues are glycine (42%) and alanine (25%), with smaller amounts of amino acids with bulkier side chains (Arg, Tyr, Gln, Ser, Leu, and Pro). Both spidroin proteins contain two major types of repeating units: polyalanine sequences (5 to 10 residues) and glycine-enriched motifs such as polyGlyAla, GlyProGlyGly*X* (where *X* is often Gln), and GlyGly*X* (where *X* can be various amino

acids). In mature silk protein. the polyalanine and polyGlyAla sequences form antiparallel *β*-pleated sheets, the microcrystalline structures that give silk its tensile strength (**Figure 5A**). The *β*-pleated sheets are connected by glycine-enriched sequences that form random coil, *β*-spirals (similar to *β*-turns), and GlyGly*X* helical structures that together constitute an amorphous and elastic matrix.

Spider Silk Fiber Assembly

Production of dragline spider silk fiber provides a rare opportunity to observe protein folding as it occurs in a living organism. The spinning process (**Figure 5B**) begins in the ampullate gland in the spider's abdomen where epithelial cells secrete spidroin into the gland's lumen. Silk protein, referred to as silk feedstock or *dope*, is highly concentrated (as high as 50%). At this stage, spidroin's globular conformation (about 30% *α*-helices) ensures its solubility in water and prevents aggregation. The silk dope is squeezed through the ampullate gland and the narrow funnel that connects to the spinning duct. Here the flowing dope begins to assume the properties of a liquid crystal as long axes of the protein molecules are forced into parallel orientation. The tapered S-shaped spinning duct has three segments. As the protein moves through the segments, nascent silk polymer forms as a result of increasing shear stress (force applied by the parallel duct wall) and several biochemical environment changes. Within the duct, Na^+ and Cl^- are extracted, and phosphate and K^+ are pumped in. An increase in the K^+:Na^+ ratio, combined with the secretion of phosphate and H^+, is believed to cause the conversion of *α*-helical conformations to *β*-pleated sheets. At first randomly oriented, the *β*-pleated sheets are eventually forced into parallel alignment with the long axis of the filament. In the third segment of the duct, large amounts of water, released from the silk protein as hydrophobic interactions increase, are pumped out by epithelial cells. The valve at the end of the duct is believed to act as a clamp that grips the silk and as a means of restarting the spinning process if the silk breaks. The silk polymer then enters one of numerous spigots within a spinneret (**Figure 5C**). As the silk filament emerges and the remaining water evaporates, it is solid. The filaments from numerous spigots wrap around each other to form a cable-like fiber. The diameter and strength of the fiber depend on the muscular tension within the spinneret valve and how fast the spider draws it out.

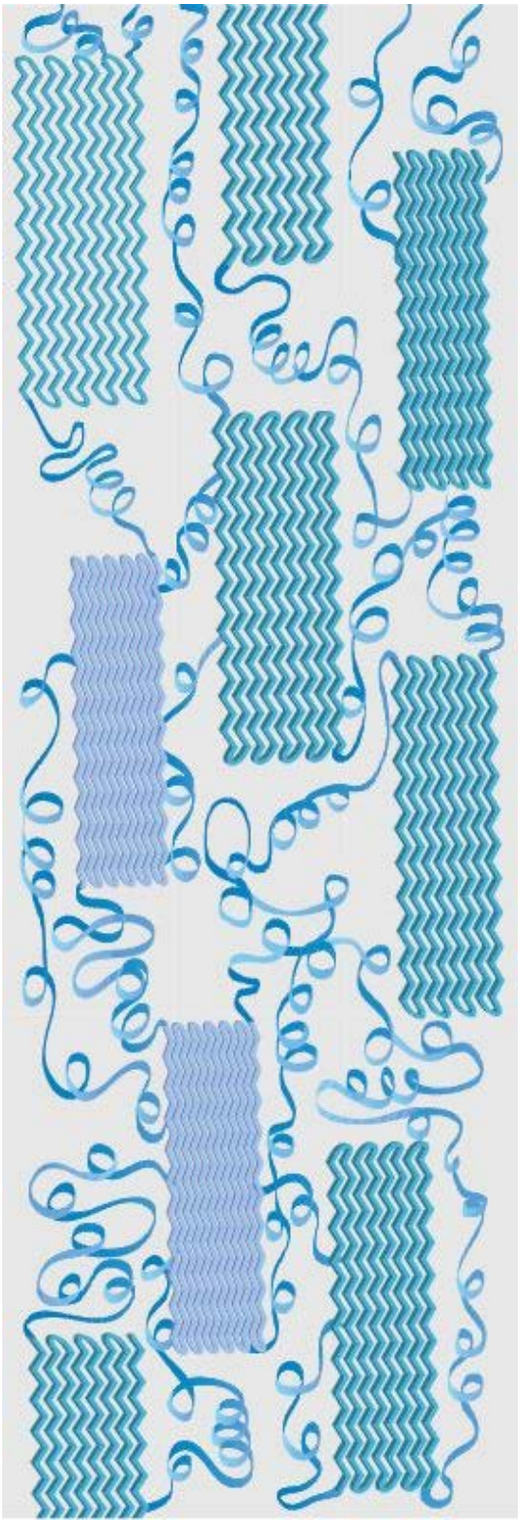


FIGURE 5A

Diagrammatic View of a Spider Silk Filament Segment

The structure of spider silk is not known with precision. It is known that two types of β-pleated sheet (highly ordered and less ordered) are responsible for the strength of spider silk. They are linked to each other by polypeptide sequences that form the random coil, left-hand helices, and β-spirals that provide elasticity. Silks differ in their β-pleated sheet and random coil content. For example, dragline silk has a higher content of β-pleated sheet than does capture silk.

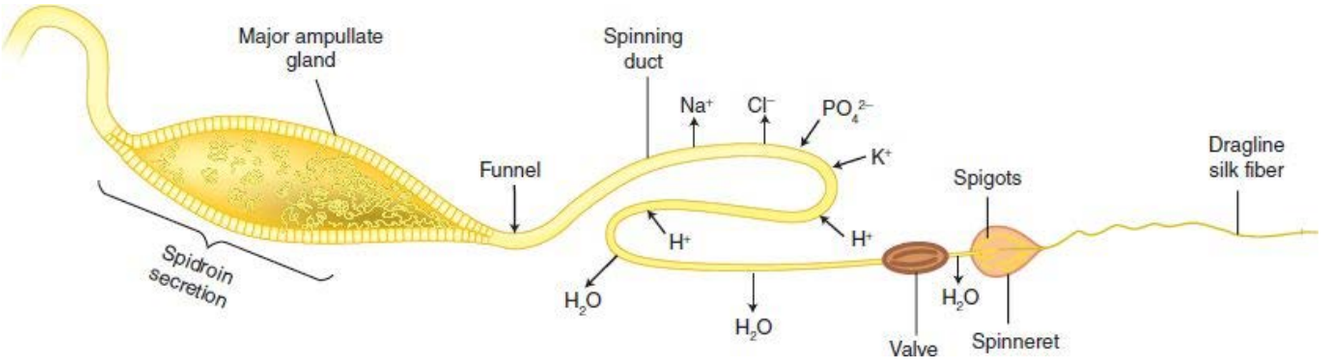


FIGURE 5B

Processing of Spider Dragline Silk

After the spidroins are secreted into the lumen of the major ampullate gland, they move toward the funnel, where they exit into the beginning of the spinning duct. As a result of shear stress and other forces (e.g., squeezing of the wall of the ampullate gland and pulling of the silk fiber out of the spinneret by the spider), the spidroins in the silk dope are compressed and forced to align along their long axes. As the silk polymer progresses down the tapering duct, biochemical changes (e.g., in Na^+, K^+, and H^+ concentrations) cause the conversion of α-helices into hydrophobic β-pleated sheets that expel H_2O. After passing through the valve, the polymer is forced through one of several spigots. Several emerging filaments are twisted together to form a silk fiber that is pulled out of the spinneret by the spider.

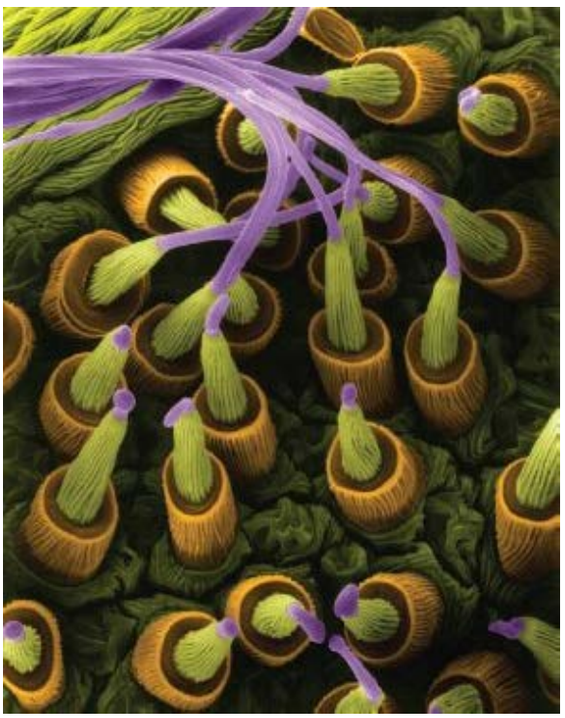


FIGURE 5C

Illustration of the Silk Spinning Spigots of a Spider Spinneret

Note that emerging filaments are twisting together to form a fiber.

SUMMARY Biodegradable, lightweight, strong spider silk has an enormous number of potential applications. Intense, and as yet unsuccessful, research efforts have focused on duplicating the natural process by which spiders produce this remarkable fiber.

Biochemistry IN THE LAB

Protein Technology

Living organisms produce a stunning variety of proteins. Consequently, it is not surprising that considerable time, effort, and funding have been devoted to investigating their properties. Since Frederick Sanger determined the amino acid sequence of bovine insulin in 1953, the structures of several thousand proteins have been elucidated.

In contrast to the 10 years required for insulin, current technologies allow protein sequence determination within a few days by mass spectrometry. The amino acid sequence of a protein can also be generated from its DNA or mRNA coding sequence if this information is available. After a brief review of protein purification methods, mass spectrometry is described. An older means of determining the primary sequence of polypeptides, the Edman degradation method, is described in an online Biochemistry in the Lab box—Protein Sequencing Analysis: The Edman Degradation. Note that all the techniques for isolating, purifying, and characterizing proteins exploit differences in charge, molecular weight, and binding affinities. Many of these technologies apply to the investigation of other biomolecules.

Purification

Protein analysis begins with isolation and purification. Extraction of a protein requires cell disruption and homogenization (see Biochemistry in the Lab: Cell Technology in Chapter 2). This process is often followed by differential centrifugation and, if the protein is a component of an organelle, by density gradient centrifugation. After the protein-containing fraction has been obtained, several relatively crude methods may be used to enhance purification. In **salting out**, high concentrations of salts such as ammonium sulfate [$(NH_4)_2SO_4$] are used to precipitate proteins. Because each protein has a characteristic salting-out point, this technique removes many impurities. (Unwanted proteins that remain in solution are discarded when the liquid is decanted.) When proteins are tightly bound to membrane, organic solvents or detergents often aid in their extraction. Dialysis (**Figure 5D**) is routinely used to remove low-molecular-weight impurities such as salts, solvents, and detergents.

As a protein sample becomes progressively more pure, more sophisticated methods are used to achieve further purification. Among the most commonly used techniques are chromatography and electrophoresis.

Chromatography

Originally devised to separate low-molecular-weight substances such as sugars and amino acids, chromatography has become an invaluable tool in protein purification. A wide variety of chromatographic techniques are used to separate protein mixtures on the basis of molecular properties such as size, shape, and weight or certain binding affinities. Often, several techniques must be used sequentially to obtain a demonstrably pure protein.

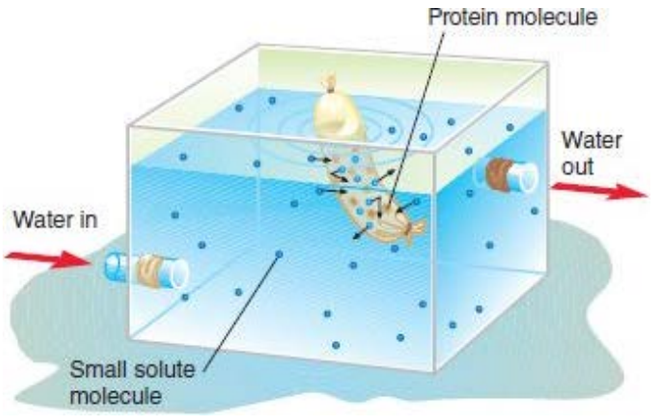


FIGURE 5D

Dialysis

Proteins are routinely separated from low-molecular-weight impurities by dialysis. When a dialysis bag (an artificial semipermeable membrane) containing a cell extract is suspended in water or a buffered solution, small molecules pass out through the membrane's pores. If the solvent outside the bag is continually renewed, all low-molecular-weight impurities are removed from the inside.

In all chromatographic methods, the protein mixture is dissolved in a liquid known as the **mobile phase**. As the protein molecules pass across the **stationary phase** (a solid matrix), they separate from each other because they are differently distributed between the two phases. The relative movement of each molecule results from its capacity to remain associated with the stationary phase while the mobile phase continues to flow.

Three chromatographic methods commonly used in protein purification are gel-filtration chromatography, ion-exchange chromatography, and affinity chromatography. **Gel-filtration chromatography** (**Figure 5E**) is a form of size-exclusion chromatography in which particles in an aqueous solution flow through a column (a hollow tube) filled with gel and are separated according to size. Molecules that are larger than the gel pores are excluded and therefore move through the column quickly. Molecules that are smaller than the gel pores diffuse in and out of the pores, thereby retarding their movement through the column. Differences in the rates of particle movement separate the protein mixture into bands, which are then collected separately.

Ion-exchange chromatography separates proteins on the basis of their charge. Anion-exchange resins, which consist of positively charged materials, bind reversibly with a protein's negatively charged groups. Similarly, cation-exchange resins bind positively charged groups. After proteins that do not bind to the resin have been removed, the protein of interest is recovered by an appropriate change in the solvent pH (altering the protein's net charge) and/or salt concentration.

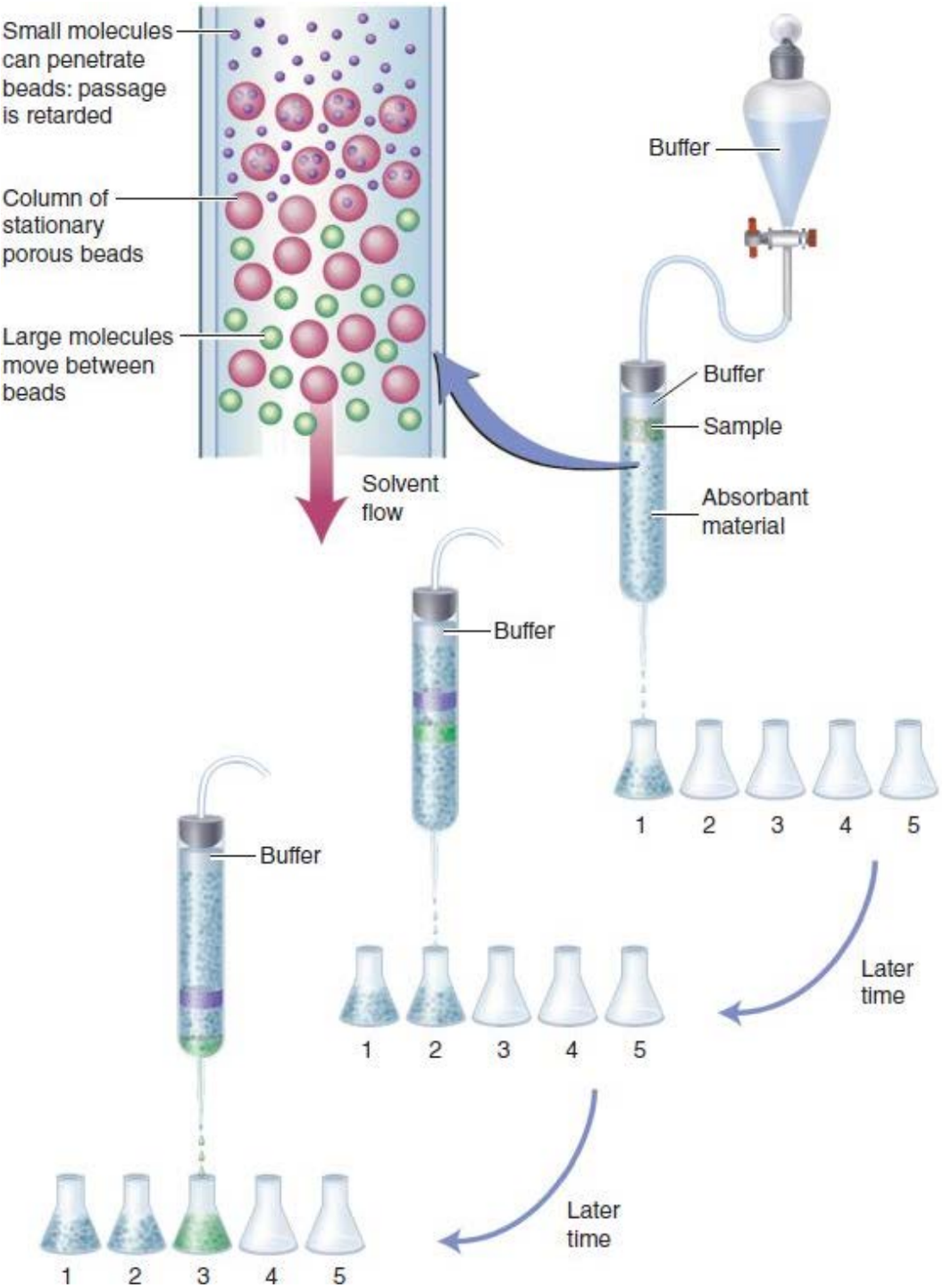


FIGURE 5E

Gel-Filtration Chromatography

In gel-filtration chromatography, the stationary phase is a gelatinous polymer, with pore sizes selected by the experimenter to separate molecules according to their sizes. The sample is applied to the top of the column and is eluted with buffer (the mobile phase). As elution proceeds, larger molecules travel faster through the gel than smaller molecules, whose progress is slowed because they can enter the pores. If fractions are collected, the larger molecules appear in the earlier fractions and later fractions contain smaller molecules.

Affinity chromatography is an efficient method of protein purification that takes advantage of the unique biological properties of proteins. That is, it uses a reversible noncovalent binding affinity between a specific protein and a special molecule (the ligand). The ligand is covalently bound to an

insoluble matrix, which is placed in a column. After nonbinding protein molecules have passed through the column, the protein of interest is removed by altering the conditions that affect protein-ligand binding (i.e., pH or salt concentration). Affinity chromatography is commonly used in the purification of recombinant proteins. (A recombinant protein is produced in a host organism by the transcription of a recombinant DNA sequence [pp. 712–15], followed by translation.) The recombinant protein is typically a fusion protein that contains an affinity tag such as GST (glutathione-S-transferase; see p. 549) that is grafted either to the N- or C-terminal of the protein of interest. GST is a small protein (26 kDa) that has a high binding affinity for GSH (p. 549). When a protein mixture is mixed with beads with GSH attached, the GST fusion protein will bind to the beads. After the beads are gently washed to remove other proteins, the beads are then washed with free GSH, which causes the detachment of the fusion protein. The protein of interest is recovered by the enzyme-catalyzed hydrolysis of a specific site between the protein of interest and the GST tag.

High-performance liquid chromatography (HPLC) is an extremely sensitive method that uses a pressurized liquid solvent containing a sample mixture through a column filled with an absorbent material to separate and identify components. Various forms of HPLC are used in protein analysis. Examples include ion-exchange and bioaffinity chromatography. HPLC is often used in association with mass spectrometry (p. 187).

Electrophoresis

Because proteins are electrically charged, they move in an electric field. In this process, called **electrophoresis**, molecules separate from each other because of differences in their net charge. As described previously, molecules with a positive net charge migrate toward the negatively charged electrode (cathode). Molecules with a net negative charge will move toward the positively charged electrode (anode). Molecules with no net charge will not move at all.

Electrophoresis, one of the most widely used techniques in biochemistry, is usually carried out using gels such as polyacrylamide or agarose. The gel, functioning much as it does in gel-filtration chromatography, also acts to separate proteins on the basis of their molecular weight and shape. Consequently, gel electrophoresis is highly effective at separating complex mixtures of proteins or other molecules.

Bands resulting from a gel electrophoretic separation may be treated in several ways. Specific bands may be excised from the gel after visualization with ultraviolet light. Each protein-containing slice is then eluted with buffer and prepared for further analysis. Because of its high-resolving power, gel electrophoresis is also used to assess the purity of protein samples. Staining gels with a dye such as Coomassie brilliant blue is a common method for quickly assessing the success of a purification step.

SDS–polyacrylamide gel electrophoresis (SDS-PAGE) is a widely used variation of electrophoresis that can be used to determine molecular weight (**Figure 5F**). SDS (sodium dodecyl sulfate), a negatively charged detergent, binds to the hydrophobic regions of protein molecules, causing the proteins to denature and assume rod-like shapes. Because most molecules bind SDS in a ratio roughly proportional to their molecular weights, during electrophoresis SDS-treated proteins migrate toward the anode (+ pole) only in relation to their molecular weight.

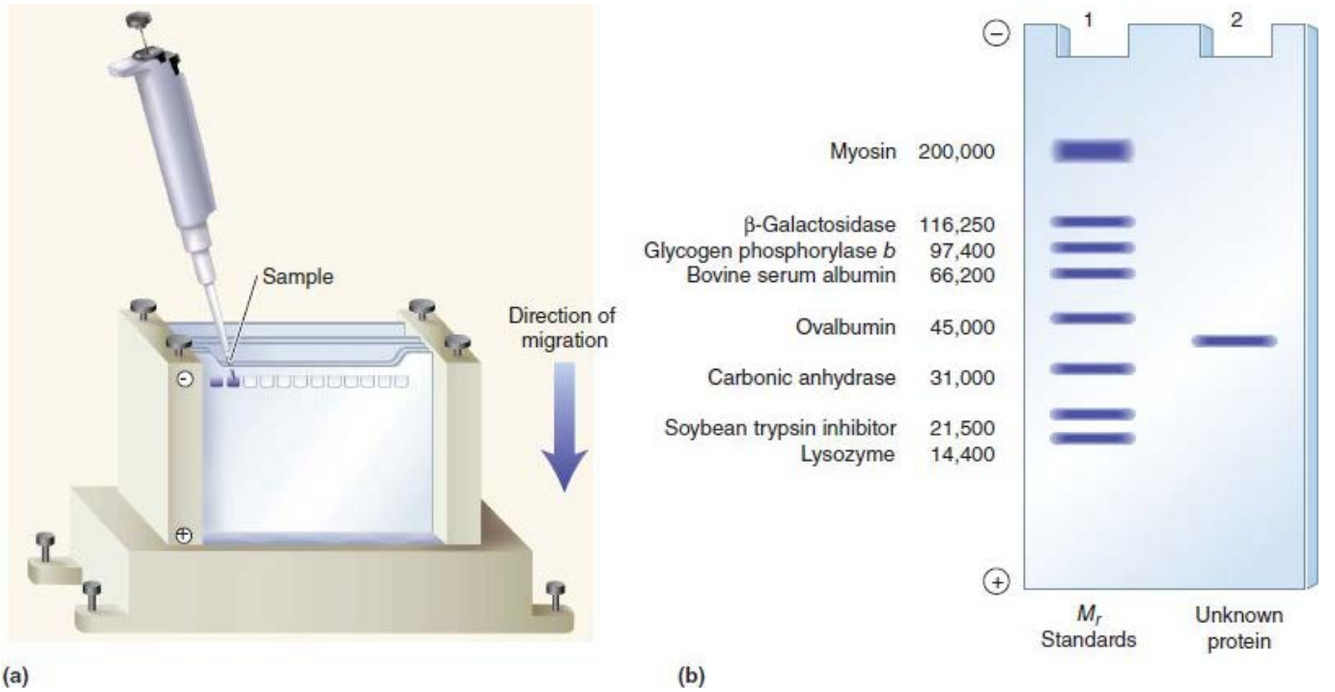


FIGURE 5F

Gel Electrophoresis

(a) Gel apparatus. The samples are loaded into wells. After an electric field is applied, the proteins move into the gel. (b) Molecules separate and move in the gel as a function of molecular weight and shape.

Mass Spectrometry

Mass spectrometry (MS) is a powerful, sensitive technique for separating, identifying, and determining the mass of molecules. It exploits differences in their mass-to-charge (*m*/*z*) ratios. In a mass spectrometer, ionized molecules flow through a magnetic field (**Figure 5G**). The magnetic field force deflects the ions depending on their *m*/*z* ratios, with lighter ions being more deflected from a straight-line path than heavier ions. A detector measures the deflection of each ion. In addition to protein identity and mass determinations, MS is also used to detect bound cofactors and protein modifications. Because MS analysis involves the ionization and vaporization of the substances to be investigated, its use in the analysis of thermally unstable macromolecules such as proteins and nucleic acids did not become feasible until methods such as electrospray ionization and matrix-assisted laser desorption ionization (MALDI) had been developed. In electrospray ionization, a solution containing the protein of interest is sprayed in the presence of a strong electrical field into a port in the spectrometer. As the protein droplets exit the injection device, typically an ultrafine glass tube, the protein molecules become charged. In MALDI, a laser pulse vaporizes the protein, which is embedded in a solid matrix. Once the sample has been ionized, its molecules, now in the gas phase, are separated according to their individual *m*/*z* ratios. A detector within the mass spectrometer produces a peak for each ion. In a computer-assisted process, information concerning each ion's mass is compared against data for ions of known structure and used to determine the sample's molecular identity.

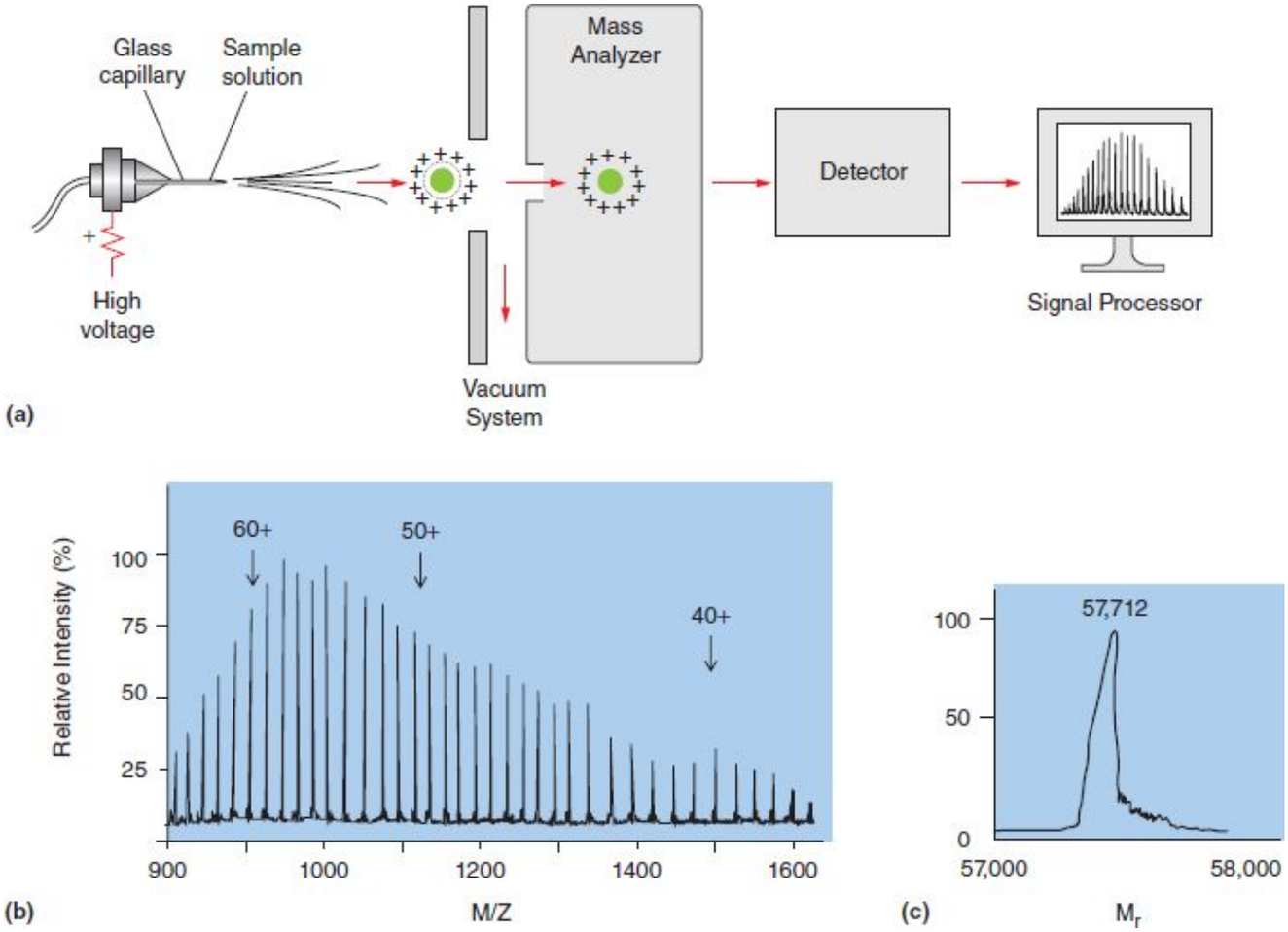


FIGURE 5G

Mass Spectrometry

(a) The principal steps in electrospray ionization. The sample (a protein dissolved in a solvent) is injected via a glass capillary into the ionization chamber. The voltage difference between the electrospray needle and the injection port results in the creation of protein ions. The solvent evaporates during this phase. The ions enter the mass spectrometer, which then measures their *m*/*z* ratios. (b) An electrospray mass spectrum showing the *m*/*z* ratios for several peaks. (c) A computer analysis of the data showing the molecular mass of the sample protein (M_r = relative molecular weight, that is, molecular mass divided by 1/12th of the mass of a single unbound carbon-12 atom).

Protein sequencing analysis makes use of tandem MS (two mass spectrometers linked in series, MS/MS). A protein of interest, often extracted from a band in a gel, is then digested by a proteolytic enzyme. Subsequently, the enzyme digest is injected into the first mass spectrometer, which separates the oligopeptides according to their *m*/*z* ratios. One by one, each oligopeptide ion is directed into a collision chamber, where it is fragmented by collisions with hot inert gas molecules. Product ions, peptides that differ from each other in size by one amino acid residue, are then sequentially directed into the second mass spectrometer. (Peptides with 30 or less amino acid residues generate characteristic fragmentation patterns that are used to identify them.) A computer identifies each peak in the spectrum and automatically determines the amino acid sequence of the peptides. The process is then repeated with oligopeptides derived from digestion with another enzyme. The computer uses the sequence information derived from both digests to determine the amino acid sequence of the original polypeptide.

Recently, a higher-resolution mass spectrometry technique, referred to as top-down MS, has been developed for protein structure analysis. In top-down MS (as opposed to the bottom-up technique of

digesting a protein before MS analysis), an intact protein or protein mixture is directly injected into an instrument called a *Fourier transform ion cyclotron resonance mass spectrometer* (FTICR-MS). In FTICR-MS signals generated by the accelerated movement of charged particles in a static magnetic field and a rapidly varying electric field are analyzed by performing a mathematical transformation called a Fourier transform to yield a mass spectrum. FTICR-MS has been used successfully to obtain high-resolution sequencing of proteins with masses under 100,000 kDa, as well as characterization of posttranslational modifications and protein–protein interactions.

In shotgun proteomics, a bottom-up technology, complex mixtures of proteins (e.g., the proteome of a specific cell type) can be analyzed. First, the mixture is digested by a protease and the resulting peptides are then separated by liquid chromatography. Tandem MS is used to identify the peptide fragments. Proteins are identified using software-enabled genomic and proteomic databases.

Protein Sequence-Based Function Prediction

Once a polypeptide has been isolated, purified, and sequenced, the next logical step is to determine its function. This endeavor usually begins with a database search of known protein sequences. BLAST (Basic Local Alignment Search Tool) is a computer program (www.ncbi.nim.nih.gov/blast) that allows fast searches of known sequences for matches to the unknown protein sequence (the query sequence). Protein sequence databases (e.g., UniProt [*Uni*versal *Prot*ein resource] www.uniprot.org) are sufficiently large that about 50% of sequence comparison queries yield matched sequences that are close enough to infer function.

X-Ray Crystallography

Much of the three-dimensional structural information about proteins has been obtained by X-ray crystallography. Because the bond distances in proteins are approximately 0.15 nm, the electromagnetic radiation used to resolve protein structure must have a short wavelength. Visible light wavelengths [(λ) = 400–700 nm] clearly do not have sufficient resolving power for biomolecules. X-rays, however, have very short wavelengths (0.07–0.25 nm).

In X-ray crystallography, highly ordered crystalline specimens are exposed to an X-ray beam (**Figure 5H**). As the X-rays hit the crystal, they are scattered by the atoms in the crystal. The diffraction pattern that results is recorded on charge-coupled device (CCD) detectors. The diffraction patterns are used to construct an electron density map. Because there is no objective lens to recombine the scattered X-rays, the three-dimensional image is reconstructed mathematically. Computer programs now perform these extremely complex and laborious computations. The three-dimensional structure of a polypeptide can also be determined using *homologous modeling*, a method that is based on the observation that three-dimensional protein structure is more conserved than protein sequences. A structural model is constructed from X-ray diffraction data of one or more homologous proteins in the Protein Data Bank (www.pdb.org).

NMR Spectroscopy

NMR is a commonly used form of spectroscopy based on the absorption of electromagnetic radiation (radio waves) by atomic nuclei with magnetic properties (i.e., isotopes that have nuclear spin such as ^{1}H, ^{13}C, and ^{15}N) that are aligned with a strong magnetic field. (*Nuclear spin* [p. P-4] is a form of angular momentum that occurs in atomic nuclei with an odd number of protons or neutrons. A spinning nucleus is similar to a miniature bar magnet.) Unlike X-ray crystallography, which is limited to proteins that crystallize easily, NMR can be used to investigate the structure of moderately

sized proteins (40–60 kDa), as well as other macromolecules, without regard to their crystallization capacity. The three-dimensional structure of protein molecules is determined using short pulses of radio waves, which disrupt the alignment of the nuclei along the external magnetic field. Since the total magnetic field experienced by each nucleus includes local magnetic fields of nearby atoms, each NMR-active nucleus will react differently to the radio waves. An NMR spectrum is a computer program–derived plot of radio pulse frequency against the energy absorbed by each of the molecule's nuclei. Once a protein's amino acid sequence is known, its folded structure is determined by analysis of several NMR experiments in which the properties and relative location of each atom in the molecule are resolved. NMR has another important advantage over X-ray crystallography: X-ray crystallography can only provide a snapshot of a crystallized protein, whereas NMR can be used in the investigation of protein dynamics (e.g., internal structural changes that are essential for a molecule's function). NMR can also provide information about a protein's responses to interactions with other molecules or changes in conditions such as temperature and pH.

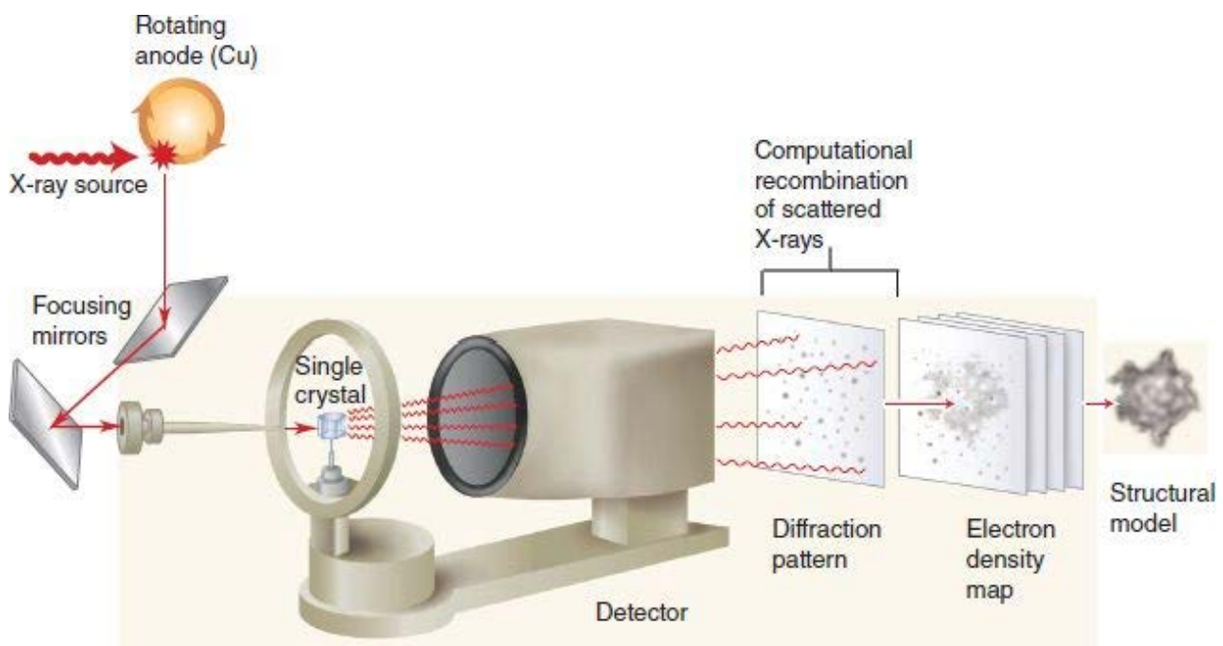


FIGURE 5H

Schematic Diagram of X-Ray Crystallography

X-rays are useful in the analysis of biomolecules because their wavelength range is similar to the magnitude of chemical bonds. Consequently, the resolving power of X-ray crystallography is equivalent to interatomic distances.

Chapter Summary

1. Polypeptides are amino acid polymers. Proteins may consist of one or more polypeptide chains.
2. Each amino acid contains a central carbon atom (the α-carbon) to which an amino group, a carboxylate group, a hydrogen atom, and an R group are attached. In addition to comprising protein, amino acids have several other biological roles. According to their capacity to interact with water, amino acids may be separated into four classes: nonpolar, polar, acidic, and basic.
3. Amino acids undergo several chemical reactions. Three reactions are especially important: peptide bond formation, cysteine oxidation, and Schiff base formation.
4. Proteins have a vast array of functions in living organisms. In addition to serving as structural materials,

proteins are involved in metabolic regulation, transport, defense, and catalysis. Some proteins are multifunctional; that is, they have two or more seemingly unrelated functions. Proteins can also be classified into families and superfamilies, according to their sequence similarities. Fibrous proteins (e.g., collagen) are long, rod-shaped molecules that are insoluble in water and physically tough. Globular proteins (e.g., hemoglobin) are compact, spherical molecules that are usually soluble in water.

5. Biochemists have distinguished four levels of protein structure. Primary structure, the amino acid sequence, is specified by genetic information. As the polypeptide chain folds, local folding patterns constitute the protein's secondary structure. The overall three-dimensional shape that a polypeptide assumes is called the tertiary structure. Proteins that consist of two or more polypeptides have quaternary structure.
6. Numerous proteins, especially molecules that participate in eukaryotic regulatory processes, are partially or completely unstructured.
7. Many physical and chemical conditions disrupt protein structure. Denaturing agents include strong acids or bases, reducing agents, organic solvents, detergents, high salt concentrations, heavy metals, temperature changes, and mechanical stress.
8. One of the most important aspects of protein synthesis is the folding of polypeptides into their biologically active conformations. Despite decades of investigation into the physical and chemical properties of polypeptide chains, the mechanism by which a primary sequence dictates the molecule's final conformation is unresolved. Many proteins require molecular chaperones to fold into their final three-dimensional conformations. Protein misfolding is now known to be an important feature of several human diseases, including Alzheimer's disease and Huntington's disease.
9. Fibrous proteins (e.g., α-keratin and collagen), which contain high proportions of α-helices or β-pleated sheets, have structural rather than dynamic roles.
10. Despite their varied functions, most globular proteins have features that allow them to bind to specific ligands or sites on certain macromolecules. These binding events involve conformational changes in the globular protein's structure.
11. The biological activity of complex multisubunit proteins is often regulated by allosteric interactions in which small ligands bind to the protein. Any change in the protein's activity is caused by changes in the interactions among the protein's subunits. Effector molecule binding can increase or decrease the function of a protein.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on amino acids, peptides, and proteins to help you prepare for exams.

Chapter 5 Review Quiz

Suggested Readings

Aebersold R, Mann M. 2016. Mass spectrometric exploration of proteome structure and function. Nature 537:347–55.

Arnaud CH. 2013. Top-down proteomics becomes reality. Chem Eng News 91(20):11–17.

Bustamonte C. 2004. Of torques, forces, and protein machines. Protein Sci 13:306l–5.

Caetano-Anolles G, et al. 2009. The origin, evolution and structure of the protein world. Biochem J 417:621–37.

Englander W, Mayne L. 2014. The nature of protein folding pathways. Proc Nat Acad Sci 111(45):15873–80.

Dunker AK, Kriwacki RW. 2011. The orderly chaos of proteins. Sci Am 304:68-73.

Everts S. 2013. Mirror molecules. Sci Am 306(5):79–81.

Heim M, Keerl D, Scheibel T. 2009. Spider silk: from soluble protein to extraordinary fiber. Angewandte Chem Int Ed 48:3584–96.

Kennedy M. 2012. Spider silk cape goes on show at V & A. The Guardian (London, UK), 24 Jan 2012.

Kim YE, et al. 2013. Molecular chaperone function in protein folding and proteostasis. Annu Rev Biochem 82:323–55.

Laptenko O, et al. 2016. The tail that wags the dog: how the disordered C-terminal domain controls the transcriptional activities of the p53 tumor suppressor. Trends Biochem Sci 41(12):1022–34.

Smith LM, Kelleher NL. 2013. Proteoform: a single term describing protein complexity. Nature Methods 10(3):186–7.

Uversky VN. 2016. p53 proteoforms and intrinsic disorder: an illustration of the protein structure–function continuum concept. Int J Mol Sci 17(11):1874.

van der Lee R, et al. 2014. Classification of intrinsically disordered regions and proteins. Chem Reviews 114:6589–631.

Wright PE, Dyson HJ. 2015. Intrinsically disordered proteins in cellular signaling and regulation. Nat Rev Mol Cell 16(1):18–29.

Key Words

affinity chromatography, *185*
aldimine, *148*
aldol condensation, *175*
aliphatic hydrocarbon, *138*
allosteric transition, *164*
allostery, *163*
Alzheimer’s disease, *172*
amino acid residue, *136*
amphipathic molecule, *168*
amphoteric molecule, *136*
apoprotein, *153*
aromatic hydrocarbon, *138*
asymmetric carbon, *141*
chaperonins, *171*
chiral carbon, *141*
conjugated protein, *153*
cooperative binding, *178*
denaturation, *167*
disulfide bridge, *147*
effector, *164*
electrophoresis, *186*
enantiomer, *141*
fibrous protein, *152*
fold, *159*

gel-filtration chromatography, *184*
globular protein, *152*
glycoprotein, *153*
heat shock protein, *152*
hemoprotein, *153*
holoprotein, *153*
homologous polypeptide, *153*
hormone, *140*
hsp70, *171*
hsp90, *171*
Huntington’s disease, *172*
intrinsically disordered protein, *165*
intrinsically disordered region, *165*
ion-exchange chromatography, *184*
isoelectric point, *143*
ligand, *160*
lipoprotein, *153*
mass spectrometry, *187*
metalloprotein, *153*
mobile phase, *184*
modular protein, *160*
modulator, *164*
molecular chaperone, *170*
molecular disease, *155*
molten globule, *170*
mosaic protein, *160*
motif, *158*
multifunctional protein, *152*
neurotransmitter, *140*
oligomer, *162*
optical isomer, *141*
peptide, *136*
peptide bond, *145*
phosphoprotein, *153*
polypeptide, *136*
primary structure, *153*
prosthetic group, *153*
protein, *136*
protein family, *152*
protein folding, *158*
protein superfamily, *152*
proteoform, *166*
protomer, *162*

quaternary structure, *153*
response element, *165*
salt bridge, *160*
salting out, *184*
Schiff base, *148*
SDS–polyacrylamide gel electrophoresis, *186*
secondary structure, *153*
stationary phase, *184*
stereoisomer, *141*
subunit, *162*
supersecondary structure, *158*
tertiary structure, *153*
zwitterion, *136*

Review Questions

SECTION 5.1

Comprehension Questions

1. Define the following terms:
   a. aliphatic hydrocarbon
   b. amino acid
   c. neurotransmitter
   d. hormone
   e. zwitterion
2. Define the following terms:
   a. chiral carbon
   b. stereoisomer
   c. enantiomer
   d. optical isomer
   e. isoelectric point
3. Define the following terms:
   a. peptide bond
   b. Schiff base
   c. disulfide bridge
   d. aldimine
   e. amphoteric
4. Define the following terms:
   a. α-carbon
   b. hydrophobic amino acid
   c. polar amino acid
   d. acidic amino acid

e. basic amino acid

5. Define the following terms:

   a. polypeptide
   b. peptide
   c. protein
   d. dipeptide
   e. hydroxyproline

6. Indicate whether each of the following amino acids is polar:

   a. lysine
   b. tyrosine
   c. leucine
   d. asparagine
   e. proline
   f. cysteine
   g. glutamic acid
   h. valine
   i. histidine
   j. glycine

Fill in the Blanks

7. Molecules with both a positive and a negative charge, but an overall neutral charge, are called _____.
8. The amide linkages of proteins are called _____.
9. Amino acid polymers consisting of more than 50 amino acids are called _____.
10. Primary amine groups of amino acids react with aldehydes or ketones to form _____.
11. Cystine is unique among the amino acids because it contains a _____ bond.
12. Molecules that can behave as an acid or as a base are called _____.
13. Nonsuperimposable mirror-image stereoisomers are referred to as _____.

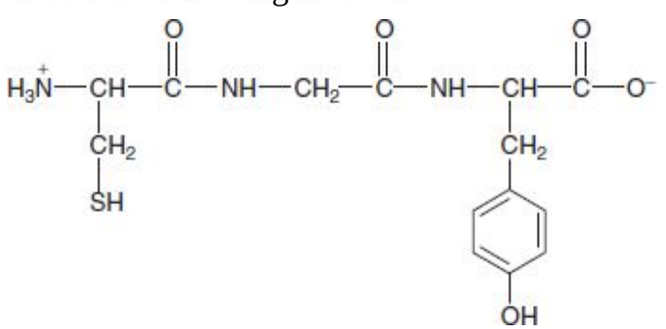
Short-Answer Questions

14. With the exception of glycine, all of the standard amino acids have one chiral center. Are there any amino acids that have two chiral centers?
15. Lysine has the following pK_a values:

    $pK_1 = 2.18$ $pK_2 = 8.95$ $pK_R = 10.79$

    Give the structure and net charge of lysine at the following pH values: 1, 4, 7, 10, and 12.

16. Consider the following molecule:



a. Name this molecule.
b. Use the three-letter symbols for the amino acids to represent this molecule.

17. Rotation about the peptide bond in glycylalanine is hindered. Draw the resonance forms of the peptide bond and explain why.
18. Distinguish between peptides, polypeptides, and proteins.
19. Draw the structure of alanylisoleucylphenylalanine. Calculate its isoelectric point.
20. Name four examples of biologically active derivatives of the standard amino acids.
21. Describe why penicillamine is used to treat cystinuria.
22. Proteins are often regulated by phosphorylation/dephosphorylation reactions. What amino acid residues can be phosphorylated?

Critical-Thinking Questions

23. The peptide bond is a stronger bond than the ester bond. What structural feature of the peptide bond gives it additional bond strength?
24. Glucose is a six-carbon polyhydroxyaldehyde. Determine the structure of the product of the reaction of glycine with glucose.
25. Consider the tripeptide Glu-Asp-Phe. What is the approximate isoelectric of this molecule? In which direction will the tripeptide move in an electric field at pH 1, 5, 10, and 12.

SECTION 5.2

Comprehension Questions

26. Define the following terms:

a. glutathione
b. atrial natriuretic factor
c. vasopressin
d. oxytocin
e. homeostasis

27. Define the following terms:

a. peroxide
b. GSSG
c. aquaporin
d. hypothalamus
e. renin

Fill in the Blanks

28. _____ is a steroid that promotes Na^+ and water retention by the kidney.
29. _____ inhibits the release of molecules that promote increased blood pressure.
30. _____ is a peptide that protects cells by reacting with peroxides.
31. _____ in the hypothalamus monitor blood Na^+ and when necessary trigger vasopressin secretion.
32. _____ is the product when the iron atom in hemoglobin is oxidized to its ferric form.

Short-Answer Questions

33. Describe the different effects of vasopressin and atrial natriuretic factor on blood pressure.
34. Describe the functions of glutathione within cells.
35. Why is the cellular ratio of GSH to GSSG in cytoplasm usually high?

Critical-Thinking Questions

36. Vasopressin and oxytocin are examples of peptide molecules with more than one function. Provide examples of these functions and then speculate as to why organisms use the same molecules in seemingly different processes.

SECTION 5.3

Comprehension Questions

37. Define the following terms:

    a. homologous polypeptides
    b. protein fold
    c. ligand
    d. oligomer
    e. protomer

38. Define the following terms:

    a. holoprotein
    b. heat shock protein
    c. protein family
    d. protein superfamily
    e. moonlighting protein

39. Define the following terms:

    a. metalloprotein
    b. primary structure
    c. secondary structure
    d. tertiary structure
    e. quaternary structure

40. Define the following terms:

    a. intrinsically disordered protein
    b. intrinsically disordered region
    c. prosthetic group
    d. apoprotein
    e. molecular chaperone

41. Define the following terms:

    a. β-hairpin
    b. supersecondary structure
    c. fibrous protein
    d. globular protein

e. mosaic protein

42. Define the following terms:

    a. salt bridge
    b. allosteric transition
    c. molecular disease
    d. protein denaturation
    e. α-helix

43. Define the following terms:

    a. β-strand
    b. β-barrel
    c. fibronectin
    d. botulinum toxin
    e. conjugated protein

Fill in the Blanks

44. The blood-clotting proteins _____ and _____ prevent blood loss when blood vessels are damaged.
45. Na^+K^+ATPase and hemoglobin are examples of _____ proteins.
46. _____ are rod-shaped proteins that are insoluble in water.
47. _____ and _____ linkages are covalent bonds between polypeptides that are the result of the oxidation of lysine side chains.
48. The function of proteins such as casein and zein is _____.
49. Keratin and fibrinogen are examples of _____ proteins.

Short-Answer Questions

50. Indicate the level(s) of protein structure to which each of the following contributes:

    a. amino acid sequence
    b. β-pleated sheet
    c. hydrogen bond
    d. disulfide bond

51. At what level of protein structure does each of the following denaturation act?

    a. heat
    b. strong acid
    c. saturated salt solution
    d. organic solvents (e.g., alcohol or chloroform)

52. A polypeptide has a high p*I* value. Suggest which amino acid residues might comprise it.
53. Outline the steps to purify a protein. What criteria are used to evaluate purity?
54. Describe the forces involved in protein folding.
55. The muscles of deep diving mammals such as whales contain exceptionally large amounts of myoglobin. How does this circumstance contribute to prolonged dives?
56. List three factors that do not foster α-helix formation.
57. Briefly outline the roles of ribosome-associated chaperones, hsp70, hsp90, and chaperonins in protein folding.

Critical-Thinking Questions

58. A mutational change alters a polypeptide by substituting three adjacent prolines for three glycines. What possible effect will this effect have on the protein's structure?
59. Why are moonlighting proteins necessary and/or desirable?
60. The folding of intrinsically disordered proteins is prevented by the presence of Ser, Lys, and Glu residues. What do these amino acids have in common? Why do they disrupt IDPs?
61. Why do some ordered proteins require molecular chaperones to fold into their active conformations, whereas others do not?
62. Hydrophobic amino acids play an important role in forming and maintaining the three-dimensional structure of ordered proteins. Can you suggest how these molecules perform this feat?
63. Proteins that are synthesized by living organisms adopt a biologically active conformation. Yet when such molecules are prepared in the laboratory, they usually fail to spontaneously adopt their active conformations. Can you suggest why?
64. When the moonlighting protein glyeraldehyde-3-phosphate dehydrogenase (GAPDH) catalyzes a key reaction in glycolysis (a metabolic pathway in cytoplasm), it does so as a homotetramer (four identical subunits). The GAPDH monomer is a nuclear DNA repair enzyme. Describe in general terms what structural properties of moonlighting proteins allow this phenomenon.
65. As a genetic engineer, you have been given the following task: alter a protein's structure by converting a specific amino acid sequence that forms an extended α-helix to one that forms a β-barrel. What types of amino acid are probably in the α-helix, and which ones would you need to substitute.
66. The active (catalytic) site of an enzyme contains the side chains of amino acid residues that are conserved because they participate in the protein's catalytic activity. The bulk of the enzyme, however, is not part of the active site. A substantial amount of energy is required to synthesize enzymes. Why are these molecules so large?
67. The synthesis of 2,3-BPG from the glycolytic intermediate 1,3-BPG (glycerate-1,3-bisphosphate) is catalyzed by the enzyme bisphosphoglycerate mutase, one of many iron-requiring enzymes. What effect on oxygen delivery to the body's tissues would be expected in the event of an iron-deficient diet?
68. A structural protein may incorporate large amounts of immobilized water as part of its structure. Can you suggest how protein molecules "freeze" the water in place and make it part of their structures?
69. What effect would hyperventilation (rapid breathing) have on the concentration of oxyhemoglobin in the bloodstream?
70. Describe the structural feature of p53 that allows it to exert its significant control over cell proliferation.
71. Why is it advantageous for p53 to be activated by factors such as ER stress, light, and hypoxia (low oxygen concentration)?

MCAT Study Questions

72. How many tetrapeptides can be synthesized from one molecule each of lysine, proline, glutamine, and cysteine?

    a. 4
    b. 12
    c. 16

d. 24

73. What is the charge on lysine at a pH of 8?

a. 0
b. +2
c. +1
d. −2

74. Peptide bond cleavage is a _____ reaction.

a. hydration
b. hydrolysis
c. addition
d. elimination

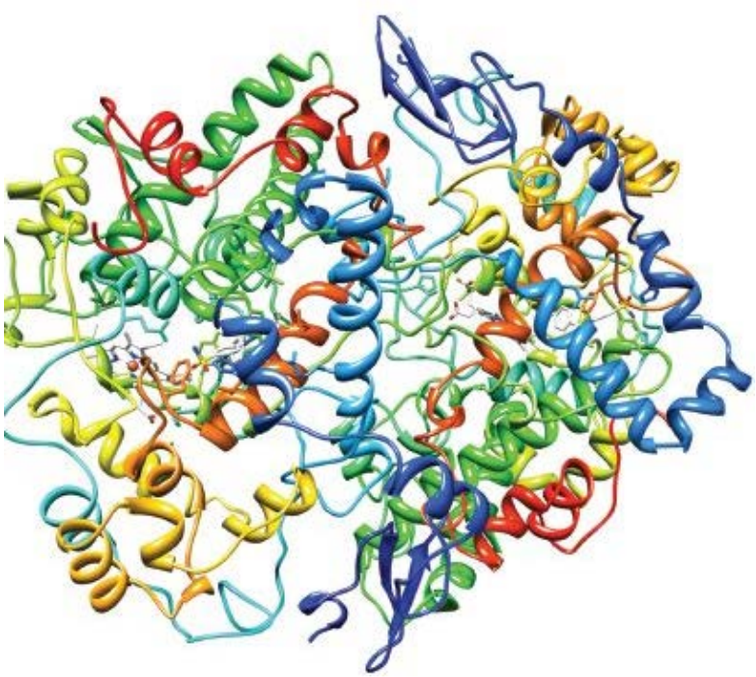
75. The disulfide bonds of ribonuclease are cleaved by which of the following?

a. digestion by chymotrypsin
b. incubation at pH 6
c. reaction with diisopropylphosphate at pH 2
d. use of β-mercaptoethanol and 8 M urea at pH 8

76. Which of the following amino acids would not be found in an α-helix?

a. Ala
b. Ser
c. Gly
d. Leu

CHAPTER 6 Enzymes



Cyclooxygenase Cyclooxygenase-2 (COX-2), shown here in its monomer form, is one of two human COX enzymes that catalyze the first reaction in the conversion of the 20-carbon fatty acid arachidonic acid into prostaglandins and thromboxanes (pp. 406–07). COX-2 is upregulated in inflammatory processes. Aspirin reduces pain and inflammation because it irreversibly inhibits COX enzymes.

OUTLINE

PAIN, ASPIRIN, AND THE COX ENZYMES

6.1 PROPERTIES OF ENZYMES
Enzyme Catalysts: The Basics
Enzymes: Activation Energy and Reaction Equilibrium
Enzymes and Macromolecular Crowding Effects
Enzyme Specificity

6.2 CLASSIFICATION OF ENZYMES

6.3 CATALYSIS
Organic Reactions and the Transition State
Transition State Stabilization
Catalytic Mechanisms
The Roles of Amino Acids in Enzyme Catalysis
The Role of Cofactors in Enzyme Catalysis

Effects of Temperature and pH on Enzyme-Catalyzed Reactions
Detailed Mechanisms of Enzyme Catalysis

6.4 ENZYME KINETICS
Michaelis–Menten Kinetics
Lineweaver–Burk Plots
Multisubstrate Reactions
Enzyme Inhibition
Enzyme Kinetics, Metabolism, and Macromolecular Crowding

6.5 ENZYME REGULATION
Genetic Control
Covalent Modification
Allosteric Regulation
Compartmentation

Biochemistry in Perspective
Alcohol Dehydrogenase: A Tale of Two Species

AVAILABLE ONLINE
Biochemistry in Perspective
Enzymes and Clinical Medicine
Biochemistry in Perspective
Quantum Tunneling and Catalysis
Biochemistry in Perspective
Humans and Enzymes—A Brief History

Pain, Aspirin, and the Cox Enzymes

Pain. Virtually all humans are familiar with the unpleasant sensations of pain (ranging from mild discomfort to agony) and the inflammatory processes that accompany it. All types of pain can be separated into two forms: acute and chronic. Acute pain, resulting from phenomena such as cuts, burns, bee stings, broken bones, and childbirth, occurs suddenly. The inflammatory and healing processes that are normally triggered by these traumas typically resolve within days or, at most, several months. Chronic pain continues for long periods of time and is often linked to long-term disorders such as rheumatoid arthritis, osteoarthritis, back pain, fibromyalgia, diabetic neuropathy (p. 617), and cancer. Chronic pain may begin with an injury, but in some cases it seems to begin without any apparent body damage.

Pain, as unpleasant as it is, is a vitally important warning system that protects the body. For example, when a barefooted pedestrian steps on a broken piece of glass, she is almost immediately motivated to remove the glass piece and then to clean and protect the wound. Without pain sensations, the same individual stepping on the glass would continue walking, thereby driving the glass deeper into the foot and risking significantly more damage.

Pain Perception

Pain signals are initiated when mechanical, thermal, or chemical stimuli activate receptors in the bare nerve endings of primary sensory neurons called *nociceptors*. Damaged cells release numerous substances, including the inflammatory peptides substance P and bradykinin, as well as nitric oxide (p. 552), histamine (p. 551), and prostaglandins (p. 406), among others. These substances contribute to pain perception and the inflammatory processes that usually lead to healing. Prostaglandin E2 (PGE2) promotes all the signs of

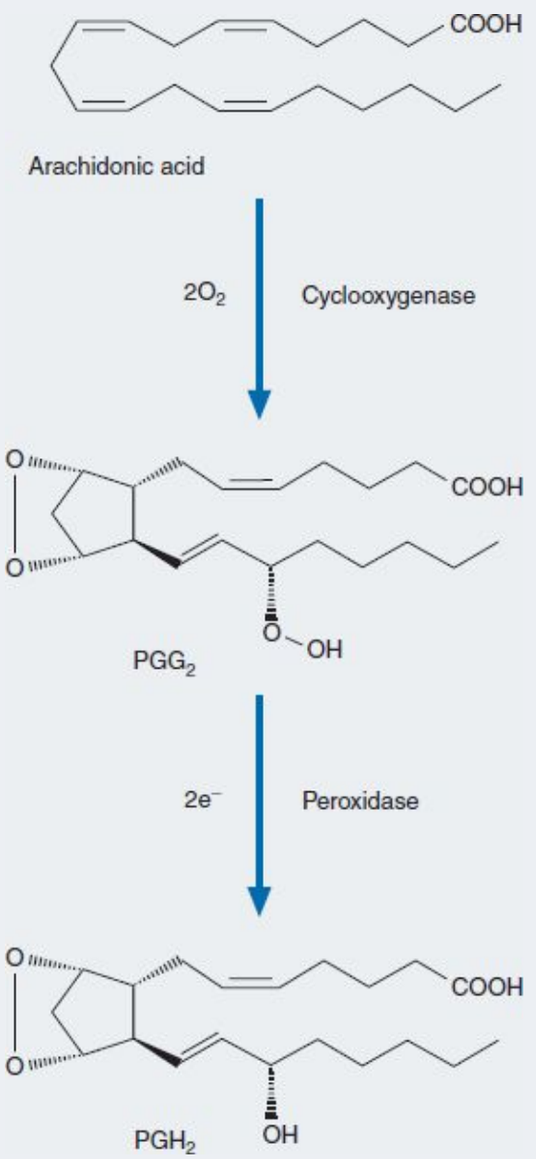
inflammation: redness, swelling, and pain. It increases pain perception during inflammation by lowering the threshold for pain signaling in sensory neurons.

Aspirin

For thousands of years pain has often been relieved by extracts of willow tree bark. Scientific investigations of willow tree bark ultimately resulted in the discovery of its active principle: a bitter substance called *salicin*, which is converted to salicylic acid in the body. In 1899, the Bayer Company began selling a less irritating form of salicylic acid in tablets called acetylsalicylic acid, or aspirin.

Aspirin, the most widely used medication in the modern world, inhibits pain perception and inflammation by preventing the conversion of a 20-carbon fatty acid called *arachidonic acid* into PGG_2 (prostaglandin G_2), the precursor of a group of molecules called the *eicosanoids* (prostaglandins and thromboxanes, p. 406), which have diverse physiological functions. Aspirin inhibits pain perception by irreversibly inhibiting an enzyme called *prostaglandin endoperoxide synthase* (PTGS).

PTGS has two active sites: cyclooxygenase and peroxidase. In the cyclooxygenase-catalyzed reaction, arachidonic acid reacts with two molecules of O_2 to yield PGG_2, a bicyclic hydroperoxide. The peroxidase activity then catalyzes the reaction of PGG_2 with the reducing agent GSH (p. 549) to yield PGH_2, the precursor of the other eicosanoids. Aspirin inhibits PTGS by acetylating a serine residue within the cyclooxygenase active site.



The COX Enzymes

There are two forms of PTGS, called COX-1 and COX-2; both of them function as dimers. COX-1 is a constitutive enzyme: in other words, it is regularly synthesized in a wide variety of tissues where it functions in numerous processes, such as cell signaling and tissue maintenance (e.g., protection of the mucosal lining of the gastrointestinal tract). In contrast, COX-2 is usually only synthesized during inflammatory processes. Aspirin inhibits both enzymes; this fact explains several possible side effects of aspirin, the most prominent of which is gastrointestinal bleeding.

Overview

BIOCHEMISTS HAVE INVESTIGATED ENZYMES (BIOLOGICAL CATALYSTS) FOR MORE THAN 140 YEARS. LONG BEFORE THEY HAD ANY REALISTIC understanding of the physical basis of the living state, biochemists instinctively appreciated the importance of enzymes. Using the technologies devised by biochemists, life scientists gradually determined the properties of biological systems. This work eventually demonstrated that almost every event in living organisms occurs because of enzyme-catalyzed reactions. Until recently, all known enzymes were proteins, but groundbreaking research led to the revelation that RNA molecules also have catalytic properties. This chapter is devoted to catalytic proteins. The characteristics of catalytic RNA molecules are described in Chapter 18.

Without enzymes, most of the thousands of biochemical reactions that sustain living processes would occur at imperceptible rates. Recent determinations of uncatalyzed (unenhanced) reaction rates in water range from 5 seconds for CO_2 hydration to 1.1 billion years for glycine decarboxylation. In contrast, enzyme-catalyzed reactions typically occur within time frames ranging from micro- to milliseconds. Enzymes are in fact the means by which living organisms channel the flow of energy and matter. Today, as a result of accumulating evidence derived from protein dynamics (conformational motion studies) and macromolecular crowding analysis, enzyme research is undergoing revolutionary changes. For example, according to long-held views, enzyme function depends almost entirely on the complementary shapes and catalytic interactions between reactant molecules and their more or less flexible binding sites. Recently, however, investigators have demonstrated that the catalytic function of certain enzymes can be linked to internal motions that extend throughout the protein molecule. Similarly, it is now recognized that enzymes function in conditions that are vastly different from those traditionally studied (i.e., purified molecules in dilute concentration). Instead, the in vivo (“in life”) milieu of enzymes is a crowded gel-like environment. As a result of recent investigations, models of enzyme kinetics are evolving, and methods of experimentation, data collection, and computer simulations are becoming more sophisticated and closer to realistic in vivo conditions. This chapter reviews the structural and functional properties of enzymes.

6.1 PROPERTIES OF ENZYMES

Enzymes have several remarkable properties (**Table 6.1**). First, the rates of enzymatically catalyzed reactions are often phenomenally high. Rate increases of 10^7 to 10^{19} have been observed. Second, in

marked contrast to inorganic catalysts, the enzymes are highly specific to the reactions they catalyze, and side products are rarely formed. Finally, because of their relatively large and complex structures, enzymes can be regulated. This is an especially important consideration in living organisms, which must conserve energy and raw materials.

TABLE 6.1 Key Characteristics of Enzymes

• Increase reaction rates
• Obey the laws of thermodynamics (i.e., no effect on K_{eq} values)
• Catalyze the forward and backward reactions of reversible reactions
• Usually present in low concentrations because they are not consumed by reactions
• Are controlled via regulatory mechanisms
• Transition state of reacting substrates bound in enzyme active sites

Enzyme Catalysts: The Basics

How do enzymes work? The answer to this question requires a review of the role of catalysts. By definition, a **catalyst** enhances the rate of a chemical reaction but is not permanently altered by the reaction. Catalysts perform this feat because they decrease the activation energy required for a chemical reaction. In other words, catalysts provide an alternative reaction pathway that requires less energy (**Figure 6.1**). The free energy of activation, $\Delta G^{\ddagger}$, is defined as the amount of energy required to convert 1 mol of **substrate** (reactant) molecules from the ground state (the stable, low-energy form of a molecule) to the **transition state**, which is the structural form of the reactant molecule that has the highest energy along the reaction coordinate. Since the transition state possesses more energy than the reactant or the product, it is also the least stable. In the reaction in which ethanol is oxidized to form acetaldehyde

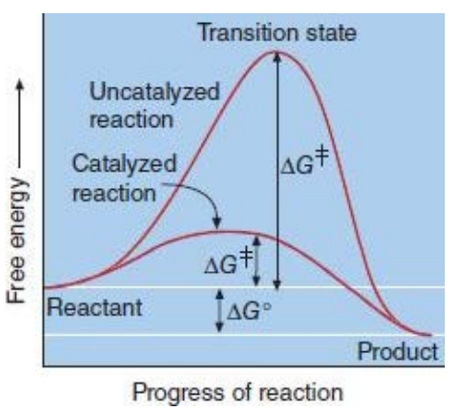
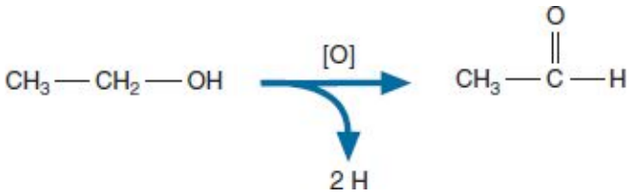


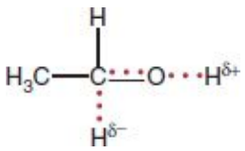
FIGURE 6.1

A Catalyst Reduces the Activation Energy of a Reaction

A catalyst alters the free energy of activation $\Delta G^{\ddagger}$, not the standard free energy $\Delta G°$ of the reaction. The transition state occurs at the apex of both reaction pathways.



this transition state might look like



Note that the alcoholic H is beginning to leave as a H^+ ion and the methylene H is beginning to leave as a hydride ion ($H:^-$).

Enzymes: Activation Energy and Reaction Equilibrium

To proceed at a viable rate, most chemical reactions require an initial input of energy. At temperatures above absolute zero (0 K, or –273.1°C), all molecules possess vibrational energy, which increases as the molecules are heated. Consider the following spontaneous reaction:

$$A + B \rightarrow C$$

As the temperature rises, vibrating molecules A and B are more likely to collide. A chemical reaction occurs when the colliding molecules possess a minimum amount of energy called the **activation energy** (E_a) or, more commonly in biochemistry, the *free energy of activation* ($\Delta G^{\ddagger}$). Not all collisions result in chemical reactions because only a fraction of the molecules have sufficient energy or the correct orientation to react (i.e., to break bonds or rearrange atoms into product molecules). Increasing collisions by raising the temperature or increasing reactant concentrations can improve product formation rates. In living systems, however, elevated temperature is unrealistic because of structural damage to biomolecules, and the concentration of most reactants is relatively low. Living organisms use enzymes to circumvent these restrictions. Each type of enzyme contains a unique, intricately shaped binding surface called an **active site**. Each active site is a cleft or crevice in a large protein molecule into which substrate molecules can bind via weak reversible noncovalent interactions in a catalysis-promoting orientation. The active site is more than a binding site, however. Several of the amino acid side chains that line the active site actively participate in the catalytic process.

The shape and charge distribution of an enzyme's active site constrains the motions and allowed conformations of the substrate, forcing it to adopt a conformation more like that of the transition state. In other words, the structure of the active site is used to optimally orient the substrate. As a result, the enzyme–substrate complex converts to product and free enzyme without the high-energy requirement of the constrained transition state. Consequently, the reaction rate increases significantly over that of the uncatalyzed reaction. Several other factors (described in Section 6.3) also contribute to rate enhancement.

Enzymes, like all catalysts, cannot alter the equilibrium of the reaction, but they can increase the rate toward equilibrium. Consider the following reversible reaction:

$$A \leftrightharpoons B$$

Without a catalyst, the reactant A is converted into the product B at a certain rate. Because this is a reversible reaction, B is also converted into A. The rate expression for the forward reaction is $k_F[A]^n$, and the rate expression for the reverse reaction is $k_R[B]^m$. The superscripts n and m represent the order of a reaction. Reaction order reflects the mechanism by which A is converted to B and vice versa. A reaction order of 2 for the conversion of A to B indicates that it is a bimolecular process and that two molecules of A must collide for the reaction to occur (Section 6.3). At equilibrium, the rates for the forward and reverse reactions must be equal:

$$k_F[A]^n = k_R[B]^m \tag{1}$$

which rearranges to

$$\frac{k_F}{k_R} = \frac{[B]^m}{[A]^n} \tag{2}$$

The ratio of the forward and reverse constants is the equilibrium constant:

$$K_{eq} = \frac{[B]^m}{[A]^n} \tag{3}$$

For example, in Equation (3), if $m = n = 1$, and $k_F = 1 \times 10^{-3}\ s^{-1}$, and $k_R = 1 \times 10^{-6}\ s^{-1}$, then

$$K_{eq} = \frac{10^{-3}}{10^{-6}} = 10^3$$

At equilibrium, therefore, the ratio of products to reactants is 1000 to 1.

In a catalyzed reaction, both the forward rate and the backward rate are increased, but the K_{eq} (in this case, 1000) remains unchanged. If the catalyst increases both the forward and the reverse rates by a factor of 100, then the forward rate becomes 100,000 and the reverse rate becomes 100. Because of the dramatic increase in the rate of the forward reaction made possible by the catalyst, equilibrium is approached in seconds or minutes instead of hours or days.

Enzymes and Macromolecular Crowding Effects

Enzymes operate in living organisms in crowded conditions (p. 39). However, enzyme-catalyzed reactions have been traditionally investigated using dilute, buffered solutions. This strategy is based on the simplifying assumption of an ideal solution. Ideal solutions, for example, contain solutes in such low concentration that interactions such as steric repulsion or attractive forces are nonexistent. Reactions that occur in living organisms, however, deviate from ideality. In such circumstances, equilibrium constants are based not on solute concentrations but on activities, quantities called *effective concentrations* that take intermolecular interactions into account. The effective concentration or *activity* (a) of a solute is equal to

$$a = \gamma c \tag{4}$$

where γ is a correction factor called the *activity coefficient*, which depends on the size and charge of the species and on the ionic strength of the solution in which the species is reacting, and c is the concentration in moles per liter.

The impact of this phenomenon can be considerable. For example, the oxygen-binding capacity of

hemoglobin, the predominant red blood cell protein, differs by several orders of magnitude depending on whether it is measured within red blood cells or in dilute buffer. The equilibrium constant for a reaction under nonideal conditions is given by

$$K^{\circ}_{eq} = \frac{\gamma_B [B]}{\gamma_A [A]} = K^{i}_{eq}\Gamma \quad (5)$$

where K_{eq}^{i} is the ideal constant, and Γ is the nonideality factor, the ratio of the activity coefficients of the products and reactants.

Over the past two decades, it has become increasingly apparent that the assumption of ideal conditions needs to be reevaluated. Consequently, many investigators now use high-molecular-weight "crowding agents" such as dextran (a glucose polymer produced by some bacteria) or serum albumin to simulate intracellular conditions in enzyme studies. Enzyme assays in the presence of crowding agents are closer to those of direct in vivo measurements, but the environment of the assay is still too homogeneous and differs significantly from the crowded heterogeneous conditions in vivo. It is a challenge for today's biochemists to construct assays and models that duplicate in vivo conditions.

Enzyme Specificity

Enzyme specificity is an enzyme property that is partially accounted for by the lock-and-key model, introduced by Emil Fischer in 1890. Each enzyme binds to a single type of substrate because the active site and the substrate have complementary structures. The substrate's overall shape and charge distribution allow it to enter and interact with the enzyme's active site. In a modern variation of the lock-and-key model, Daniel Koshland's *induced-fit model*, the flexible structure of proteins is taken into account (**Figure 6.2**). In this model, substrate does not fit precisely into a rigid active site. Instead, noncovalent interactions between the enzyme and substrate change the three-dimensional structure of the active site, conforming the shape of the active site to the shape of the substrate in its transition state conformation.

Although the catalytic activity of some enzymes depends only on interactions between active site amino acids and the substrate, other enzymes require nonprotein components for their activities. Enzyme **cofactors** may be ions, such as Mg^{2+} or Zn^{2+}, or complex organic molecules, referred to as **coenzymes**. The protein component of an enzyme that lacks an essential cofactor is called an **apoenzyme**. Intact enzymes with their bound cofactors are referred to as **holoenzymes**.

KEY CONCEPTS

- Enzymes are catalysts.
- Catalysts modify the rate of a reaction because they provide an alternative reaction pathway that requires less activation energy than the uncatalyzed reaction.
- Most enzymes are proteins.

The activities of some enzymes can be regulated. Adjustments in the rates of enzyme-catalyzed reactions allow cells to respond effectively to environmental changes. Organisms may control enzyme activities directly, principally through the binding of activators or inhibitors, the covalent modification of enzyme molecules, or indirectly, by regulating enzyme synthesis. (Control of enzyme synthesis requires gene expression changes, a topic covered in Chapters 18 and 19.)

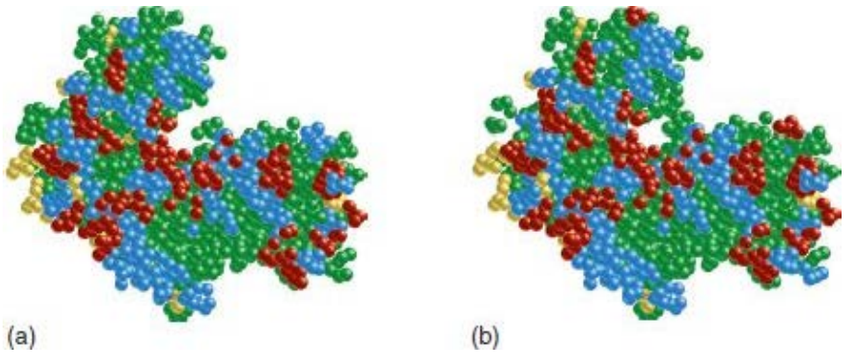


FIGURE 6.2

The Induced-Fit Model

Substrate binding causes enzymes to undergo conformational change. Hexokinase, a single polypeptide with two domains, is shown (a) before and (b) after glucose binding. The domains move relative to each other to close around a glucose molecule (not shown).

QUESTION 6.1

The hexokinases are a class of enzymes that catalyze the ATP-dependent phosphorylation of hexoses (sugars with six carbons). The hexokinases will bind only D-hexose sugars and not their L-counterparts. In general terms, describe the features of enzyme structure that make this specificity possible.

6.2 CLASSIFICATION OF ENZYMES

In the early days of biochemistry, enzymes were named at the whim of their discoverers. Often, enzyme names provided no clue to their function (e.g., trypsin), and sometimes several names were used for the same enzyme. Enzymes were often named by adding the suffix *-ase* to the name of the substrate. For example, urease catalyzes the hydrolysis of urea. To eliminate confusion, the International Union of Biochemistry (IUB) instituted a systematic naming scheme for enzymes. Each enzyme is now classified and named according to the type of chemical reaction it catalyzes. In this scheme, an enzyme is assigned a four-number classification and a two-part name, called a *systematic name*. In addition, a shorter version of the systematic name, called the *recommended name*, is suggested by the IUB for everyday use. For example, alcohol:NAD^+ oxidoreductase (EC 1.1.1.1) is usually referred to as alcohol dehydrogenase. (The letters EC are an abbreviation for the Enzyme Commission of the IUB.) Because many enzymes were discovered before the institution of the systematic nomenclature, the old well-known names have been retained in quite a few cases.

The following are the six major enzyme categories:

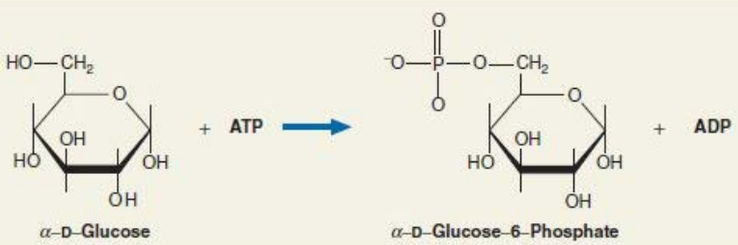
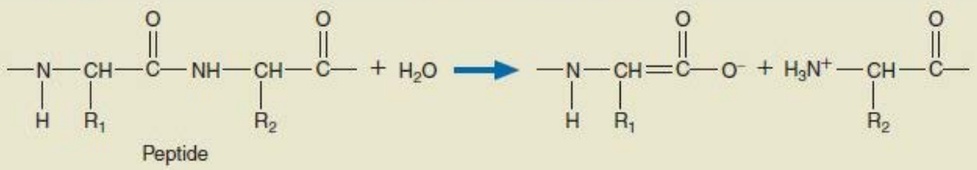
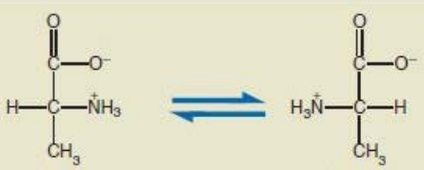
1. **Oxidoreductases.** Oxidoreductases catalyze oxidation-reduction reactions in which the oxidation state of one or more atoms in a molecule is altered. Oxidation-reduction in biological systems involves one- or two-electron transfer reactions accompanied by the compensating change in the amount of hydrogen and oxygen in the molecule. Prominent examples include the redox reactions facilitated by the dehydrogenases and the reductases. For example, alcohol dehydrogenase catalyzes the oxidation of ethanol and other alcohols, and ribonucleotide reductase catalyzes the reduction of ribonucleotides to form deoxyribonucleotides. The oxygenases, oxidases, and peroxidases are among the enzymes that use O_2 as an electron acceptor.
2. **Transferases.** Transferases are enzymes that transfer molecular groups from a donor molecule

to an acceptor molecule. Such groups include amino, carboxyl, carbonyl, methyl, phosphoryl, and acyl (RC=O). Common trivial names for the transferases often include the prefix *trans*; the transcarboxylases, transmethylases, and transaminases are examples.

3. **Hydrolases.** Hydrolases catalyze reactions in which the cleavage of bonds such as C—O, C—N, and O—P is accomplished by the addition of water. The hydrolases include esterases, phosphatases, and proteases.
4. **Lyases.** Lyases catalyze reactions in which groups (e.g., H_2O, CO_2, and NH_3) are removed by elimination to form a double bond or are added to a double bond. Decarboxylases, hydratases, dehydratases, deaminases, and synthases are examples of lyases.
5. **Isomerases.** A heterogeneous group of enzymes, the isomerases catalyze several types of intramolecular rearrangements. The sugar isomerases interconvert *aldoses* (aldehyde-containing sugars) and *ketoses* (ketone-containing sugars). The epimerases catalyze the inversion of asymmetric carbon atoms, and the mutases catalyze the intramolecular transfer of functional groups.
6. **Ligases.** Ligases catalyze bond formation between two substrate molecules. For example, DNA ligase links DNA strand fragments together. The names of many ligases include the term *synthetase*. Several other ligases are called carboxylases.

Table 6.2 presents an example from each enzyme class.

TABLE 6.2 Selected Examples of Enzymes

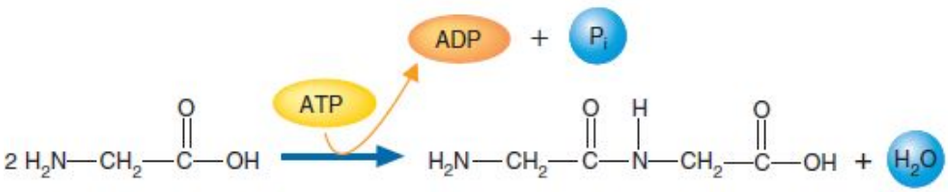
Enzyme Class	Example	Reaction Catalyzed
Oxidoreductase	Alcohol dehydrogenase	
Transferase	Hexokinase	
Hydrolase	Chymotrypsin	
Lyase	Pyruvate decarboxylase	
Isomerase	Alanine racemase	
Ligase	Pyruvate carboxylase	

QUESTION 6.2

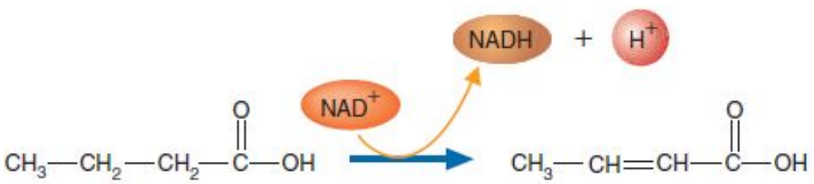
Which type of enzyme catalyzes each of the following reactions?

(a)

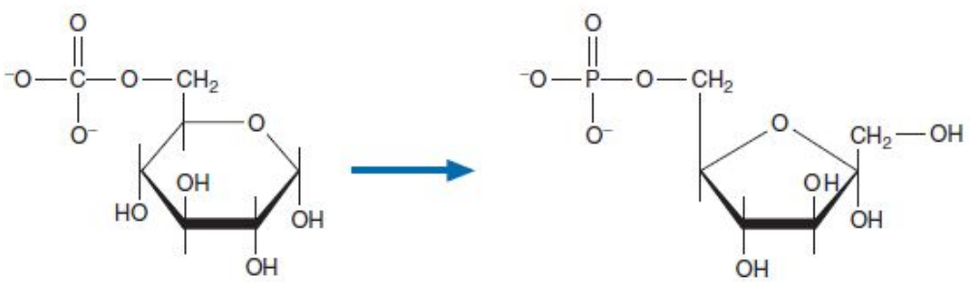
(b)



(c)



(d)

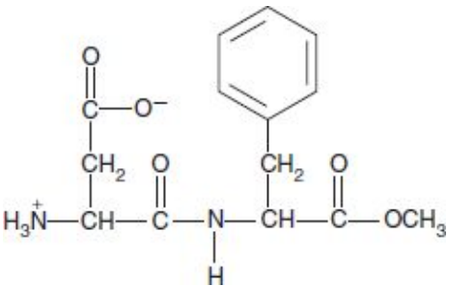


(e)

(f)

QUESTION 6.3

Aspartame, an artificial sweetener, has the following structure:



Once consumed in food or beverages, aspartame is degraded in the digestive tract to its component molecules. Predict what the products of this process are. What classes of enzymes are involved?

6.3 CATALYSIS

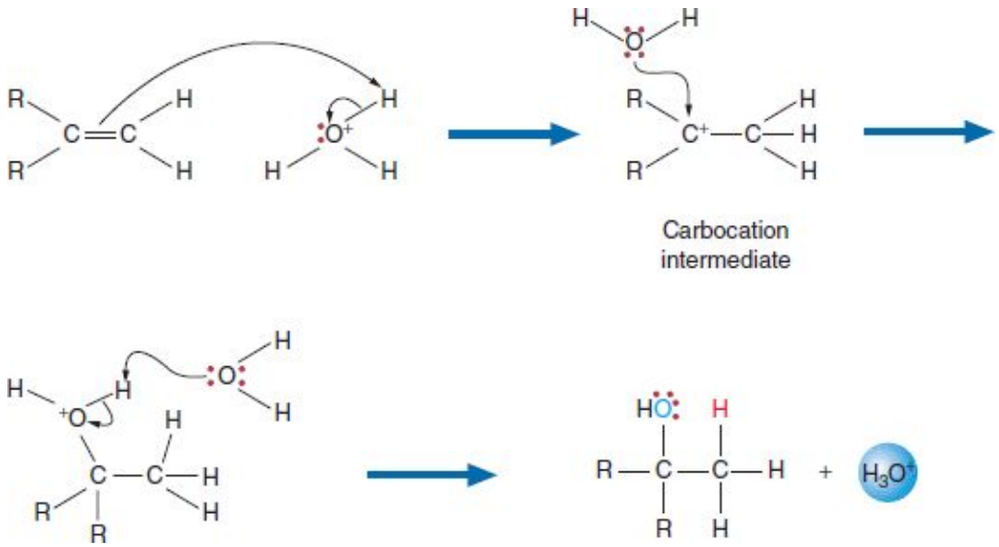
However valuable thermodynamic calculations and kinetic studies (pp. 218–32) are, they do not reveal any information about the actual catalytic mechanisms of enzymes. [A *mechanism* describes how bonds are broken and new ones formed in the conversion of substrate(s) to product(s).] Enzyme mechanism investigations seek to relate enzyme activity to the structure and function of the active site. Scientists who study enzyme mechanisms use such methods as X-ray crystallography, chemical inactivation of active site side chains, and modeling with simple model compounds as substrates and as inhibitors.

Organic Reactions and the Transition State

Biochemical reactions follow the same set of rules as the reactions studied by organic chemists. The essential features of both are the reaction between electron-deficient atoms (electrophiles) and electron-rich atoms (nucleophiles) and the formation of transition states. Each is discussed briefly.

Chemical bonds form when a nucleophile donates an electron pair to an electrophile. For example, in the following reaction

the π electrons of the nucleophilic double bond will react with a partially positive hydrogen atom of an electrophilic hydronium ion. As in all organic reactions, product formation takes place in several steps.



This step-by-step description is referred to as a **reaction mechanism**. The curved arrows illustrate the flow of electrons away from a nucleophile and toward an electrophile. During the course of the reaction, one or more intermediates form. An **intermediate** is a species that exists for a finite length of time (10^{-13}s or less) and then is transformed into product. In the example reaction, a carbocation intermediate forms as π electrons in the double bond attack an electrophilic hydronium ion. (A **carbocation**, or carbonium ion, contains an electron-deficient, positively charged carbon atom.) A carbocation intermediate is stabilized by the R groups that decrease the positive charge on the carbon atom. In the reaction's next step, an electron pair of the oxygen atom in a water molecule forms a σ bond with the positively charged carbon atom. In the final step, the alcohol product results from a proton transfer to another water molecule. Other examples of reactive intermediates observed in biochemical reactions include **carbanions** (nucleophilic carbon anions with three bonds and an unshared electron pair) and **free radicals** (highly reactive species with at least one unpaired electron).

Transition State Stabilization

In any chemical reaction, only molecules that reach the activated condition known as the transition state (see p. 198 and **Figure 6.1**) can convert into product molecules. The conversion of a stable reactant molecule into an activated one is analogous to the energy-requiring process of rolling a boulder up a hill. Once the boulder has reached the top, only a slight push will cause it to slide down the other side, releasing energy as it does so. The rate of a reaction is determined by the relative number of reactant molecules that possess sufficient energy to overcome the activation energy barrier (E_a) (see **Figure 6.1**). Reaction rates increase if E_a can be lowered.

The capacity to lower E_a has been attributed to transition state stabilization. In *transition state theory* it is assumed that the enzyme binds more strongly to the transition state than to the original substrate molecule. Chemical reactions occur in the following manner:

$$E + S \leftrightharpoons ES \leftrightharpoons ETS \rightarrow E + P$$

where ETS = enzyme-transition state complex.

The dissociation constants for the ES complex and the ETS complex are K_S = [E][S]/[ES] and K_T = [E][TS]/[ETS], respectively. The rate accelerations achieved by enzymes require that $K_T < K_S$. In other words, an enzyme stabilizes the transition state by binding it more tightly than to the substrate. Enzymes can bind more effectively to the transition state when the active site is complementary in its shape and electrostatic structure to the transition state's structure.

Triose phosphate isomerase (TPI) is an example of an enzyme whose kinetic and mechanistic properties have been researched in some detail. TPI is an exceptionally efficient enzyme that is composed of two identical subunits. It catalyzes the reversible conversion of dihydroxyacetone phosphate (DHAP) to glyceraldehyde-3-phosphate (GAP), which is a reaction in glycolysis (Chapter 8), the biochemical pathway that degrades the sugar glucose. The enzyme's reaction mechanism involves the formation of an enediol intermediate and two transition states. (An *enediol* is a molecule containing the atomic arrangement —C(OH)

C(OH)—. It is produced by proton migration from the CH of a CHOH group to the oxygen of a carbonyl group to yield a carbon–carbon double bond where each carbon bears an OH group. See p. 254.)

The active site of TPI possesses numerous precisely oriented amino acid side chains that either participate directly in the reaction mechanism (Glu 165 and His 95) or stabilize the substrate, transition states, or the enediol intermediate (**Figure 6.3**). The conversion of DHAP into the enediol involves the deprotonation of C-1 by a nucleophilic glutamate residue and the donation of a proton to the C-2 carbonyl oxygen by an electrophilic histidine residue. Once the enediol is formed, the abstraction of a proton by the C-2 atom from the protonated glutamate residue and the loss of a proton from the C-1 hydroxyl group yield the product GAP. The energy diagram for a two-step reaction in **Figure 6.4** illustrates the difference between an intermediate (a reactive species with a finite lifetime) and a transition state (an unstable species with maximum free energy).

Evidence that enzyme mechanisms involve tight binding of transition states has been acquired with the use of transition state analogues, molecules that so closely resemble a transition state that they bind to an enzyme's active site more tightly than substrate molecules. The molecule 2-phosphoglycolic acid (PGA) resembles TPI's enediolate 1 transition state. PGA was found to bind so tightly to the active site of TPI, most notably to Glu 165 and His 65 side chains, that it functions as a potent inhibitor of the enzyme.

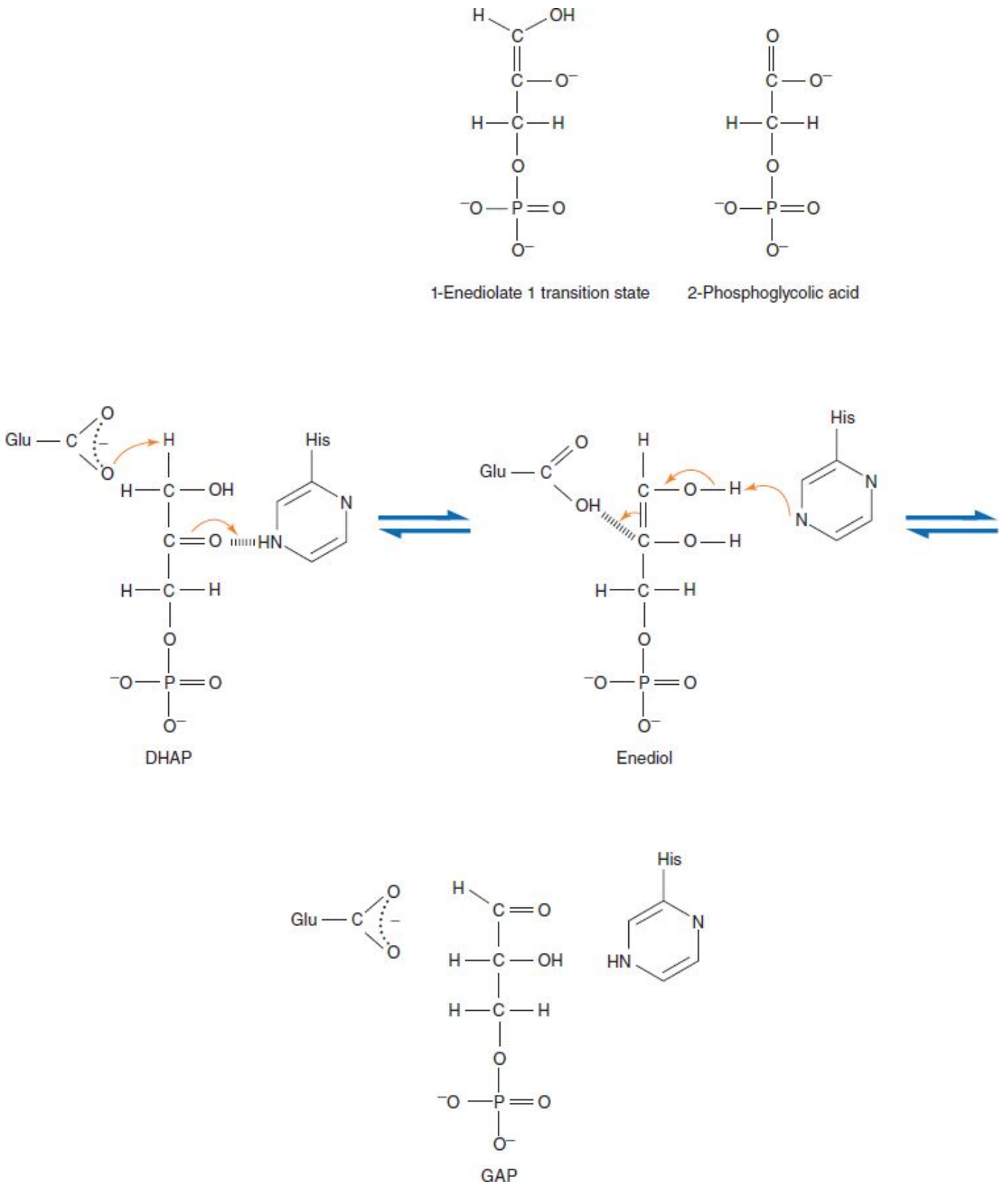


FIGURE 6.3

The Triose Phosphate Isomerase Mechanism

The isomerization reaction in which dihydroxyacetone phosphate (DHAP) is converted to glyceraldehyde-3-phosphate (GAP) is an example of acid–base catalysis that involves a glutamate (Glu 165) and a histidine (His 95) within the active site. The reaction begins when the nucleophilic glutamate carboxyl group abstracts a proton from the substrate and a histidine residue donates a proton to yield the enediol intermediate. The enediol collapses as the histidine removes a proton from the C-1 hydroxyl group and C-2 abstracts a proton from the protonated glutamate to form the GAP product.

3D animation of 2-Phosphoglycolic acid

3D animation of DHAP

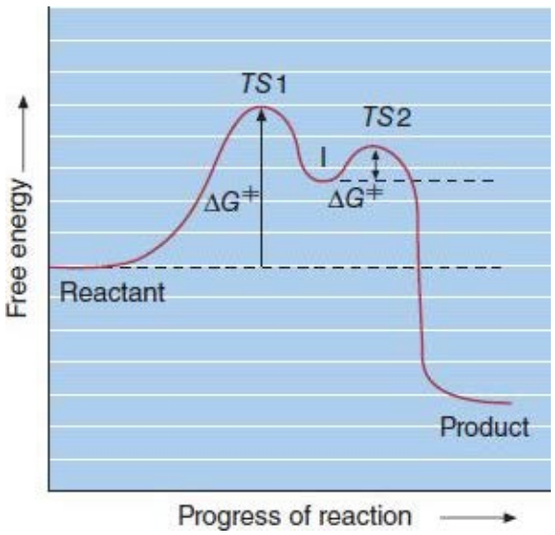


FIGURE 6.4

Energy Diagram for a Two-Step Reaction

There are two transition states (*TS*1 and *TS*2) in a two-step reaction. The overall rate of the reaction is determined by the rate of the step with the highest E_a ($\Delta G^{\ddagger}$), in this case the first step. The reaction intermediate (I), a molecular entity formed from the reactant molecule(s) that reacts further to form the product, exists in the valley between the transition state peaks.

Catalytic Mechanisms

Despite extensive research, the mechanisms of only a few enzymes are known in significant detail. It is known that enzymes achieve significantly higher catalytic rates than other catalysts because their active sites possess structures that are uniquely suited to promote catalysis. Several factors contribute to enzyme catalysis. The most important of these factors are proximity and orientation effects, electrostatic effects, acid–base catalysis, and covalent catalysis. Quantum tunneling, which occurs when hydrogen is transferred, is discussed in an online Biochemistry in Perspective essay. It should be noted that none of these catalytic factors is mutually exclusive. In varying degrees, they all participate in each type of catalytic mechanism.

PROXIMITY AND ORIENTATION EFFECTS For a biochemical reaction to occur, the substrate must come into close proximity to catalytic functional groups (side chain groups involved in a catalytic mechanism) within the active site. This binding, which is possible because of complementary noncovalent bonding interactions between active site functional groups and the substrate, allows the precise orientation of the substrate to the catalytic groups. Reactions occur faster when substrates are advantageously positioned.

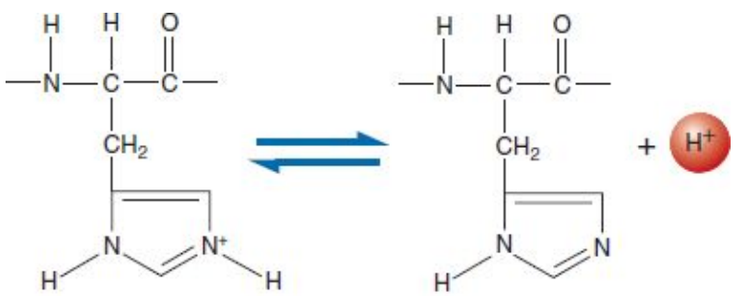
Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on quantum tunneling.

ELECTROSTATIC EFFECTS Recall that the strength of electrostatic interactions is inversely related to the hydration of participating species (Chapter 3). Hydration shells increase the distance between charge centers and reduce electrostatic attraction. The local dielectric constant is often low within active sites because of water exclusion. For example, water exclusion results from a conformational change in TPI that occurs when an enediol intermediate is in the active site. As a result, a lid-like flexible loop in the enzyme moves so as to shield the active site from solvent. The shutting of the lid also traps and stabilizes the intermediate. The charge distribution in the relatively anhydrous active site facilitates the optimum positioning of substrate molecules and influences their chemical reactivity. In addition, weak electrostatic interactions, such as those between permanent and induced dipoles in both the active site and the substrate, are believed to contribute to catalysis.

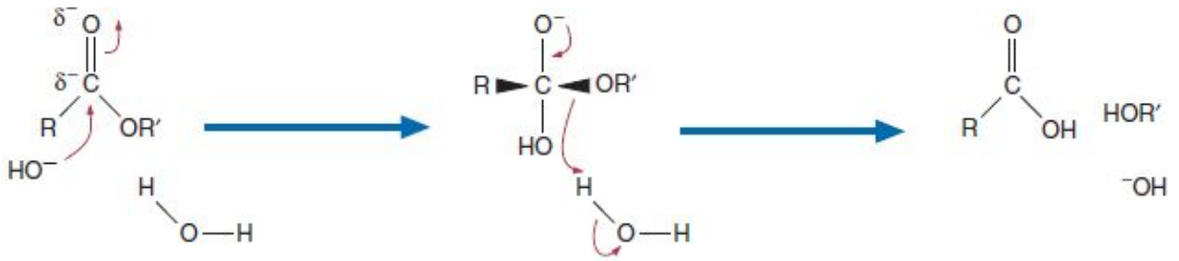
GENERAL ACID–BASE CATALYSIS Acid–base catalysis (proton transfer) is an important factor in chemical reactions. For example, consider the hydrolysis of an ester:

Because water is a weak nucleophile, ester hydrolysis is relatively slow in neutral solution. Ester hydrolysis takes place much more rapidly if the pH is raised. As the hydroxide ion attacks the polarized carbon atom of the carbonyl group (**Figure 6.5a**), a tetrahedral intermediate is formed. As the intermediate breaks down, a proton is transferred from a nearby water molecule. The reaction is complete when the alcohol is released. However, hydroxide ion catalysis is not practical in living systems. Enzymes use several functional groups that behave as general bases to transfer protons efficiently. Such groups can be precisely positioned in relation to the substrate (**Figure 6.5b**). Ester hydrolysis can also be catalyzed by a general acid (**Figure 6.5c**). As the oxygen of the ester's carbonyl group binds to the proton, the carbon atom becomes more electrophilic. The ester then becomes more susceptible to the nucleophilic attack of a water molecule.

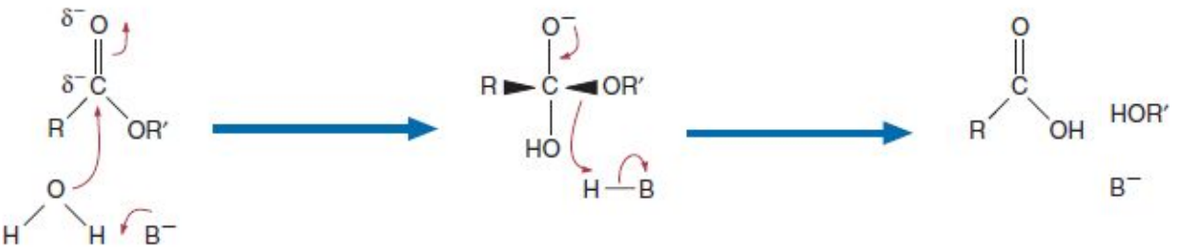
Within enzyme active sites, the functional groups on the side chains of histidine (imidazole), aspartate and glutamate (carboxylate), tyrosine (hydroxyl), cysteine (sulfhydryl), and lysine (amine) can act either as proton donors (called **general acids**) or as proton acceptors (called **general bases**) depending on their state of protonation. Each of these side chain groups has a characteristic pK_a value that may be influenced by nearby charged or polar atoms. For example, the side chain of histidine often participates in concerted acid–base catalysis because its pK_a range is close to physiological pH. The protonated imidazole ring can serve as a general acid, and the deprotonated imidazole ring can serve as a general base:



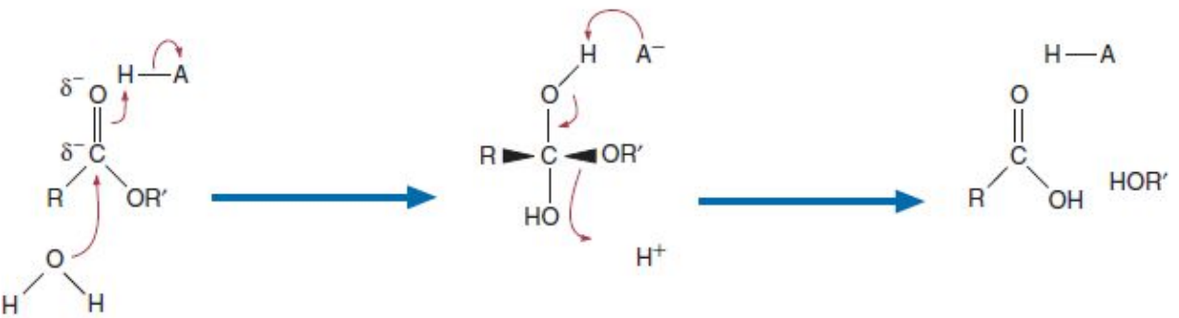
Histidine



(a) Hydroxide ion catalysis



(b) General base catalysis



(c) General acid catalysis

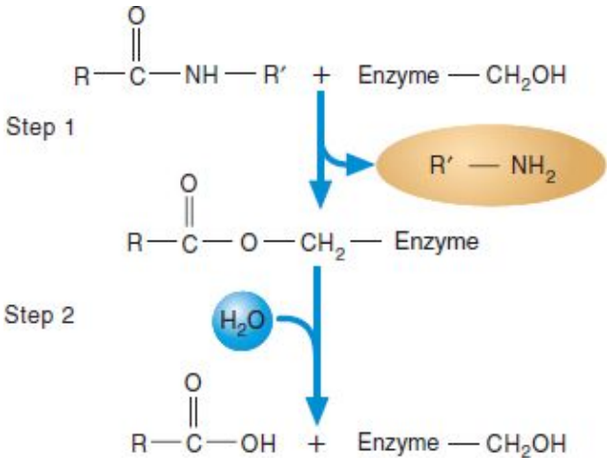
FIGURE 6.5

Ester Hydrolysis

Esters can be hydrolyzed in several ways: (a) catalysis by free hydroxide ion; (b) general base catalysis, in which an amino acid side chain in an enzyme active site accepts a proton; and (c) general acid catalysis, in which an enzyme active site side chain protonates the substrate. A colored arrow represents the movement of an electron pair during each mechanism.

Within the active site of the serine proteases (a class of proteolytic enzymes such as trypsin and chymotrypsin; see pp. 216–17), the close proximity of an aspartate carboxylate group to histidine raises the latter's pK_a. As a consequence, histidine, acting as a general base, abstracts a proton from a nearby serine side chain, thus converting the oxygen of serine into a better nucleophile.

COVALENT CATALYSIS In some enzymes, a nucleophilic side chain group forms an unstable covalent bond with an electrophilic group on the substrate. The serine proteases use the —CH_2—OH group of serine as a nucleophile to hydrolyze peptide bonds. During the first step, the nucleophile (i.e., the oxygen of serine) attacks the carbonyl group of the peptide substrate. As the ester bond is formed, the peptide bond is broken. The resulting acyl-enzyme intermediate is hydrolyzed by water in a second reaction:



Several other amino acid side chains may act as nucleophiles. The sulfhydryl group of cysteine and the carboxylate groups of aspartate and glutamate can play this role.

The Roles of Amino Acids in Enzyme Catalysis

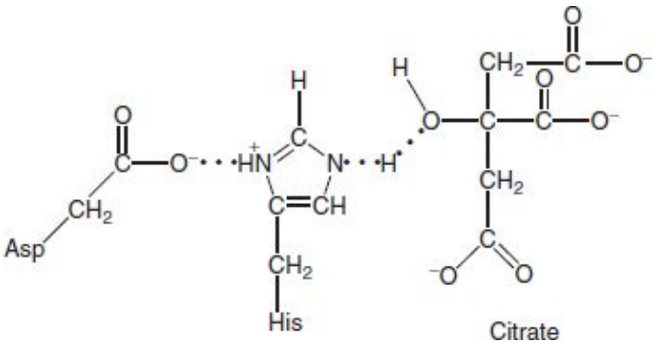
The active sites of enzymes are lined with amino acid side chains that are in close proximity as a result of the protein-folding process. Together these side chains create a microenvironment that is conducive to catalysis. The functions of active site side chains fall into two major categories: catalytic and noncatalytic. Catalytic residues directly participate in the catalytic mechanism, whereas noncatalytic residues have support functions. Of the 20 amino acids found in proteins (**Figure 5.2**), only those with polar and charged side chains actually participate in catalysis. These amino acids (and their side chain groups) are as follows: serine, threonine, and tyrosine (hydroxyl); cysteine (thiol); glutamine and asparagine (amide); glutamate and aspartate (carboxylate); lysine (amine); arginine (guanidinium); and histidine (imidazole).

Extensive research has revealed that catalytic mechanisms require the precise positioning of one or more catalytic units that are composed of either two or three amino acid side chains called dyads or triads, respectively. Although the number of enzymes is vast, the catalytic units are composed of relatively few combinations of amino acids. Commonly observed examples include the arginine–arginine, carboxylate–carboxylate, and carboxylate–histidine dyads.

An arginine–arginine dyad is a catalytic unit in adenylate kinase, an enzyme that catalyzes the transfer of phosphoryl groups from ATP to other nucleotides. The polarizing effect of the two arginines on the phosphate group's oxygens has the effect of converting phosphate into a good leaving group.

Carboxylate–carboxylate dyads occur in the active sites of the aspartic proteases, a family of proteolytic enzymes such as pepsin, which animals use to digest dietary protein. The close proximity of the two negatively charged aspartate carboxyl groups raises the pK_a of one of the aspartates, making it less acidic and more basic. Its enhanced capacity to accept a proton initiates a general acid–general base hydrolytic mechanism.

The functional properties of aspartate–histidine dyads result from a polarized imidazolium ring caused by the close proximity of the aspartate's negatively charged carboxylate group. In the active site of aconitase, the enzyme that catalyzes the isomerization of citrate to form isocitrate (p. 344), histidine acts as a general acid and protonates the −OH group of the citrate, making it a better leaving group.



3D animation of citrate

The resulting HOH is held in the active site by histidine long enough to be able to attack the intermediate and generate the isomer of citrate (isocitrate). An aspartate–histidine dyad is also, most notably, a component of the well-researched serine protease triad (Asp-His-Ser, p. 216). The close proximity of the aspartate carboxylate group to the imidazole group of histidine raises the latter group's pK_a, thus facilitating its ability to remove a proton from serine. (At physiological pH, serine hydroxyl groups are ordinarily poor nucleophiles.) The deprotonated serine is thus converted into a better nucleophile.

The functions of noncatalytic side groups, which include substrate orientation and transition state stabilization, are subtle in comparison with those of catalytic residues. For example, the substrate specificity of chymotrypsin (pp. 216–17), manifested in the cleavage of peptide bonds on the C-terminal side of the aromatic amino acids tryptophan, tyrosine, and phenylalanine, is made possible by the relatively large size of a hydrophobic pocket within the active site that both accommodates and orients the aromatic side chains. The **oxyanion** intermediate (a molecular species with a negatively charged oxygen: see p. 216 and **Figure 6.9**) that forms during the catalytic mechanism is stabilized by interactions between the substrate's peptide bond carbonyl group and the backbone amide hydrogens of serine and glycine residues.

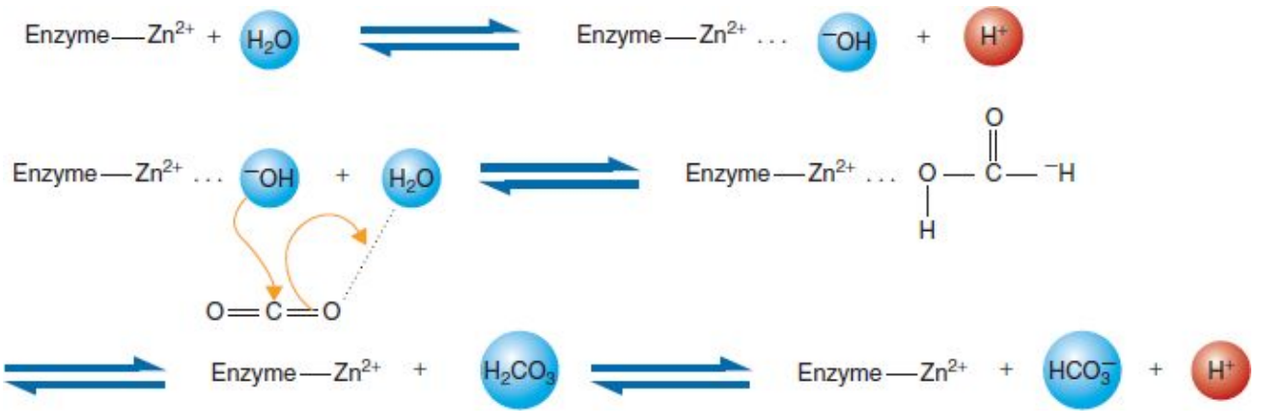
The Role of Cofactors in Enzyme Catalysis

In addition to active site amino acid side chains, many enzymes require nonprotein cofactors, that is, metal cations and the coenzymes. Each group has distinctive structural properties and chemical reactivities.

METALS The important metals in living organisms fall into two classes: the alkali and alkaline earth metals (e.g., Na^+, K^+, Mg^{2+}, and Ca^{2+}) and the transition metals (e.g., Zn^{2+}, Fe^{2+}, and Cu^{2+}). The alkali and alkaline earth metals in enzymes are loosely bound and usually have structural roles. In contrast, the transition metals play key roles in catalysis either bound to functional groups such as carboxylate, imidazole, or hydroxyl groups, or as components of prosthetic groups such as Fe^{2+} in heme.

Several properties of transition metals make them useful in catalysis. Metal ions provide a high concentration of positive charge that is especially useful in binding small molecules. Because transition metals act as *Lewis acids* (electron pair acceptors, see p. P-19), they are effective electrophiles. (Amino acid side chains are poor electrophiles because they cannot accept unshared pairs of electrons.) Because the directed d shell valences of metals allow them to interact with two or more ligands, metal ions help orient the substrate within the active site. As a consequence, the enzyme–metal ion complex polarizes the substrate and promotes catalysis. For example, carbonic

anhydrase (**Figure 6.6a**) is the enzyme that catalyzes the reversible hydration of CO_2 to form bicarbonate (HCO_3^-). Its active site contains a zinc (Zn^{2+}) cofactor that is coordinated with three histidine side chains (**Figure 6.6b**). The zinc ion polarizes a water molecule, resulting in a Zn^{2+}-bound OH group. The OH group (acting as a nucleophile) attacks CO_2, converting it into HCO_3^-:



Finally, because transition metals have two or more valence states, they can mediate oxidation-reduction reactions by reversibly gaining or losing electrons. For example, the reversible oxidation of Fe^{2+} to form Fe^{3+} is important in the function of cytochrome P_{450}, a type of enzyme that processes toxic substances in animals. [See the Biochemistry in Perspective essays in Chapters 10 (Myocardial Infarct: Ischemia and Reperfusion) and 12 (Biotransformation).]

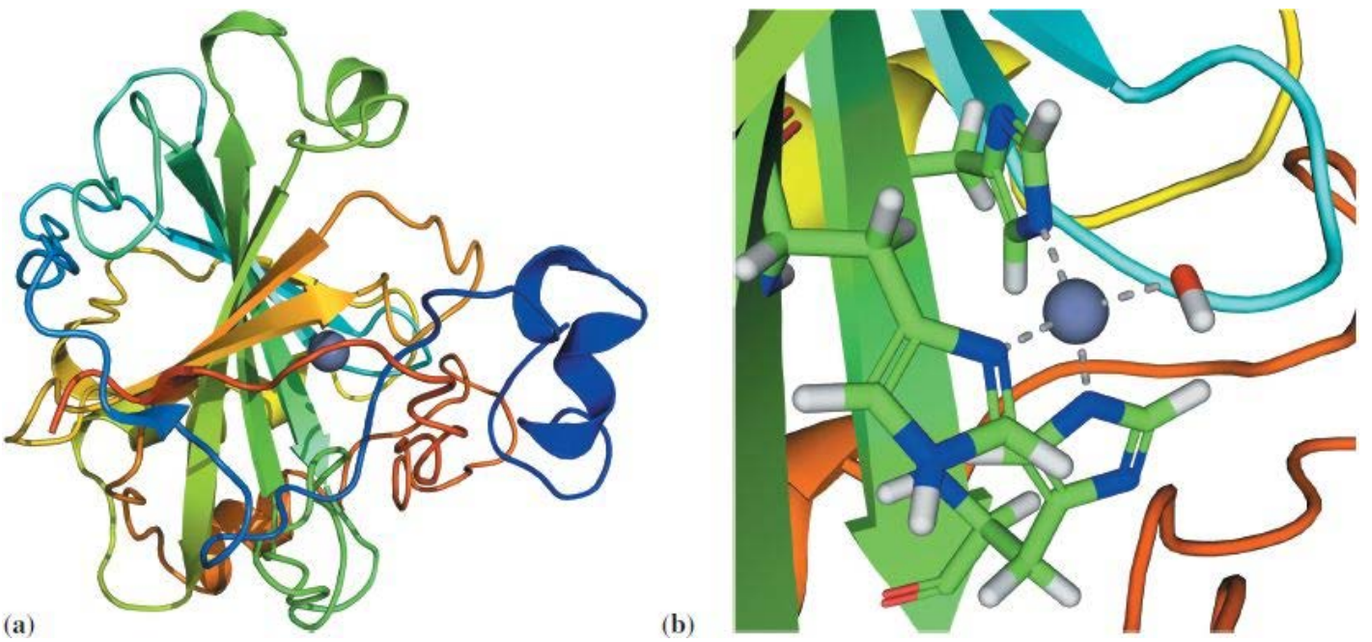


FIGURE 6.6

Structure of α-Carbonic Anhydrase

(a) Found predominantly in animals, the α-carbonic anhydrases (CAs) catalyze the reversible hydration of CO_2. CAs are usually zinc-dependent metalloenzymes with a deep cavity in which the active site is composed in part by an antiparallel β-sheet. (b) The zinc ion of carbonic anhydrase is located at the bottom of a cone-shaped cavity, where it is coordinated with three histidyl residues. In the absence of substrate molecules, the zinc ion is also coordinated with a water molecule. In this illustration, the zinc ion is shown bound to a hydroxide group, the product of the acceptance of a proton from a water molecule by a fourth histidyl residue (not shown). The hydroxide group will subsequently attack a CO_2 molecule.

Menkes' Syndrome

QUESTION 6.4

Copper is a cofactor in several enzymes, including lysyl oxidase and superoxide dismutase. Ceruloplasmin, a deep-blue glycoprotein, is the principal copper-containing protein in blood. It is used to transport Cu^{2+} and maintain appropriate levels of Cu^{2+} in the body's tissues. Ceruloplasmin also catalyzes the oxidation of Fe^{2+} to Fe^{3+}, an important reaction in iron metabolism. Because the metal is widely found in foods, copper deficiency is rare in humans. Deficiency symptoms include anemia, leukopenia (reduction in blood levels of white blood cells), bone defects, and weakened arterial walls. The body is partially protected from exposure to excessive copper (and several other metals) by metallothionein, a small, metal-binding protein that possesses a large proportion of cysteine residues. Certain metals (most notably zinc and cadmium) induce the synthesis of metallothionein in the intestine and liver.

In *Menkes syndrome* intestinal absorption of copper is defective. How can affected infants be treated to avoid the symptoms of the disorder, which include seizures, retarded growth, and brittle hair?

Wilson's Disease

QUESTION 6.5

In a rare inherited disorder, called *Wilson's disease*, excessive amounts of copper accumulate in liver and brain tissue. A prominent symptom of the disease is the deposition of copper in greenish-brown layers surrounding the cornea, called Kayser–Fleischer rings. A defective ATP-dependent protein that transports copper across cell membranes causes Wilson's disease. Apparently, the copper transport protein is required to incorporate copper into ceruloplasmin and to excrete excess copper. In addition to a diet low in copper, Wilson's disease is treated with zinc sulfate and the chelating agent penicillamine (p. 148). Describe how these treatments work. [*Hint:* Metallothionein has a greater affinity for copper than for zinc.]

COENZYMES Coenzymes are organic molecules that provide enzymes with chemical versatility because they either possess reactive groups not found on amino acid side chains or can act as carriers for substrate molecules. Some coenzymes are only transiently bound to the enzyme and are essentially cosubstrates, whereas others are tightly bound by covalent or noncovalent bonds. Unlike ordinary catalysts, coenzyme structures are changed by the reactions in which they participate. Their catalytically active forms must be regenerated before another catalytic cycle can occur. Transiently bound coenzymes are usually regenerated by a reaction catalyzed by another enzyme, whereas tightly bound coenzymes are regenerated during a step in the catalytic cycle. Most coenzymes are derived from vitamins. **Vitamins** (organic nutrients required in small amounts in the human diet) are divided into two classes: water-soluble and lipid-soluble. In addition, there are certain vitamin-like

substances that organisms can synthesize in sufficient amounts to facilitate enzyme-catalyzed reactions. Examples include lipoic acid, carnitine, coenzyme Q, biopterin, *S*-adenosylmethionine, and *p*-aminobenzoic acid.

Vitamins

Coenzymes can be classified according to function into three groups: electron transfer, group transfer, and high-energy transfer potential. Coenzymes involved in redox (electron or hydrogen transfer) reactions include nicotinamide adenine dinucleotide (NAD^+), nicotinamide adenine dinucleotide phosphate ($NADP^+$), flavin adenine dinucleotide (FAD), flavin mononucleotide (FMN), coenzyme Q (CoQ), and tetrahydrobiopterin (BH_4). Coenzymes such as thiamine pyrophosphate (TPP), coenzyme A (CoASH), and pyridoxal phosphate are involved in the transfer of aldehyde, acyl, and amino groups, respectively. One-carbon transfers, a diverse set of reactions in which carbon atoms are transferred in various oxidation states between substrates, require biotin, tetrahydrofolate (TH_4) or *S*- adenosylmethionine (SAM). Nucleotides, known for their high-energy transfer potential, function as coenzymes in that they activate metabolic intermediates and/or serve as phosphate donors or as carriers for small molecules. Examples of the latter include UDP-glucose, an intermediate in glycogen synthesis (p. 316), and CDP (cytidine diphosphate)-ethanolamine, an intermediate in the synthesis of certain lipids (p. 472). In such cases, the nucleotide serves as a molecular carrier and a good leaving group in subsequent reactions. The numerous noncovalent bonds that form when the nucleotide binds within the enzyme's active site assure that the attached metabolite is correctly positioned. Note that the structures of coenzymes such as NAD^+, $NADP^+$, FAD, FMN, and CoASH also contain adenine nucleotide components. **Table 6.3** lists the vitamins and vitamin-like substances, their coenzyme forms, and the reactions they facilitate, along with the pages on which their structural and functional properties are described.

KEY CONCEPTS

- The amino acid side chains in the active site of enzymes catalyze proton transfers and nucleophilic substitutions. Other reactions require nonprotein cofactors, that is, metal cations and the coenzymes.
- Metal ions are effective electrophiles, and they help orient the substrate within the active site. In addition, certain metal cations mediate redox reactions.
- Coenzymes are organic molecules that have a variety of functions in enzyme catalysis.

QUESTION 6.6

Identify each of the following as a cofactor, coenzyme, apoenzyme, holoenzyme, or none of these. Explain each answer.

a. Zn^{2+}
b. NAD^+
c. active alcohol dehydrogenase
d. CDP-ethanolamine
e. alcohol dehydrogenase lacking Zn^{2+}
f. biotin

g. FMN

TABLE 6.3 Vitamins, Vitamin-like Molecules, and Their Coenzyme Forms

Molecules	Coenzyme Form	Reaction or Process Promoted	See Page
Water-Soluble Vitamins			
Thiamine (B_1)	Thiamine pyrophosphate	Decarboxylation, aldehyde group transfer	342
Riboflavin (B_2)	FAD and FMN	Redox	335
Pyridoxine (B_6)	Pyridoxal phosphate	Amino group transfer	534
Nicotinic acid (niacin)	NAD and NADP	Redox	335
Pantothenic acid	Coenzyme A	Acyl transfer	341
Biotin	Biotin	Carboxylation	461
Folic acid	Tetrahydrofolic acid	One-carbon group transfer	545
Vitamin B_{12}	Deoxyadenosylcobalamin, methylcobalamin	Intramolecular rearrangements	545
Ascorbic acid (vitamin C)	Unknown	Hydroxylation	393
Lipid-Soluble Vitamins			
Vitamin A	Retinal	Vision, growth, and reproduction	392
Vitamin D	1, 25-Dihydroxycholecalciferol	Calcium and phosphate metabolism	418
Vitamin E	Unknown	Lipid antioxidant	392
Vitamin K	Unknown	Blood clotting	417
Vitamin-like Molecules			
Coenzyme Q	Coenzyme Q	Redox	364
Biopterin	Tetrahydrobiopterin	Redox	543
S-adenosylmethionine	*S*-adenosylmethionine	Methylation	546
Lipoic acid	Lipoamide	Redox	342

Effects of Temperature and pH on Enzyme-Catalyzed Reactions

Any environmental factor that disturbs protein structure may change enzymatic activity. Enzymes are especially sensitive to changes in temperature and pH.

TEMPERATURE All chemical reactions are affected by temperature. In general, as temperature increases reaction rates increase; that is, the number of collisions increases, and more molecules have sufficient energy to enter into the transition state. The rates of enzyme-catalyzed reactions also increase with increasing temperature. However, enzymes are proteins that become denatured at high temperatures. An enzyme's *optimum temperature*, the temperature at which it operates at maximum efficiency (**Figure 6.7**), is determined in the laboratory under specified conditions of pH, ionic strength, and solute concentrations. In living organisms there is no single definable optimal temperature for enzymes.

PH Hydrogen ion concentration, expressed as a pH value, affects enzymes in several ways. First, catalytic activity is related to the ionic state of the active site. Changes in hydrogen ion concentration can affect the ionization of active site groups (p. 101). For example, the catalytic activity of a certain enzyme requires the protonated form of an active site side chain amino group. If the pH becomes so alkaline that the group loses its proton, the enzyme's activity may be depressed. In addition, substrates may be affected. If a substrate contains an ionizable group, a change in pH may alter its capacity to bind to the active site. Second, changes in ionizable groups may change the tertiary structure of the enzyme. Drastic changes in pH often lead to denaturation.

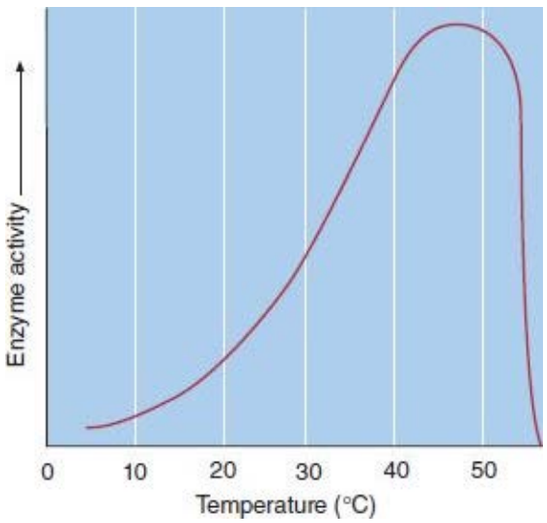


FIGURE 6.7

Effect of Temperature on Enzyme Activity

Modest increases in temperature increase the rate of enzyme-catalyzed reactions because of an increase in the number of collisions between enzyme and substrate. Eventually, increasing the temperature decreases the reaction velocity. Catalytic activity is lost when heat denatures the enzyme.

Although a few enzymes tolerate large changes in pH, most enzymes are active only within a narrow pH range. For this reason, living organisms employ buffers to closely regulate pH. The pH value at which an enzyme's activity is maximal is called the **pH optimum** (**Figure 6.8**). The pH optima of enzymes, determined in the laboratory, vary considerably. For example, the optimum pH of pepsin, a proteolytic enzyme produced in the chief cells of the stomach, is approximately 2. The stomach's low pH is the result of the secretion of hydrochloric acid (HCl) by parietal cells. The optimum pH for the proteolytic enzyme chymotrypsin is 8. Produced in the pancreas as a larger molecule (p. 233), chymotrypsin is one of several digestive enzymes that are secreted into the small intestine as a component of pancreatic juice. Pancreatic juice also contains sodium bicarbonate ($NaHCO_3$), which neutralizes the acidity of the partially digested food entering the small intestine from the stomach, thereby raising its pH to about 8.

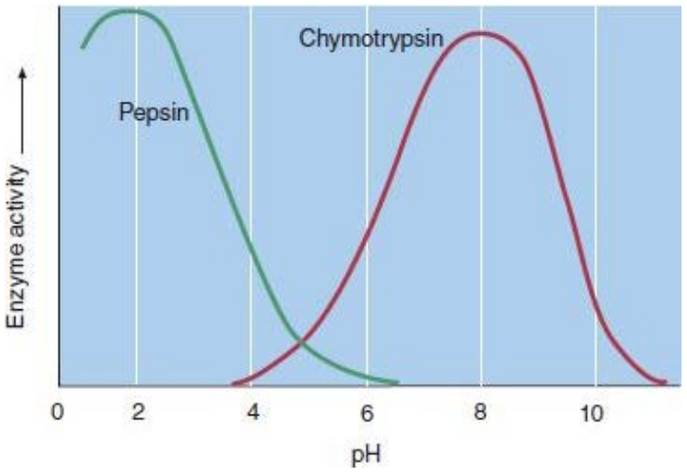


FIGURE 6.8

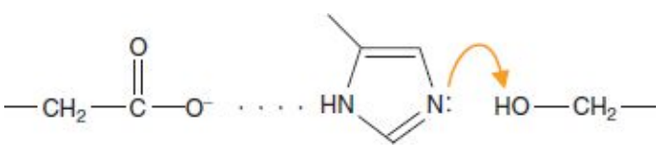
Effect of pH on Two Enzymes

Each enzyme has a certain pH at which it is most active. A change in pH can alter the ionizable groups within the active site or affect the enzyme's conformation.

Detailed Mechanisms of Enzyme Catalysis

Each of the thousands of enzymes that have been investigated has a unique structure, substrate specificity, and reaction mechanism. The mechanisms of a variety of enzymes have been investigated intensively over the past several decades. The subsections that follow describe the catalytic mechanisms of two well-characterized enzymes.

CHYMOTRYPSIN Chymotrypsin is a 27,000 Da protein that belongs to the serine proteases. The active sites of all serine proteases contain a characteristic set of amino acid residues, often referred to as the serine protease triad (p. 210). In the chymotrypsin numbering system, these are Asp 102, His 57, and Ser 195. Studies of crystallized enzyme bound to substrate analogues reveal that these residues are close to each other in the active site. The active site serine residue, one component of a catalytic triad, plays an especially important role in the catalytic mechanisms of this group of enzymes. Serine proteases are irreversibly inhibited by diisopropylfluorophosphate (DFP). In DFP-inhibited enzymes, the inhibitor is covalently bound only to Ser 195 and not to any of the other 29 serines. The special reactivity of Ser 195 is attributed to the proximity of His 57 and Asp 102. The imidazole ring of His 57 lies between the carboxyl group of Asp 102 and the —CH_2OH group of Ser 195. The carboxyl group of Asp 102 polarizes His 57, thus allowing it to act as a general base (i.e., the abstraction of a proton by the imidazole group is facilitated):



Removing the proton from the serine OH group converts it into a more effective nucleophile.

Chymotrypsin catalyzes the hydrolysis of peptide bonds adjacent to aromatic amino acids. The probable mechanism for this reaction is illustrated in **Figure 6.9**. Step (a) of the figure shows the initial enzyme–substrate complex. The alignment of the amino acid residues within the active site, including the catalytic triad of Asp 102, His 57, and Ser 195, creates a temporarily unoccupied position called the oxyanion hole. In addition, the active site provides a hydrophobic binding pocket for the substrate's phenylalanine side chain. The nucleophilic hydroxyl oxygen of Ser 195 launches a

nucleophilic attack on the amide carbonyl carbon of the substrate. The tetrahedral sp intermediate that forms (step b) from the planar sp^3 hybridized amide carbon is sufficiently distorted that the newly created oxyanion is stabilized by hydrogen bonds to the amide hydrogens of Ser 195 and Gly 193 and enters the oxyanion hole. It is believed that the formation of the tetrahedral intermediate, which resembles the transition state, and its preferential binding by the enzyme, is responsible for the catalytic efficiency of the serine proteases.

The tetrahedral intermediate subsequently decomposes to form the covalently bound acyl-enzyme intermediate (step c). The residue His 57, acting as a general acid, is believed to facilitate this decomposition. In steps (d) and (e), the two previous steps are reversed. With water acting as an attacking nucleophile, a tetrahedral (oxyanion) intermediate is formed. By step (f) (the final enzyme-product complex), the bond between the serine oxygen and the carbonyl carbon has been broken. Serine is again hydrogen-bonded to His 57.

ALCOHOL DEHYDROGENASE The alcohol dehydrogenases (ADHs) (**Figure 6.10**) are a diverse group of enzymes that are distributed throughout the three domains of living organisms (Bacteria, Archaea, and Eukarya). They catalyze the reversible oxidation of alcohols to form aldehydes or ketones. In the following reaction, the oxidation of ethanol, two electrons and two protons are removed from the alcohol molecule. The coenzyme NAD^+ acts as a hydride ion ($H:^-$) acceptor.

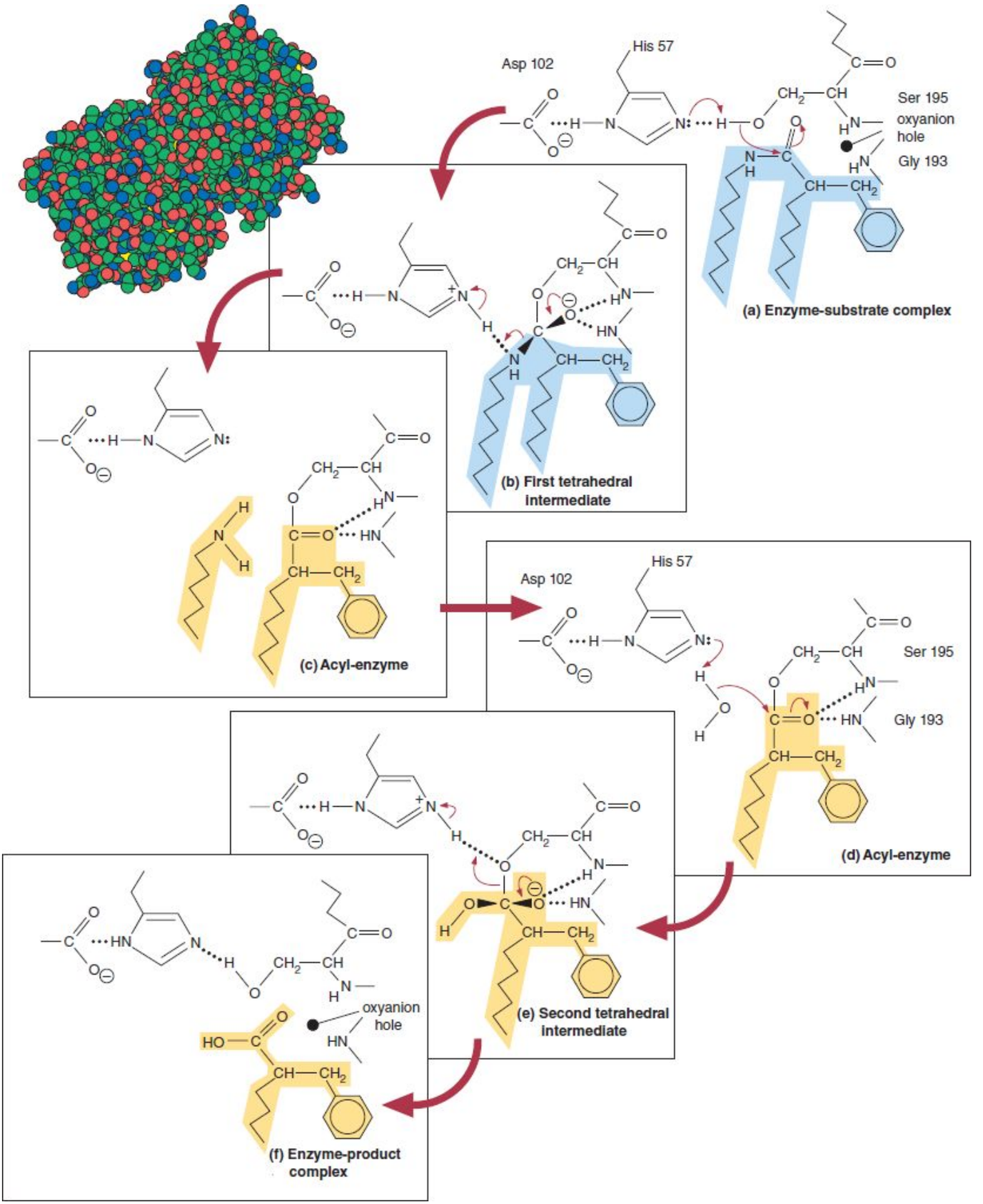


FIGURE 6.9

The Probable Mechanism of Action of Chymotrypsin

There is a fast acylation step during which the carbonyl end of the target peptide bond of the substrate is transferred to the enzyme to form an acyl-enzyme adduct and the amino end of the substrate leaves the active site. A second slow deacylation of the enzyme follows, releasing the carbonyl end of the substrate.

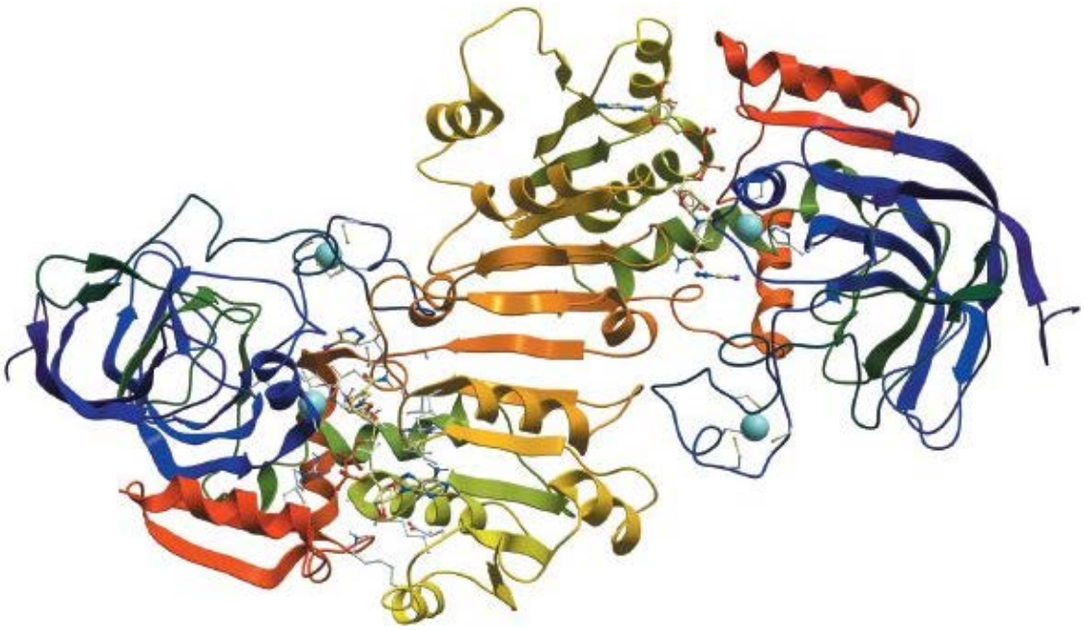


FIGURE 6.10

Human Alcohol Dehydrogenase

In humans, the alcohol dehydrogenases are oxidoreductases that function as dimers with a zinc ion complexed within the active site in each subunit. The subunits are aligned antiparallel to each other. When the dimer forms, the two NAD-binding domains, each of which contains β-strands, form a 12-stranded β-pleated sheet.

Most ADHs consist of two or four subunits, each of which contains two zinc ions. One zinc ion, located in the active site, is crucial for catalysis, whereas the other has a structural function that promotes the enzyme's stability. The active site of alcohol dehydrogenase also contains two cysteine residues (Cys 48 and Cys 174) and a histidine residue (His 67), all of which are coordinated to a zinc ion (**Figure 6.11a**). After NAD^+ binds to the active site, the substrate ethanol enters and binds to the Zn^{2+} as the alcoholate anion (**Figure 6.11b**). The electrostatic effect of Zn^{2+} stabilizes the transition state. As the intermediate decomposes, the hydride ion is transferred from the substrate to the nicotinamide ring of NAD^+. After the aldehyde product is released from the active site, NADH also dissociates.

KEY CONCEPTS

- Each enzyme has a unique structure, substrate specificity, and reaction mechanism.
- Each mechanism is affected by catalysis-promoting factors that are determined by the structure of the substrate and the enzyme's active site.

6.4 ENZYME KINETICS

Recall from Chapter 4 (p. 119) that the value of ΔG for a specific reaction can predict whether it is spontaneous but cannot be used to predict its rate. The rate or **velocity** of a biochemical reaction is defined as the change in the concentration of a reactant or product per unit time. The initial velocity v_0 of the reaction A $\rightarrow$ P, where A and P are substrate and product molecules, respectively, is

$$v_0 = \frac{-\Delta[\mathrm{A}]}{\Delta t} = \frac{\Delta[\mathrm{P}]}{\Delta t} \tag{6}$$

where [A] = concentration of substrate
[P] = concentration of product
t = time

Initial velocity (v_0) is the velocity of a reaction when [A] greatly exceeds the concentration of enzyme E [E], and the reaction time is very short. Measurements of v_0 are made immediately after the mixing of enzyme and substrate because it can be assumed that the reverse reaction (i.e., conversion of product into substrate) has not yet occurred to any appreciable extent.

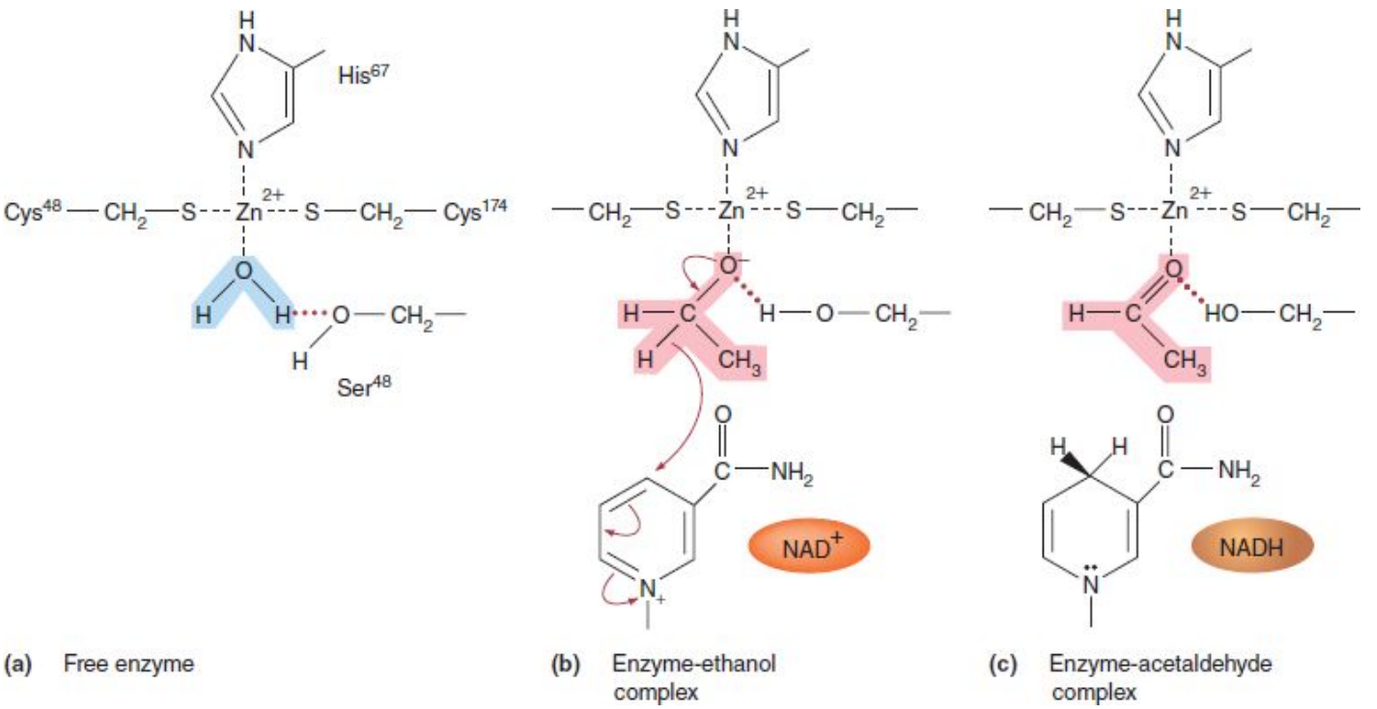


FIGURE 6.11

Functional Groups of the Active Site of Alcohol Dehydrogenase

(a) Without a substrate, a molecule of water is one of the ligands of the Zn^{2+} ion.
(b) The substrate ethanol probably binds to the Zn^{2+} as the alcoholate anion, displacing the water molecule. (c) NAD^+ accepts a hydride ion from the substrate, and the aldehyde product is formed.

The quantitative study of enzyme catalysis, referred to as **enzyme kinetics**, provides information about reaction rates. Kinetic studies also measure the affinity of enzymes for substrates and inhibitors and provide insight into reaction mechanisms. Enzyme kinetics provides insight into the forces that regulate metabolic pathways. The rate of the reaction A → P is proportional to the frequency with which the reacting molecules form product. The reaction rate is

$$v_0 = k[\mathrm{A}]^x \tag{7}$$

where v_0 = initial rate

k = a rate constant that depends on the reaction conditions (e.g., temperature, pH, and ionic strength)

x = the order of the reaction

Combining Equations (6) and (7), we have

$$\frac{\Delta [A]}{\Delta t} = k\,[A]^x \tag{8}$$

Order, defined as the sum of the exponents on the concentration terms in the rate expression, is determined empirically, that is, by experimentation (**Figure 6.12**). Determining the order of a reaction allows an experimenter to draw certain conclusions regarding the reaction's mechanism. A reaction is said to follow *first-order kinetics* if the rate depends on the first power of the concentration of a single reactant and suggests that the rate-limiting step is a unimolecular reaction (i.e., no molecular collisions are required). In such a reaction (A → P), it is assumed that

$$\text{Rate} = k[A]^1 \tag{9}$$

If [A] is doubled, the rate is observed to double. Reducing [A] by half results in halving the observed reaction rate. In first-order reactions, the concentration of the reactant is a function of time, so k is expressed in units of s^{-1}. In any reaction, the time required for one-half of the reactant molecules to be consumed is called a *half-life* ($t_{1/2}$).

In the reaction A + B → P, if the order of A and B is 1 each, then the reaction is said to be *second-order*, and A and B must collide for product to form (a biomolecular reaction):

$$\text{Rate} = k[A]^1[B]^1 \tag{10}$$

In this circumstance, the reaction rate depends on the concentrations of the two reactants. In other words, both A and B take part in the reaction's rate-determining step. Second-order rate constants are measured in units of $M^{-1}s^{-1}$.

Sometimes second-order reactions involve reactants such as water that are present in great excess:

$$A + H_2O \rightarrow P$$

The second-order rate expression is

$$\text{Rate} = k[A]^1\,[H_2O]^1 \tag{11}$$

Because water is present in excess and $[H_2O]$ is essentially constant, however, the reaction appears to be first-order. Such reactions are said to be *pseudo-first-order*. Hydrolysis reactions in biochemical systems are assumed to be pseudo-first-order because of the ready availability of the second reactant, H_2O, in aqueous environments.

Another possibility is that only one of the two reactants is involved in the rate-determining step and it alone appears in the rate expression. For the above reaction, if rate = $k[A]^2$, then the rate-limiting step involves collisions between A molecules. The water is involved in a fast, non-rate-limiting step in the reaction mechanism.

When the addition of a reactant does not alter a reaction rate, the reaction is said to be *zero-order* for that reactant. For the reaction A → P, the experimentally determined rate expression under such conditions is

$$\text{Rate} = k[A]^0 = k \tag{12}$$

The rate is constant because the reactant concentration is high enough to saturate all the catalytic sites on the enzyme molecules. An example of an order determination is given in Problem 6.1.

Reaction order can also be characterized in another way. A theoretical term can be used to

characterize simple reactions: *molecularity* is defined as the number of colliding molecules in a single-step reaction. A *unimolecular* reaction A → B has a molecularity of one, whereas the *bimolecular* reaction A + B → C + D has a molecularity of two.

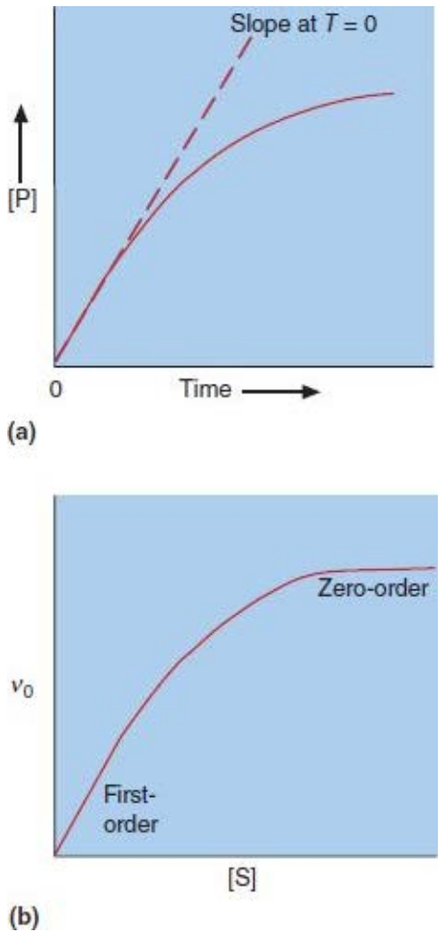


FIGURE 6.12

Enzyme Kinetic Studies

(a) Conversion of substrate to product per unit time. The slope of the curve at $t = 0$ equals the initial rate of the reaction. (b) Plot of initial velocity v versus substrate concentration [S]. The rate of the reaction is directly proportional to substrate concentration only when [S] is low. When [S] becomes high enough to saturate the enzyme, the rate of the reaction is zero-order with respect to substrate. At intermediate substrate concentrations, the reaction has a mixed order (i.e., the effect of substrate on reaction velocity is in transition).

Michaelis–Menten Kinetics

In 1913, Leonor Michaelis and Maud Menten proposed one of the most useful models in the systematic investigation of enzyme rates. The concept of the enzyme–substrate complex, first enunciated by Victor Henri in 1903, is central to Michaelis–Menten kinetics. When the substrate S binds in the active site of an enzyme E, an intermediate complex (ES) is formed. ES complex formation lowers the energy of the transition state and facilitates the product formation state. After a brief time, the product dissociates from the enzyme. This process can be summarized as follows:

KEY CONCEPTS

- Enzyme kinetics is the quantitative study of enzyme catalysis.
- Kinetic studies measure reaction rates and the affinity of enzymes for substrates and inhibitors.

- Kinetics also provides insight into reaction mechanisms.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \tag{13}$$

where k_1 = rate constant for ES formation

k_{-1} rate constant for ES dissociation

$= k_2$ rate constant for product formation and release from the active site

=

WORKED PROBLEM 6.1

Consider the following reaction:

Given the following rate data, determine the order in each reactant and the overall order of the reaction.

Initial Concentrations (mol/L)

Ethanol	NAD^+	Rate (mmol/s)
0.1	0.1	1×10^2
0.2	0.1	2×10^2
0.1	0.2	2×10^2
0.2	0.2	4×10^2

SOLUTION

The overall initial rate expression is

$$\text{Rate} = k[\text{ethanol}]^x[NAD^+]^y$$

To evaluate x and y, determine the effect on the rate of the reaction of increasing the concentration of one reactant while keeping the concentration of the other constant.

For this experiment, doubling the concentration of ethanol doubles the rate of the reaction; therefore, x is 1. Doubling the concentration of NAD^+ doubles the rate of the reaction. So y is also 1. The rate expression then is

$$\text{Rate} = k[\text{ethanol}]^1[NAD^+]^1$$

The reaction is first-order in both reactants and second-order overall. ■

Equation (13) ignores the reversibility of the step in which the ES complex is converted into enzyme and product. This simplifying assumption is allowed if the reaction rate is measured while [P] is still very low. Recall that initial velocities are measured in most kinetic studies. In addition, many enzymes have little affinity for the product, so the reverse reaction is less possible.

According to the Michaelis–Menten model, as currently conceived, it is assumed that (1) k_{-1} is negligible compared with k_1 and (2) the rate of formation of ES is equal to the rate of its degradation over most of the course of the reaction (i.e., the [ES] remains the same throughout the reaction). The second premise is referred to as the *steady state assumption*. The general expression for the velocity of the reaction is

$$\text{Rate} = \frac{\Delta P}{\Delta t} = k_2 [ES] \tag{14}$$

To be useful, a reaction rate must be defined in terms of [S] and [E]. The rate of formation of ES is equal to k_1[E][S], whereas the rate of ES dissociation is equal to (k_{-1} + k_2)[ES]. The steady state assumption equates these two rates:

$$k_1[E][S] = (k_{-1} + k_2)[ES] \tag{15}$$

$$[ES] = \frac{[E][S]}{(k_{-1} + k_2) / k_1} \tag{16}$$

Michaelis and Menten introduced a new constant, K_m (now referred to as the *Michaelis constant*):

$$K_m = \frac{k_{-1} + k_2}{k_1} \tag{17}$$

They also derived the equation

$$v = \frac{V_{max}[S]}{[S] + K_m} \tag{18}$$

where V_{max} = maximum velocity that the reaction can attain.

This equation, now referred to as the *Michaelis–Menten equation*, has proven to be very useful in defining certain aspects of enzyme behavior. For example, when [S] is equal to K_m, the denominator in Equation (18) is equal to 2[S], and v is equal to $V_{max}/2$ (**Figure 6.13**). The experimentally determined value K_m (measured in moles per liter of substrate) is considered a constant that is characteristic of the enzyme and the substrate under specified conditions. It may reflect the affinity of the enzyme for its substrate. (If k_2 is much smaller than k_{-1}, that is, $k_2 << k_{-1}$, then the K_m value approximates k_1. In this circumstance, K_m is the dissociation constant for the ES complex.) The lower the value of K_m, the lower the [S] required to reach 1/2 V_{max} and, therefore, the greater the "affinity" of the enzyme for the substrate. For many enzymes, K_m values reflect normal in vivo substrate concentrations.

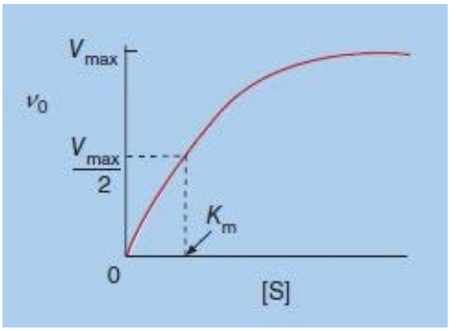


FIGURE 6.13

Initial Reaction Velocity v_0 and Substrate Concentration [S] for a Typical Enzyme-Catalyzed Reaction

The enzyme has half-maximal velocity at substrate concentration K_m.

An enzyme's kinetic properties can also be used to determine its catalytic efficiency. The **turnover number** (k_{cat}) of an enzyme is defined as

$$k_{cat} = \frac{V_{max}}{[E_t]} \tag{19}$$

where k_{cat} = number of substrate molecules converted to product per unit time by an enzyme molecule under saturating conditions

$[E_t]$ = total enzyme concentration

Under physiological conditions, [S] is usually significantly lower than K_m. A more useful measure of catalytic efficiency is obtained by rearranging Equation (19) as

$$V_{max} = k_{cat}[E_t] \tag{20}$$

and substituting this function in the Michaelis–Menten equation (Equation 18):

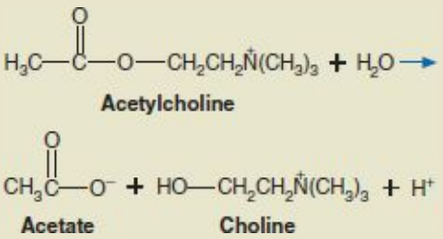
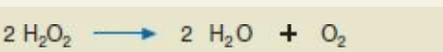
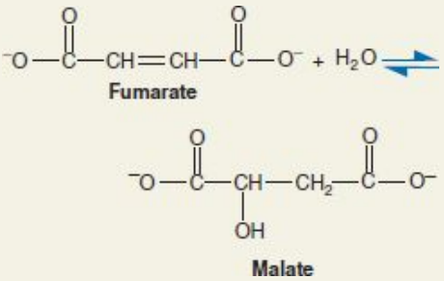
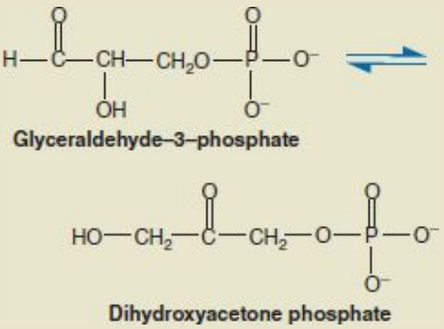
$$v = \frac{k_{cat}[E_t][S]}{K_m + [S]} \tag{21}$$

When [S] is very low, $[E_t]$ is approximately equal to [E], and Equation (21) reduces to

$$v = (k_{cat}/K_m)[E][S] \tag{22}$$

In Equation (22) the term k_{cat}/K_m, also referred to as the **specificity constant**, is the second-order rate constant for a reaction in which [S] << K_m, a common occurrence in biological systems. In this reaction, [S] is sufficiently low that the value of k_{cat}/K_m reflects the relationship between catalytic rate and substrate-binding affinity. A substrate with a high specificity constant will have a low K_m (high affinity) and a high kinetic efficiency (high turnover number). The specificity constant can be useful when comparing different substrates for the same enzyme. When [S] is low and the $[E_t]$ can be accurately measured, Equation (22) holds. Examples of k_{cat}, K_m, and k_{cat}/K_m values of selected enzymes are provided in **Table 6.3**.

TABLE 6.4 The Values of k_{cat}, k_m, and k_{cat}/k_m for Selected Enzymes

Enzyme	Reaction Catalyzed	k_{cat} (s^{-1})	K_m (M)	k_{cat}/K_m ($M^{-1}s^{-1}$)
Acetylcholinesterase		1.4×10^4	9×10^{-5}	1.6×10^8
Carbonic anhydrase		4×10^5	0.026	1.5×10^7
Catalase		4×10^7	1.1	4×10^7
Fumarase		8×10^2	5×10^{-6}	1.6×10^8
Triose phosphate isomerase		4.3×10^3	4.7×10^{-4}	2.4×10^8

Source: Adapted from A. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, 2nd ed., W. H. Freeman, New York, 1999.

It should be noted that the upper limit for an enzyme's k_{cat}/K_m value cannot exceed the maximal value of the rate at which the enzyme can bind to substrate molecules (k_1). This limit is imposed by the rate of diffusion of substrate into an enzyme's active site. The *diffusion control limit* on enzymatic reactions is approximately 10^8 to 10^9 $M^{-1}s^{-1}$. Several enzymes, such as those listed in **Table 6.4**, have k_{cat}/K_m values that approach the diffusion control limit. Because such enzymes convert substrate to product virtually every time the substrate diffuses into the active site, they are said to have achieved *catalytic perfection*. Living organisms overcome the diffusion control limit for the enzymes in biochemical pathways that do not achieve this high degree of catalytic efficiency by organizing them into multienzyme complexes. In these complexes, the active sites of the enzymes are in such close proximity to each other that diffusion is not a factor in the transfer of substrate and product molecules.

KEY CONCEPTS

- The Michaelis–Menten kinetic model explains several aspects of the behavior of many enzymes.
- Each enzyme has a characteristic K for a particular substrate under specified conditions.

m

Enzyme activity is measured in *international units* (IU). One IU is defined as the amount of enzyme that produces 1 μmol of product per minute. An enzyme's *specific activity*, a quantity that is used to monitor enzyme purification, is defined as the number of international units per milligram of protein. [A new unit for measuring enzyme activity called the *katal* has recently been introduced. One katal (kat) indicates the amount of enzyme that transforms 1 mole of substrate per second. One katal is equal to 6×10^7 IU.]

Worked problem 6.2

Consider the Michaelis–Menten plot illustrated in **Figure 6.14**. Identify the following points on the curve:

a. V_{max}
b. K_m

SOLUTION

Refer to **Figure 6.13**.

a. The maximum rate the enzyme can attain is V_{max}. Further increases in substrate concentration do not increase the rate.
b. K_m = [S] at ■

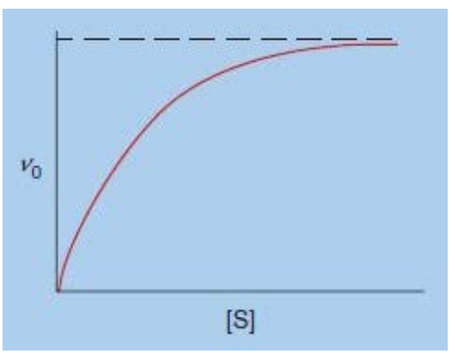


FIGURE 6.14

A Michaelis–Menten Plot

Lineweaver–Burk Plots

The K_m and V_{max} values for an enzyme are determined by measuring initial reaction velocities at various substrate concentrations. Approximate values of K_m and V_{max} can be obtained by constructing a graph, as shown in **Figure 6.13**. A more accurate determination of these values results from an algebraic transformation of the data. The Michaelis–Menten equation, whose graph is a hyperbola,

$$v = \frac{V_{max}[S]}{[S] + K_m}$$

can be rearranged by taking its reciprocal:

$$\frac{1}{v_0} = \frac{K_m}{V_{max}} \frac{1}{[S]} + \frac{1}{V_{max}}$$

The reciprocals of the initial velocities are plotted as functions of the reciprocals of substrate concentrations. In such a graph, referred to as a *Lineweaver–Burk double-reciprocal plot*, the straight line that is generated has the form $y = mx + b$, where y and x are variables ($1/v$ and $1/[S]$, respectively) and m and b are constants (K_m/V_{max} and $1/V_{max}$, respectively). The slope of the straight line is K_m/V_{max} (**Figure 6.15**). As indicated in **Figure 6.15**, the intercept on the vertical axis is $1/V_{max}$. The intercept on the horizontal axis is $-1/K_m$. More precise values of V_{max} and K_m are obtained using computer-based nonlinear curve fitting of the data.

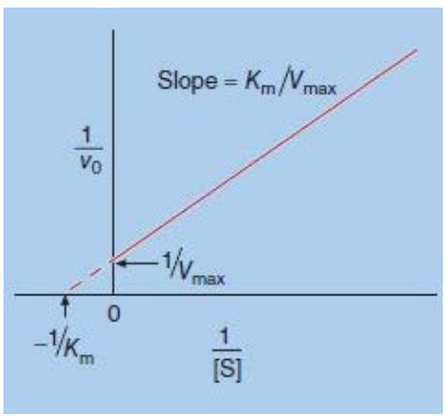


FIGURE 6.15

Lineweaver–Burk, or Double-Reciprocal, Plot

If an enzyme obeys Michaelis–Menten kinetics, a plot of the reciprocal of the reaction velocity $1/v_0$ as a function of the reciprocal of the substrate concentration $1/[S]$ will fit a straight line. The slope of the line is K_m/V_{max}. The intercept on the vertical axis is $1/V_{max}$. The intercept on the horizontal axis is $-1/K_m$.

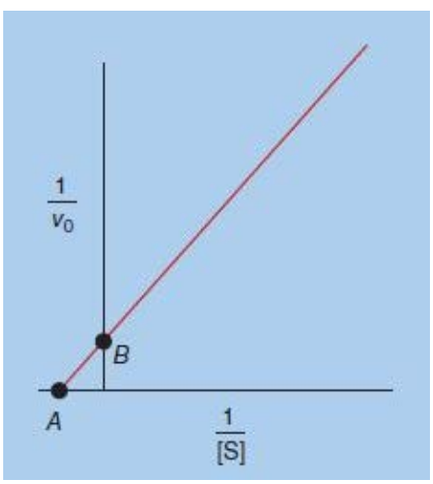


FIGURE 6.16

A Lineweaver–Burk Plot

Problem 6.3 is an example of a kinetics problem using the Lineweaver–Burk plot.

Worked problem 6.3

Consider the Lineweaver–Burk plot in **Figure 6.16**. Identify:

a. $-1/K_m$
b. $1/V_{max}$
c. K_m/V_{max}

SOLUTION

a. $A = -1/K_m$
b. $B = 1/V_{max}$
c. K_m/V_{max} = slope ■

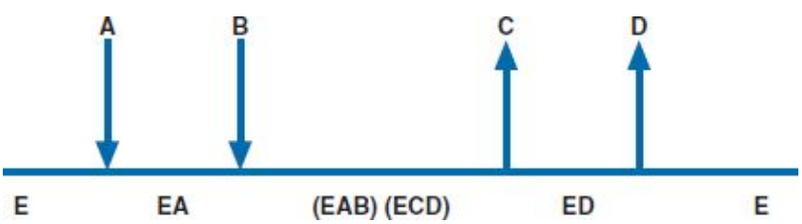
Multisubstrate Reactions

Most biochemical reactions involve two or more substrates. The most common multisubstrate reaction, the bisubstrate reaction, is represented as

$$A + B \leftrightharpoons C + D$$

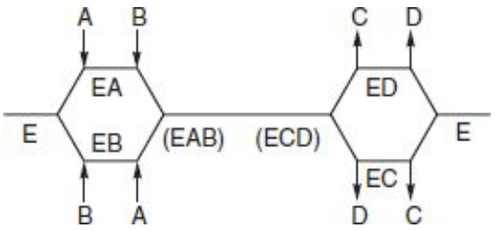
In most of these reactions, there is a transfer of a specific functional group from one substrate to the other (e.g., phosphate or methyl groups) or oxidation- reduction reactions in which redox coenzymes (e.g., NAD^+/NADH or FAD/$FADH_2$) are substrates. The kinetic analysis of bisubstrate reactions is, by necessity, more complicated than that of one-substrate reactions. Often, however, the determination of the K_m for each substrate (when the other substrate is present in a saturating amount) is sufficient for most purposes. On the basis of kinetic analysis, multisubstrate reactions can be divided into two classes: sequential and double displacement.

SEQUENTIAL REACTIONS A sequential reaction cannot proceed until all substrates have been bound in the enzyme's active site. There are two possible sequential mechanisms: ordered and random. In a bisubstrate ordered mechanism, the first substrate must bind to the enzyme before the second for the reaction to proceed to product formation. The notation for such a reaction, introduced by W. W. Cleland, is



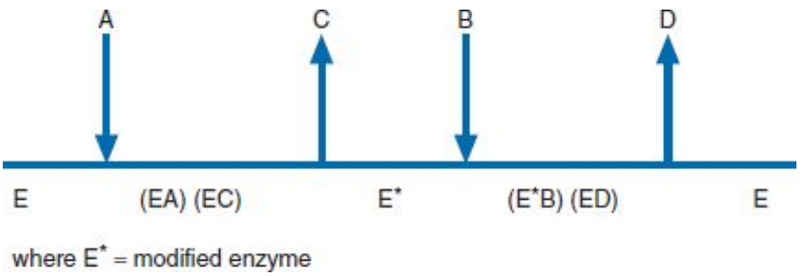
where A and B are substrates, C and D are products, and E is an enzyme.
Substrate A binds before substrate B, and a ternary complex (EAB) is formed. Once the products-containing ternary complex forms (ECD), as a result of catalysis, the enzyme releases first product C and then product D. When lactate is oxidized to yield pyruvate (p. 293) in the redox reaction catalyzed by lactate dehydrogenase, the coenzyme NAD^+ enters the active site first, followed by lactate.

In a random sequential mechanism, the substrates can bind in any order, and the products can be released in any order.



In the reversible reaction catalyzed by creatine kinase (p. 379), in which creatine reacts with ATP to yield phosphocreatine, the entrance of the substrates into the active site and the release of the products is random.

DOUBLE-DISPLACEMENT REACTIONS In a double-displacement, or "ping-pong" mechanism, the first product is released before the second substrate binds.



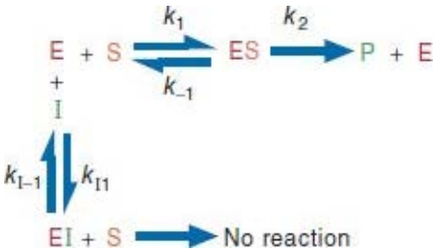
In such a reaction, the enzyme is altered by the first phase of the reaction. The enzyme is then restored to its original form during the second phase, in which the second substrate is converted to product. Transamination reactions (pp. 533–36), interconversions of α-amino acids and α-keto acids, catalyzed by a group of transaminases, are examples of double-displacement reactions.

Enzyme Inhibition

The activity of enzymes can be inhibited. Molecules that reduce an enzyme's activity, called **inhibitors**, include many drugs, antibiotics, food preservatives, poisons, and metabolites of normal biochemical processes. The investigations of enzyme inhibition and inhibitors carried out by biochemists are important for several reasons. First, in living systems, enzyme inhibition is an important means by which metabolic pathways are regulated. Numerous small biomolecules are used routinely to modulate the rates of specific enzymatic reactions so that the needs of the organism are consistently met. Second, numerous clinical therapies are based on enzyme inhibition. For example, many antibiotics and other drugs reduce or eliminate the activity of specific enzymes. The most effective AIDS treatment is a multidrug therapy that includes protease and reverse transcriptase inhibitors, molecules that disable viral enzymes required to make new virus. Finally, investigations of enzyme inhibition have enabled biochemists to develop techniques for probing the physical and chemical architecture, as well as the functional properties, of enzymes.

Enzyme inhibition can be reversible or irreversible. **Reversible inhibition** occurs when increasing substrate levels or removing the inhibitor compound while the enzyme remains intact can counteract the inhibitory effect of a compound. Reversible inhibition will be **competitive** if the inhibitor and the substrate bind to the same site, **noncompetitive** if the inhibitor binds to a site other than the active site, and **uncompetitive** if the inhibitor-binding site is created after the substrate is bound to the enzyme. **Irreversible inhibition** occurs when inhibitor binding permanently impairs the enzyme, usually through a covalent reaction that chemically modifies the enzyme.

COMPETITIVE INHIBITORS Competitive inhibitors bind reversibly to free enzyme, not the ES complex, to form an enzyme-inhibitor (EI) complex.



The substrate and the inhibitor compete for the same site on the enzyme. As soon as the inhibitor binds, substrate access to the active site is blocked. The concentration of EI complex depends on the concentration of free inhibitor and on the dissociation constant K_I:

$$K_I = \frac{[E][I]}{[EI]} = \frac{k_{I-1}}{k_{I1}}$$

where K_I is a measure of the enzyme's binding affinity for the inhibitor.

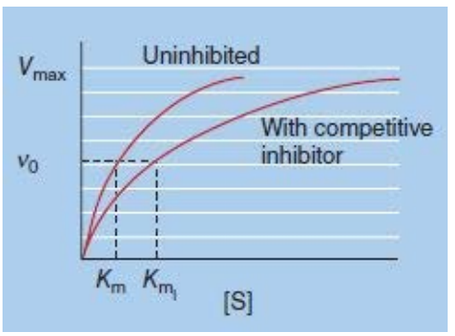


FIGURE 6.17

Michaelis–Menten Plot of Uninhibited Enzyme Activity versus Competitive Inhibition

Initial velocity v_0 is plotted against substrate concentration [S]. With competitive inhibition, V_{max} stays constant and K_m increases.

Because the EI complex readily dissociates, the enzyme is again available for substrate binding. The enzyme's activity declines with an increased K_m and unchanged V_{max} (**Figure 6.17**) in the presence of a competitive inhibitor because the EI complex does not participate in the catalytic process. Because V_{max} is unchanged, the effect of a competitive inhibitor on activity is reversed by increasing the concentration of substrate. At high [S], all the active sites are filled with substrate, and reaction velocity reaches the value observed without an inhibitor. The Michaelis–Menten equation for competitive inhibition that takes the formation of [EI] into account is obtained by incorporating a term for the effect of the inhibitor on K_m:

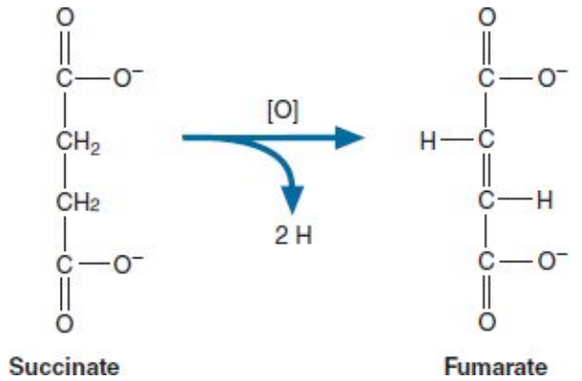
$$v = \frac{V_{max}[S]}{[S] + \alpha K_m}$$

where $\alpha = 1 + [I]/K_i$.

The term α is a function of both the competitive inhibitor's concentration and its affinity for the enzyme active site. The K_m value (the [S] at which $v = 1/2\ V_{max}$) increases in the presence of the inhibitor by the factor α. The term αK_m is often referred to as the apparent K_m (K_m^{app}).

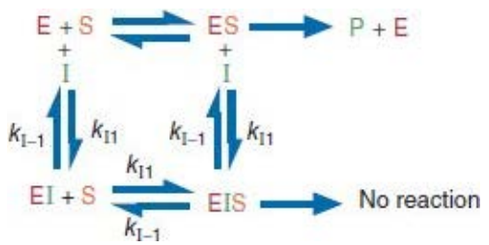
Substances that behave as competitive inhibitors (i.e., reduce an enzyme's apparent affinity for

substrate) are often similar in structure to the substrate. Such molecules include reaction products or unmetabolizable analogues or derivatives of substrate molecules. Succinate dehydrogenase, an enzyme in the Krebs citric acid cycle (Chapter 9), catalyzes a redox reaction that converts succinate to fumarate.



This reaction is inhibited by malonate (**Figure 6.18**). Malonate binds to the enzyme's active site but cannot be converted to product. Molecules that resemble the structure of a substrate's transition state, referred to as *transition state analogues*, can be especially efficient competitive inhibitors. They bind in the enzyme active site with a higher affinity than does the substrate. Oseltamivir (Tamiflu), used to prevent and treat influenza, is converted in the liver into a transition state analogue of sialic acid (a nine-carbon sugar). Tamiflu's binding in the active site of the viral enzyme neuraminidase inhibits cleavage of sialic acid from certain host cell proteins, thereby preventing the release of new virus from infected cells.

NONCOMPETITIVE INHIBITORS In some enzyme-catalyzed reactions, an inhibitor can bind to both the enzyme and the enzyme–substrate complex:



In such circumstances, referred to as **noncompetitive** inhibition, the inhibitor binds to a site other than the active site. Inhibitor binding results in a modification of the enzyme's conformation that prevents product formation (**Figure 6.19**). Noncompetitive inhibitors have little or no structural resemblance to substrate, but they may influence substrate binding if their binding sites are in close proximity to the substrate binding site. Analysis of reactions inhibited by noncompetitive inhibitors is often complex because usually two or more substrates are involved. Thus, the characteristics of the inhibition that are observed may depend in part on factors such as the order in which the different substrates bind.

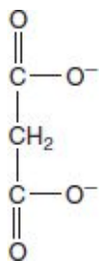


FIGURE 6.18

Structure of Malonate

3D animation of malonate

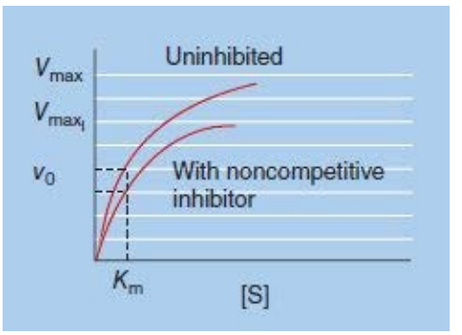


FIGURE 6.19

Michaelis–Menten Plot of Uninhibited Enzyme Activity versus Noncompetitive Inhibition

Initial velocity v_0 is plotted against substrate concentration [S]. With noncompetitive inhibition V_{max} decreases and K_m remains unchanged.

There are two forms of noncompetitive inhibition: pure and mixed. In pure noncompetitive inhibition (the inhibitor binds far from the active site), a rare phenomenon, both α and α’ (a measure of the affinity of the enzyme–substrate complex for the inhibitor) are equivalent. Mixed noncompetitive inhibition (the inhibitor binds close to the active site and to the enzyme–substrate complex) is typically more complicated because the α and α’ values are different. The Michaelis–Menten equation that describes mixed noncompetitive inhibition is

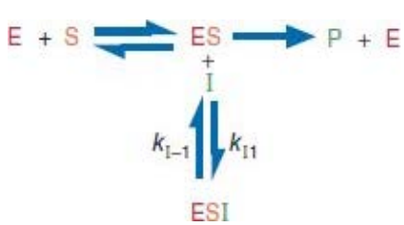
$$v = \frac{V_{max}[S]}{\alpha'[S] + \alpha K_m}$$

where $\alpha = 1 + [I]/K_I$

$\alpha' = 1 + [I]/K'_I$

In pure noncompetitive inhibition (i.e., $\alpha = \alpha'$), the values of V_{max} change, but K_m remains unaffected. In mixed noncompetitive inhibition, the values of both K_m ($K_m^{app} = \alpha/\alpha' K_m$) and V_{max} ($V_{max}^{app} = V_{max}/\alpha'$) are altered.

UNCOMPETITIVE INHIBITORS In uncompetitive inhibition (**Figure 6.20**), which is considered a rare type of noncompetitive inhibition, the inhibitor binds only to the enzyme–substrate complex, not to the free enzyme. As a consequence, the inhibitor is ineffective at low substrate concentrations because very little of the ES complex is present.



The dissociation constant for the binding step of an uncompetitive inhibitor to an enzyme is

$$K_I' = \frac{[ES][I]}{[ESI]}$$

Uncompetitive inhibition is most easily observed at high [S]. When I binds to ES, the substrate is not free to dissociate from the enzyme and the K_m ($K_m^{app} = K_m/\alpha'$) decreases, giving the appearance of higher substrate affinity. The Michaelis–Menten equation that describes uncompetitive inhibition is

$$v = \frac{V_{max}[S]}{\alpha'[S] + K_m}$$

where $\alpha' = 1 + [I]/K_I'$. Because an uncompetitive inhibitor binds only to ES, the equilibrium E + S ⇋ ES is shifted to the right, with the result of increasing the amount of ES available to bind to the inhibitor. The value of V_{max} ($V_{max}^{app} = V_{max}/\alpha'$) is lowered by a factor of α'. Uncompetitive inhibition is most commonly observed for enzymes that bind more than one substrate.

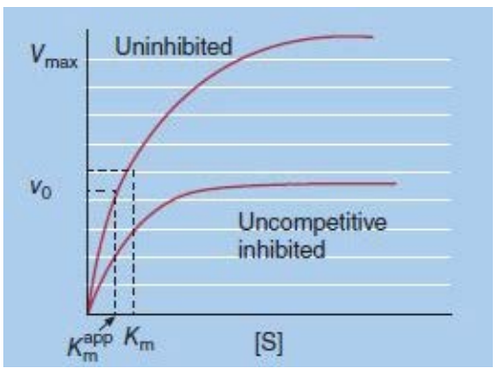


FIGURE 6.20

Michaelis–Menten Plot of Uninhibited Enzyme Activity versus Uncompetitive Inhibition

Initial velocity v_o is plotted against substrate concentration [S]. With uncompetitive inhibition both K_m and V_{max} values decrease.

KINETIC ANALYSIS OF ENZYME INHIBITION Competitive, noncompetitive, and uncompetitive inhibition can be distinguished with double-reciprocal plots (**Figure 6.21 a–d**). In two sets of rate determinations, enzyme concentration is held constant. The first experiment establishes the velocity and kinetic parameters (K_m and V_{max}) of the uninhibited enzyme. In the second experiment, a constant amount of inhibitor is included in each enzyme assay. **Figure 6.21** illustrates the different effects that inhibitors have on enzyme activity. Competitive inhibition increases the K_m of the enzyme, but the V_{max} is unchanged. (This is shown in the double-reciprocal plot as a shift in the horizontal intercept.) In pure noncompetitive inhibition (i.e., $\alpha = \alpha'$) V_{max} is lowered (i.e., the vertical intercept is shifted); the K_m is unchanged in this rare case because the *k* values for the binding to E and ES are the same. In mixed noncompetitive inhibition (i.e., α and α' are not equal), the inhibitor-binding site and the active site are close together, and their binding activities interfere with each other. Consequently, both V_{max} and K_m change. In such cases, the plots of $1/v$ versus $1/[S]$ will intersect to the left of the vertical axis. In uncompetitive inhibition, both K_m and V_{max} are changed, although their ratio (i.e., the slope K_m/V_{max}) remains the same.

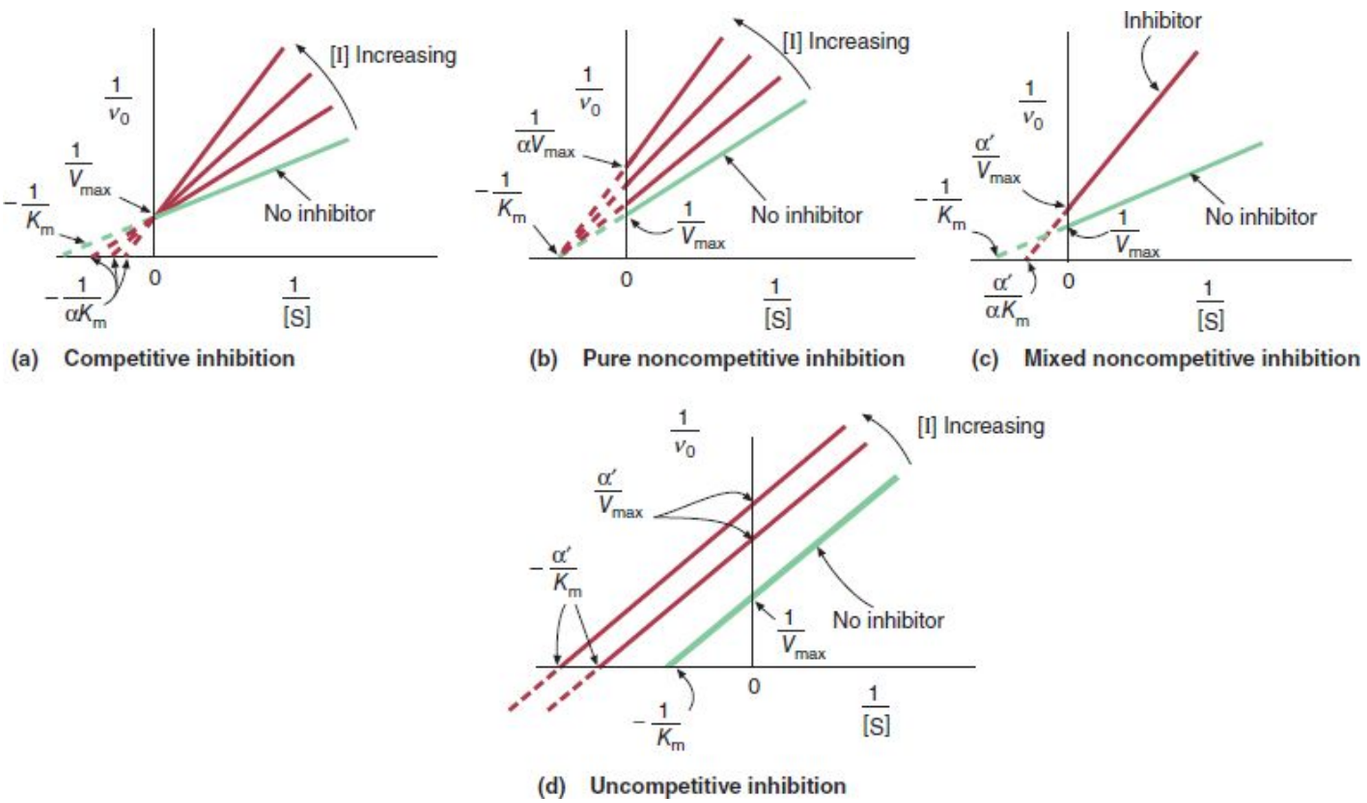


FIGURE 6.21

Kinetic Analysis of Enzyme Inhibition

(a) Competitive inhibition. Plots of $1/v$ versus $1/[S]$ in the presence of several concentrations of the inhibitor result in no change in the vertical intercept (i.e., $1/V_{max}$, so V_{max} is unchanged) and a decrease in the vertical intercept (i.e., $-1/K_m$, so K_m is increased). The Lineweaver–Burk equation for competitive inhibition is $1/v = (\alpha K_m/V_{max})(1/[S]) + 1/V_{max}$. (b) Noncompetitive inhibition. Plots of $1/v$ versus $1/[S]$ in the presence of several concentrations of the inhibitor intersect at the same point on the horizontal axis, $-1/K_m$. In pure noncompetitive inhibition, the constants for ES and EIS are assumed to stay the same. (c) In mixed noncompetitive inhibition, the plots of $1/v$ versus $1/[S]$ will intersect to the left of the vertical axis because α is not equal to α'. The Lineweaver–Burk equation for mixed noncompetitive inhibition is $1/v = (\alpha K_m/Vmax)(1/[S] + \alpha'/V_{max}$. (d) Uncompetitive inhibition. Plots of $1/v$ versus $1/[S]$ in the presence of several concentrations of the inhibitor result in a decrease in K_m (α'/K_m) and a decrease in V_{max} (α'/V_{max}), such that the slope of the line is unchanged but shifts up and to the left. The Lineweaver–Burk equation for uncompetitive inhibition is $1/v = (K_m/V_{max})(1/[S]) + \alpha'/V_{max}$.

IRREVERSIBLE INHIBITION In reversible inhibition, the inhibitor can dissociate from the enzyme because it binds through noncovalent bonds. Irreversible inhibitors usually bond covalently to the enzyme, often to a side chain group in the active site. For example, enzyme active sites containing free sulfhydryl groups can react with alkylating agents such as iodoacetate:

Glyceraldehyde-3-phosphate dehydrogenase, an enzyme in the glycolytic pathway (Chapter 8), is inactivated by alkylation with iodoacetate. Enzymes that use sulfhydryl groups to form covalent bonds with metal cofactors are often irreversibly inhibited by heavy metals (e.g., mercury and lead). The anemia that is symptomatic of lead poisoning is caused in part because lead binds to a sulfhydryl group of ferrochelatase. Ferrochelatase catalyzes the insertion of Fe^{2+} into the heme prosthetic group of hemoglobin.

Problem 6.4 is concerned with enzyme inhibition.

KEY CONCEPTS

- Reversible inhibition of enzymes can be competitive, noncompetitive, or uncompetitive.
- Competitive inhibitors reversibly compete with substrate for the same site on free enzyme.
- Noncompetitive inhibitors can bind to both the enzyme and the enzyme–substrate complex because it binds to a site outside the active site.
- Uncompetitive inhibitors bind only to the enzyme–substrate complex, not the free enzyme.
- Irreversible inhibitors usually bind covalently to the enzyme.

WORKED PROBLEM 6.4

Consider the Lineweaver–Burk plot illustrated in **Figure 6.22**. Identify the type of inhibitory action shown by compounds B, C, and D.

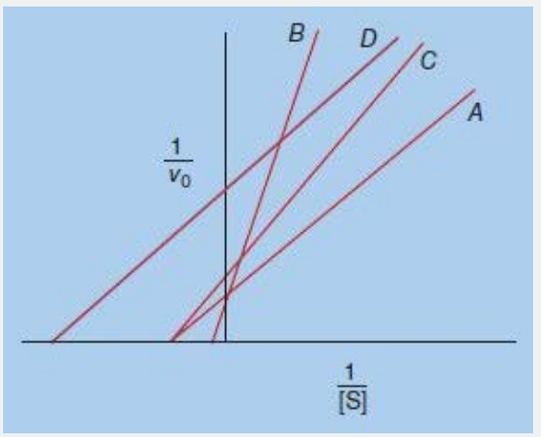


FIGURE 6.22

A Lineweaver–Burk Plot

Line *A* = normal enzyme-catalyzed reaction
Line *B* = compound B added
Line *C* = compound C added
Line *D* = compound D added

SOLUTION

Compound B is a competitive inhibitor because the K_m only has changed. Compound C is a pure noncompetitive inhibitor because the V_{max} only has changed. Compound D is an uncompetitive inhibitor because both K_m and V_{max} have changed. ■

WORKED PROBLEM 6.5

An enzyme has a K_m of 10 μM. When 5 μM of a competitive inhibitor is added, K_I is determined to be 2.5 μM. Determine (a) the value of α and (b) K_m^{app}.

SOLUTION

(a) The value of α is determined as follows:

$$\alpha = 1 + [I]/K_I = 1 + 5\ \mu M/2.5\ \mu M = 1 + 2 = 3$$

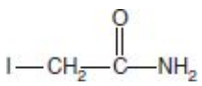
(b) K_m^{app}, the measured K_m value when the enzyme is inhibited, is calculated as follows:

$\alpha = K_m^{app}/K_m$

$K_m^{app} = \alpha K_m = (3)\ (10\ mM) = 30\ mM$ ■

QUESTION 6.7

Iodoacetamide is an irreversible inhibitor of several enzymes that have a cysteine residue in their active sites. After examining its structure, predict the product of the reaction of iodoacetamide with such an enzyme.



ALLOSTERIC ENZYMES Although the Michaelis–Menten model is an invaluable tool, it does not explain the kinetic properties of many enzymes. For example, plots of reaction velocity versus substrate concentration for many enzymes with multiple subunits are often sigmoidal rather than hyperbolic, as predicted by the Michaelis–Menten model (**Figure 6.23**). Such effects are seen in an important group of enzymes called the **allosteric enzymes**. Note that the substrate-binding curve in **Figure 6.23** resembles the oxygen-binding curve of hemoglobin (p. 178).

The properties of allosteric enzymes are discussed on pp. 233–35.

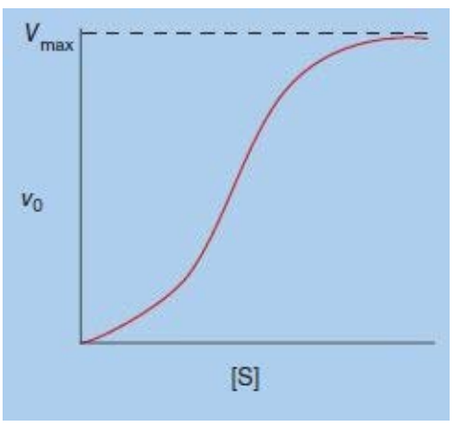


FIGURE 6.23

The Kinetic Profile of an Allosteric Enzyme

The sigmoidal binding curve displayed by many allosteric enzymes resembles the curve for the cooperative binding of O_2 to hemoglobin.

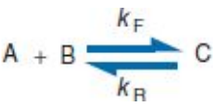
QUESTION 6.8

Drinking methanol can cause blindness in humans, as well as a severe acidosis that may be life

threatening. Methanol is toxic because it is converted in the liver to formaldehyde and formic acid by the enzymes alcohol dehydrogenase and aldehyde dehydrogenase. Methanol poisoning is treated with dialysis and infusions of bicarbonate and ethanol. Explain why each treatment is used.

Enzyme Kinetics, Metabolism, and Macromolecular Crowding

The ultimate goal of enzyme kinetic investigations is the development of a realistic understanding of metabolism in living organisms. Despite a wealth of knowledge, gathered over many decades, about biochemical pathways and the in vitro ("in glass," i.e., measured in the laboratory) kinetic properties of enzymes, insight into metabolic processes in "live" cells is elusive. An important reason for this lack of success is now believed to lie in the assumptions of the law of mass action. According to this rule, when an equilibrium constant for a reaction such as



is calculated, it is assumed that the forward rate is *linear* (i.e., is directly proportional) to the concentrations of A and B, and that the rate of the reverse reaction (k_R) is linear with respect to the concentration of C. For the K_{eq} value derived for this or any reaction to be valid, it is further assumed that (1) the system in which the reaction occurs is homogeneous (i.e., there is uniform mixing of its contents) and (2) interacting molecules move randomly and independently of each other.

It is now understood that under crowded in vivo conditions, neither of these assumptions holds true. The crowded interior of cells is highly heterogeneous (i.e., an enormous variety of molecular species are present) and filled with macromolecules, membranes, cytoskeletal components, and other obstacles that impede molecular movement. This phenomenon is called **macromolecular crowding**. Among its observed effects within cells are increased values for the effective concentrations of macromolecules, enhanced protein folding and binding affinities, changes in reaction rates and equilibrium constants, and decreased diffusion rates. These effects are *nonlinear*; that is, they are difficult to predict. For example, the in vivo kinetic values (K_m, V_{max}, and K_{eq}) of each biochemical reaction depend on the activity coefficient of its enzyme and the diffusion rates of substrate and effector molecules, neither of which is easy to describe in the heterogeneous environment of the cell. The slow diffusion of some substrates is overcome by the segregation of reactions into microcompartments where substrate concentration is relatively high. *Metabolons*, multienzyme complexes composed of pathway enzymes, are also used. Pathway intermediates are directly transferred, or "channeled," within metabolons from one active site to the next. These microenvironments have their own unique ionic composition that promotes the efficient operation of the resident enzymes. It is a significant challenge to biochemists to identify these local conditions and duplicate them in vitro to facilitate the investigation of specific enzymes and determine realistic kinetic parameters. Discrepancies between results derived in vivo, in vitro, and in silico (by computer simulation) indicate that an enzyme isolated from its metabolic pathway and other intersecting pathways does not function with the same kinetic parameters.

Despite the daunting complexities of living cells, systems biologists such as metabolic engineers and industrial scientists have endeavored to develop new strategies to investigate metabolic processes. Their goals include improvements in the metabolic capacities of the microorganisms used to produce commercial products. Among the most important tools used by these scientists are dynamic computer simulations of metabolic processes that are based on both in vivo and in vitro data. The mathematical model on which the computer simulation is based is constructed from kinetic data and equations that define *metabolic flux*, the rate of flow of metabolites, such as substrates, products, and intermediates, along biochemical pathways. (Despite the limitations of in vitro kinetic

values just described, they are recognized as invaluable baseline data during the early phases of model development.) The ultimate goal of a computer simulation is to approximate the in vivo steady state behavior of metabolic processes and predict how they will respond to parameter changes such as nutrient availability or mutant forms of enzymes. Insights already gained from computer simulations (e.g., see the discussion of glycolysis and jet engines in Chapter 8) guarantee that in silico modeling and pathway simulation will form an increasingly significant part of future enzyme kinetics investigations.

6.5 ENZYME REGULATION

Living organisms have sophisticated mechanisms for regulating their vast networks of biochemical pathways. Regulation is essential for several reasons:

1. **Maintenance of an ordered state.** Regulation of each pathway results in the production of the substances required to maintain cell structure and function in a timely fashion and without wasting resources.
2. **Conservation of energy.** Cells ensure that they consume just enough nutrients to meet their energy requirements by constantly controlling energy-generating reactions.
3. **Responsiveness to environmental changes.** Cells can make relatively rapid adjustments to changes in temperature, pH, ionic strength, and nutrient concentrations because they can increase or decrease the rates of specific reactions.

The regulation of biochemical pathways is achieved primarily by adjusting the concentrations and activities of certain enzymes. Control is accomplished by (1) genetic control, (2) covalent modification, (3) allosteric regulation, and (4) compartmentation.

Genetic Control

The synthesis of enzymes in response to changing metabolic needs, a process referred to as **enzyme induction**, allows cells to respond efficiently to changes in their environment. For example, *E. coli* cells grown without the sugar lactose initially cannot metabolize this nutrient when it is introduced into the bacterium's growth medium. The introduction of lactose in the absence of glucose, however, activates the genes that code for enzymes needed to utilize lactose as an energy source. After all the lactose has been consumed, synthesis of these enzymes is terminated.

Covalent Modification

Some enzymes are regulated by the reversible interconversion between their active and inactive forms. Several covalent modifications of enzyme structure cause these changes in function. Many such enzymes have specific residues that may be phosphorylated and dephosphorylated. For example, glycogen phosphorylase (Chapter 8) catalyzes the first reaction in the degradation of glycogen, a carbohydrate energy storage molecule. In a process controlled by hormones, the inactive form of the enzyme (glycogen phosphorylase b) is converted to the active form (glycogen phosphorylase a) by the addition of a phosphate group to a specific serine residue. In the nonphosphorylated enzyme, the peptide segment directly surrounding the serine residue is disordered. Phosphorylation of the serine side chain triggers the conversion of the disordered segment into α-helices, thereby resulting in a more active enzyme. Other types of reversible covalent modification include methylation, acetylation, and nucleotidylation (the covalent addition of a nucleotide).

Several enzymes are produced and stored as inactive precursors called **proenzymes** or **zymogens**.

Zymogens are converted into active enzymes by the irreversible cleavage of one or more peptide bonds. For example, chymotrypsinogen is produced in the pancreas. After chymotrypsinogen is secreted into the small intestine, it is converted to its active form, which is responsible for digesting dietary protein in several steps (**Figure 6.24**).

Allosteric Regulation

In each biochemical pathway, there are one or more enzymes whose catalytic activity can be modulated (i.e., increased or decreased) by the binding of effector molecules. Recall, for example, regulatory feedback motifs (p. 24), which occur in numerous metabolic pathways. In such pathways, increases in the concentration of product molecules alter the activity of the enzyme that catalyzes the pathway's committed step. The binding of ligands to *allosteric sites* on such an enzyme triggers rapid conformational changes that can increase or decrease its rate of substrate binding. A plot of a reaction catalyzed by an allosteric enzyme differs from those of enzymes that observe Michaelis–Menten kinetics. Instead the rate curve is sigmoidal (**Figure 6.25**). If the ligand is the same as the substrate (i.e., if the binding of a substrate influences the binding of additional substrate), the allosteric effects are referred to as *homotropic*. *Heterotropic effects* involve modulating ligands, which are different from the substrate. Most allosteric enzymes are multisubunit enzymes with multiple substrate and effector binding sites.

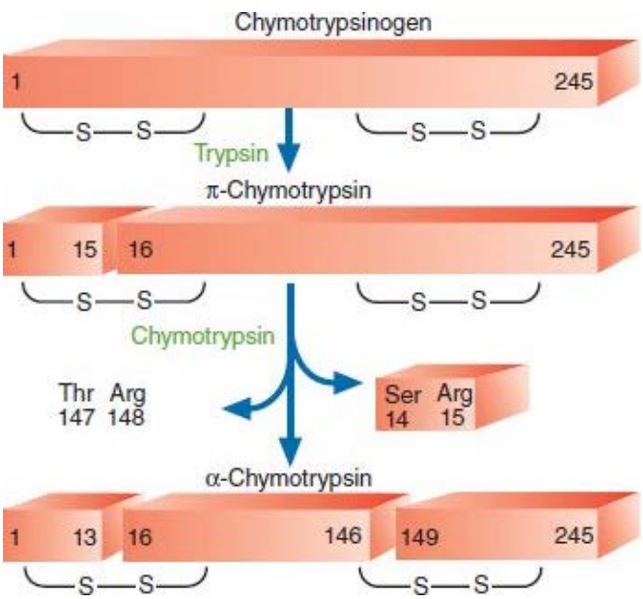


FIGURE 6.24

The Activation of Chymotrypsinogen

The inactive zymogen chymotrypsinogen is activated in several steps. After its secretion into the small intestine, chymotrypsinogen is converted into π-chymotrypsin when trypsin, another proteolytic enzyme, cleaves the peptide bond between Arg 15 and Ile 16. Later, cleaved chymotrypsin molecules activate each other by removing two depeptide segments to cause the formation of *α*-chymotrypsin.

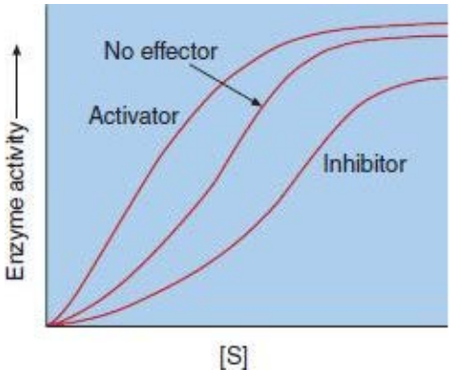


FIGURE 6.25

The Rate of an Enzyme-Catalyzed Reaction as a Function of Substrate Concentration

The activity of allosteric enzymes is affected by positive effectors (activators) and negative effectors (inhibitors).

Two theoretical models that attempt to explain the behavior of allosteric enzymes are the concerted model and the sequential model (**Figure 6.26**). In the *concerted* (or *symmetry*) model, it is assumed that the enzyme exists in only two states (Section 5.3): T(aut) and R(elaxed). Activators bind to and stabilize the R conformation, whereas inhibitors bind to and stabilize the T conformation. The term *concerted* is applied to this model because the conformations of all the protein's subunits are believed to change simultaneously when the first effector binds. (This rapid concerted change in conformation maintains the protein's overall symmetry.) The binding of an activator shifts the equilibrium in favor of the R form. An inhibitor shifts the equilibrium toward the T conformation.

The concerted model explains some properties of allosteric enzymes but not others. For example, the concerted model accounts for **positive cooperativity**, in which the first ligand increases subsequent ligand binding but not **negative cooperativity**, a phenomenon observed in enzymes in which the binding of the first ligand reduces the affinity of the enzyme for similar ligands. In addition, the concerted model makes no allowances for hybrid conformations. Examples of enzymes whose behavior appears to be consistent with the concerted model are aspartate transcarbamoylase (ATCase) in *E. coli* and phosphofructokinase (PFK).

ATCase catalyzes the first step in a reaction pathway that leads to the synthesis of the pyrimidine nucleotide cytidine triphosphate (CTP) (**Figure 6.27**). CTP acts as an inhibitor of ATCase activity. It shifts the rate curve to the right, indicating an increase in the apparent K_m of ATCase that corresponds to a lower velocity of enzyme activity at a given concentration of substrate. CTP inhibition of ATCase is an example of negative feedback inhibition (p. 24), the process in which the product of a pathway inhibits the activity of an enzyme at or near the beginning of a biochemical pathway or at a branch point. The purine nucleotide ATP acts as an activator. ATP activation of ATCase makes sense because nucleic acid biosynthesis requires relatively equal amounts of purine and pyrimidine nucleotides. When ATP concentration is higher than that of CTP, ATCase is activated. When ATP concentration is lower than that of CTP, the net effect on ATCase is inhibitory.

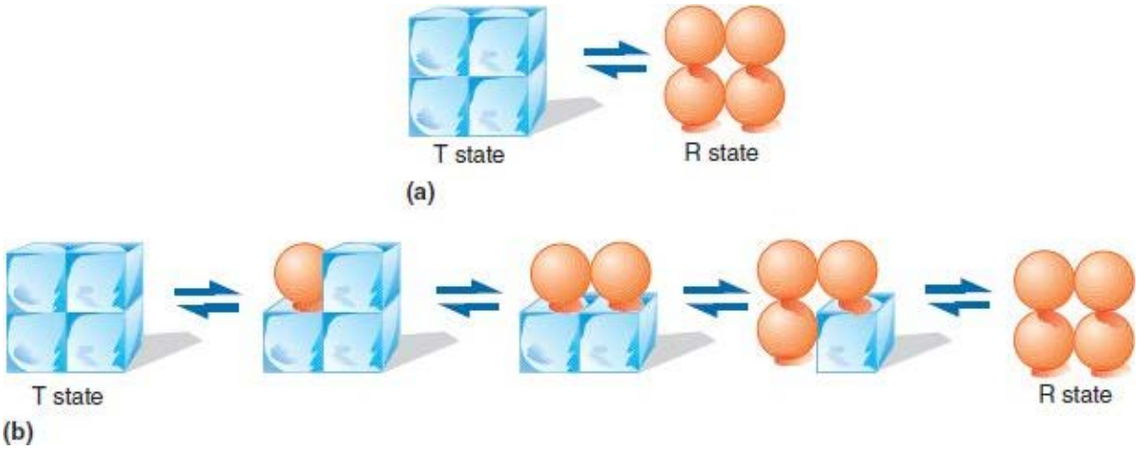


FIGURE 6.26

Allosteric Interaction Models

(a) In the concerted model, the enzyme exists in only two conformations. Substrates and activators have a greater affinity for the R state. Inhibitors favor the T state. (b) In the sequential model, one subunit assumes an R conformation as it binds to substrate. As the first subunit changes its conformation, the affinity of nearby subunits for ligand is affected.

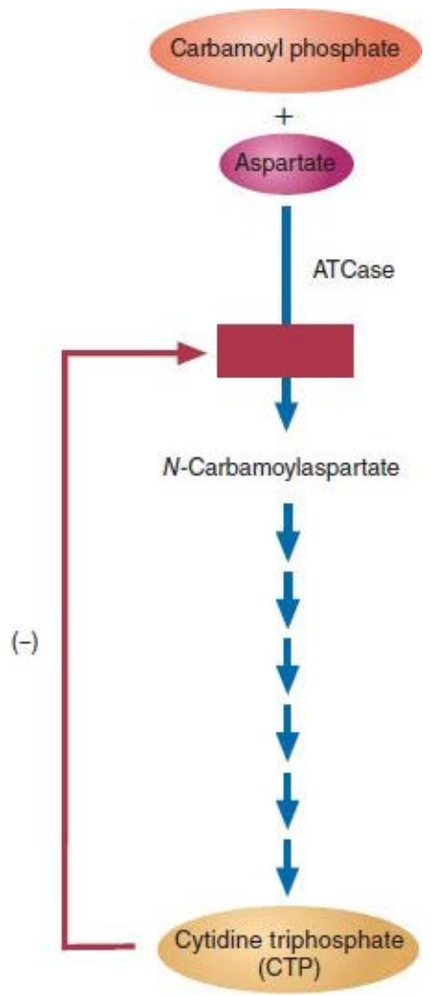


FIGURE 6.27

Feedback Inhibition

ATCase (aspartate transcarbamoylase) catalyzes the committed step in the synthesis of cytidine triphosphate (CTP). The binding of CTP, the product of the pathway, to ATCase inhibits the enzyme.

Phosphofructokinase (PFK) catalyzes an important reaction in glycolysis (p. 287): the transfer of a phosphate group from ATP to the OH group on C-1 of fructose-6-phosphate.

$$\text{Fructose-6-phosphate} + \text{ATP} \rightarrow \text{fructose-1,6-bisphosphate} + \text{ADP}$$

PFK is composed of four identical subunits, each of which has an active site and a number of allosteric sites. The R state of the enzyme is stabilized by ADP and AMP (a sensitive indicator of ATP depletion and the cell's need for energy). The T state is stabilized by molecules such as ATP and other molecules that are indicators that cellular energy is high.

In the sequential model, first proposed by Daniel Koshland, the binding of a ligand to one flexible subunit in a multisubunit protein triggers a conformational change that is sequentially transmitted to adjacent subunits. The more sophisticated sequential model allows for the intermediate conformations that may be more realistic representations of the operation of some enzymes. The model can also accommodate negative cooperativity (ligand binding to one subunit induces conformational changes in adjacent subunits that make ligand binding less likely). Consequently, the sequential model can be considered a general model that accounts for all allosteric possibilities. In this view, the concerted model is a simple example of the sequential model.

Two important aspects of allosteric regulation should be noted. First, the concerted and sequential models are theoretical models; that is, the behavior of many allosteric proteins appears to be more complex than can be accounted for by either model. For example, the cooperative binding of O_2 by hemoglobin (the most thoroughly researched allosteric protein) appears to exhibit features of both models. The binding of the first O_2 initiates a concerted T $\rightarrow$ R transition that involves small changes in the conformation of each subunit (a feature of the sequential model; see p. 234). In addition, hemoglobin species with only one or two bound O_2 have been observed. A second and more important issue is that there are no simple rules that explain metabolic regulation. For example, attempts to increase the metabolic flux of pathways in organisms such as yeast by engineering increased synthesis of allosteric enzymes failed. (*Flux* is the rate of turnover of molecules through a pathway.) Flux increases were observed only when all of the enzymes in the pathway were increased. It appears that pathway regulation is the result of regulatory contributions to a greater or lesser degree by all or most of the enzymes.

Compartmentation

Compartmentation, created by cellular infrastructure, is a significant means for regulating biochemical reactions because physical separation makes separate control possible. The intricate internal architecture of cells contains compartments (e.g., eukaryotic organelles) and microcompartments of diverse types (e.g., individual enzymes or multiprotein complexes attached to membranes or cytoskeletal filaments). Cellular compartmentation solves several interrelated problems:

1. **Divide and control**. The physical separation of competing reactions (i.e., those that would undo the other, such as kinases and phosphatases) allows for coordinated regulation that prevents the wasteful dissipation of resources.
2. **Diffusion barriers**. Within crowded cells, diffusion of substrate molecules is a potentially limiting factor in reaction rates. Cells circumvent this problem by creating microenvironments in which enzymes and their substrates are concentrated, as well as by metabolite channeling, the transfer of product molecules from one enzyme to the next in a multienzyme complex.
3. **Specialized reaction conditions**. Certain reactions require an environment with unique properties. For example, the low pH within lysosomes facilitates hydrolytic reactions.
4. **Damage control**. The segregation of potentially toxic reaction products protects other cellular

components.

Overall metabolic control requires the integration of all the cell's biochemical pathways, which is accomplished, in part, by transport mechanisms that transfer metabolites and signal molecules between compartments.

KEY CONCEPTS

- All biochemical pathways are regulated to maintain the ordered state of living cells.
- Regulation is accomplished by genetic control, covalent modification of enzymes, allosteric regulation, and cell compartmentation.

Tuberculosis

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on enzymes and clinical medicine.

QUESTION 6.9

Certain enzymes are involved in the metabolism of the medicinal chemicals that alter or enhance physiological processes. For example, aspirin suppresses pain, and antibiotics kill infectious organisms. Once such a drug has been consumed, it is absorbed and distributed to the tissues, where it performs its function. Eventually, drug molecules are processed (primarily in the liver) and excreted. The dosage of each drug that physicians prescribe is based on the amount required to achieve a therapeutic effect and on the drug's average rate of excretion from the body. Various reactions prepare drug molecules for excretion. Examples include oxidation, reduction, and conjugation reactions. (In conjugation reactions, small polar or ionizable groups are attached to a drug molecule to improve its solubility.) The amount of certain enzymes directly affects a patient's ability to metabolize a specific drug. For example, isoniazid is used to treat *tuberculosis*, a highly infectious, chronic, disease caused by *Mycobacterium tuberculosis*. This drug is metabolized by N-acetylation, in which an amide bond forms between the substrate and an acetyl group. The rate at which isoniazid is acetylated determines its clinical effectiveness.

Two tuberculosis patients with similar body weights and symptoms are given the same dose of isoniazid. Although both take the drug as prescribed, one patient fails to show a significant clinical improvement. The other patient is cured. Genetic factors appear to be responsible for the differences in drug metabolism. Can you suggest a reason why the two patients reacted so differently to isoniazid? How can physicians improve the percentage of patients who are cured?

Biochemistry IN PERSPECTIVE

Alcohol Dehydrogenase: A Tale of Two Species

Why can humans make and consume alcoholic beverages? The alcohol dehydrogenases have diverse functions in living organisms. Many single-celled organisms generate energy by fermentation, a set of *anaerobic* ("without oxygen") biochemical pathways that yield organic waste products such as alcohols and acids. In the bacterium *Clostridium acetobutylicum*, the conversion of butyraldehyde and acetone into the end products butanol and isopropanol, respectively, is catalyzed by two different ADHs. In World War I, *C. acetobutylicum* was used as an alternative source of these solvents when petroleum was in short supply. Another ADH, cinnamyl alcohol dehydrogenase, converts cinnamyl aldehydes into the cinnamyl alcohols, the precursors of lignin, the rigid cell wall polymer in the woody stems of angiosperms. Of all the ADHs, however, the enzymes that most interest researchers are those of *Saccharomyces cerevisiae* and *Homo sapiens*.

Yeast Alcohol Dehydrogenases *S. cerevisiae*, a type of yeast (unicellular fungus), has a unique property that humans have exploited for thousands of years: a robust capacity for converting sugar (either glucose or fructose) into ethanol, either in the presence or in the absence of molecular oxygen (O_2). Unlike most fermenting organisms, yeasts are protected by 250 genes that code for molecules that protect against ethanol toxicity. In fermentation pathways, a six-carbon sugar molecule is split in half and converted into two molecules each of pyruvate, ATP, and NADH (**Figure 6A**). For fermentation to continue, NAD^+ (the oxidized form of NADH) must be regenerated. Fermenting organisms solve this problem by converting pyruvate into one or more reduced molecules (e.g., alcohols, ketones, or acids). In ethanol-producing yeast, pyruvate decarboxylase removes a carboxylate group from pyruvate to form acetaldehyde, and alcohol dehydrogenase (a zinc-containing tetramer) converts acetaldehyde into ethanol. The highly efficient ethanol-producing process of modern *S. cerevisiae*, which humans use in beer brewing and wine making, is believed to have originated about 80 million years ago with the emergence of angiosperm species producing fleshy fruits. As a result of fierce competition with other microorganisms for fruit sugar molecules, ancestors of *S. cerevisiae* developed the capacity to rapidly produce large amounts of ethanol as a means of killing its competitors, while protecting themselves with detoxification enzymes and membrane-stabilizing molecules. *S. cerevisiae* possesses two alcohol dehydrogenases: ADH1 (which converts acetaldehyde to ethanol) and ADH2 (which converts ethanol to acetaldehyde). The role of ADH2 is described in Chapter 8.

Human Alcohol Dehydrogenases Humans can consume moderate amounts of ethanol because of ADH. Altogether humans have seven ADHs, each of which is a zinc-containing dimer. Ethanol is primarily detoxified by three ADH isoenzymes (ADH1A, ADH1B, and ADH1C). These molecules have a high affinity for ethanol even when it is present in low concentrations. The ADH1 isoenzymes are called the liver ADHs because they are present in relatively high quantities in liver, the major site of ethanol detoxification. (The small intestine, colon, lung, and adrenal gland also possess ADH1, but in lower amounts.) The other human ADHs differ from ADH1 in their substrate specificities and/or kinetic properties. ADH2, also found in the liver and the GI tract, has a significantly lower affinity for ethanol and therefore plays only a minor role in ethanol metabolism. Among its many functions, ADH3 (the ancestral vertebrate ADH) is involved in formaldehyde detoxification and in the synthesis of retinoic acid, derived from retinol (vitamin A). ADH4 oxidizes a wide variety of substrates, which include retinol, aliphatic alcohols, and hydroxylated steroids. The functions of ADH5 are unresolved.

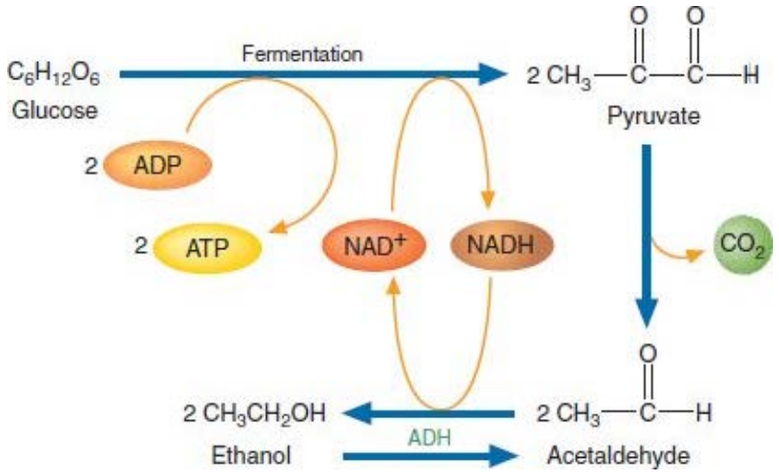


FIGURE 6A

Synthesis of Ethanol from Glucose by *S. cerevisiae*

An anaerobic biochemical pathway called fermentation yields two ATPs for each glucose molecule that is converted into two pyruvate molecules. To regenerate NAD^+, yeast cells first convert pyruvate into acetaldehyde in a reaction catalyzed by pyruvate decarboxylase. Acetaldehyde is then reversibly reduced to form ethanol and NAD^+ in a reaction catalyzed by ADH.

SUMMARY The production of alcoholic beverages by humans is possible because *S. cerevisiae* can rapidly and efficiently convert large amounts of sugar into ethanol. Humans can consume moderate amounts of toxic ethanol molecules because of the detoxifying reaction catalyzed by the liver ADH isoenzymes.

Chapter Summary

1. Enzymes are biological catalysts. They enhance reaction rates because they provide an alternative reaction pathway that requires less energy than an uncatalyzed reaction. In contrast to some inorganic catalysts, most enzymes catalyze reactions at mild temperatures. In addition, enzymes are specific to the types of reaction they catalyze. Each type of enzyme has a unique, intricately shaped binding surface called an active site. Substrate binds to the enzyme's active site, which is a small cleft or crevice in a large protein molecule. In the lock-and-key model of enzyme action, the structures of the enzyme's active site and the substrate transition state are complementary. In the induced-fit model, the protein molecule is assumed to be flexible.
2. Enzymes are classified and named according to the type of reaction each one catalyzes. There are six major enzyme categories: oxidoreductases, transferases, hydrolases, lyases, isomerases, and ligases.
3. Enzymes use the same catalytic mechanisms as nonenzymatic catalysts. Several factors contribute to enzyme catalysis: proximity and orientation effects, electrostatic effects, acid–base catalysis, and covalent catalysis. Combinations of these factors affect enzyme mechanisms.
4. Active site amino acid side chains can facilitate proton transfer and nucleophilic substitutions and stabilize the transition state. Many enzymes use nonprotein cofactors (metals and coenzymes) to facilitate reactions.
5. Enzymes are sensitive to environmental factors such as temperature and pH. Each enzyme has an optimum temperature and an optimum pH, which are determined in the laboratory.
6. Enzyme kinetics is the quantitative study of enzyme catalysis. According to the Michaelis–Menten model, when the substrate S binds in the active site of an enzyme E, an ES transition state complex is formed. During the transition state, the substrate is converted into product. After a time the product dissociates from the enzyme. The symbol V_{max} represents the maximal velocity for the reaction, and K_m is a rate constant called the Michaelis constant. Experimental determinations of K_m and V_{max} are made with Lineweaver–Burk double-reciprocal plots.
7. The turnover number (k) is a measure of the number of substrate molecules converted to product per unit

cat
time by an enzyme when it is saturated with substrate. Because [S] is relatively low under physiological conditions ([S] << K_m), the specificity constant k_{cat}/K_m is a more reliable gauge of the catalytic efficiency of enzymes.

8. Enzyme inhibition may be reversible or irreversible. In reversible inhibition, the inhibitor can dissociate from the enzyme. The most common types of reversible inhibition are competitive, noncompetitive, and uncompetitive. The assumptions of the law of mass action upon which in vitro enzyme analysis are based do not appear to hold true in the crowded and heterogeneous interior space of living organisms. Irreversible inhibitors usually bind covalently to enzymes.
9. The kinetic properties of allosteric enzymes are not explained by the Michaelis–Menten model. Most allosteric enzymes are multisubunit proteins. The binding of substrate or effector to one subunit affects the binding properties of other subunits.
10. The chemical reactions in living cells are organized into a series of biochemical pathways. The pathways are controlled primarily by adjusting the concentrations and activities of enzymes through genetic control, covalent modification, allosteric regulation, and compartmentation.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee**, where you can complete a multiple-choice quiz on water to help you prepare for exams.

Chapter 6 Review Quiz

Suggested Readings

Aqvist J, et al. 2017. Entropy and enzyme catalysis. Accounts Chem Res 50(2):199–207.

Gramser S. 2005. Alcohol and science: the party gene. Nature 438:1068–9.

Gutteridge A, Thornton JM. 2000. Understanding nature's toolkit. Trends Biochem Sci 30(11):622–9.

Petho G, Reeh PW. 2012. Sensory and signaling mechanisms of bradykinin, eicosanoids, platelet-activating factor, and nitric oxide in peripheral nociceptors. Physiol Rev 92:1699–775.

Ricciotti E, Fitzgerald GA 2011. Prostaglandins and inflammation. Arterioscler, Thromb Vasc Biol 31(5):986–1000.

Schnell S, Turner TE. 2004. Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. Prog Biophys Mol Biol 85:234–60.

Storey KB, ed. 2004. Functional metabolism: regulation and adaptation. Hoboken, NJ: Wiley-Liss.

Thomson JM, et al. 2005. Resurrecting ancestral alcohol dehydrogenases from yeast. Nat Genet. 37:630–5.

Wolfenden R, Snider MJ. 2001. The depth of chemical time and the power of enzymes as catalysts. Acc Chem Res 34(12):938–45.

Woofit M, Wolfe K. 2005. The gene duplication that greased society's wheels. Nat Genet 37:566–7.

Key Words

activation energy, *198*

active site, *199*

activity coefficient, *200*

allosteric enzyme, *231*

apoenzyme, *200*

carbanion, *205*
carbocation, *205*
catalyst, *198*
coenzyme, *200*
cofactor, *200*
competitive inhibition, *229*
enzyme, *197*
enzyme induction, *233*
enzyme kinetics, *219*
free radical, *205*
holoenzyme, *200*
hydrolase, *202*
inhibitor, *226*
intermediate, *204*
irreversible inhibition, *226*
isomerase, *202*
ligase, *202*
lyase, *202*
macromolecular crowding, *231*
negative cooperativity, *234*
noncompetitive inhibition, *227*
oxidoreductase, *201*
oxyanion, *211*
pH optimum, *215*
positive cooperativity, *234*
proenzyme, *233*
reaction mechanism, *204*
reversible inhibition, *226*
specificity constant, *222*
substrate, *198*
transferase, *201*
transition state, *198*
turnover number, *222*
uncompetitive inhibition, *228*
velocity, *218*
vitamin, *213*
zymogen, *233*

Review Questions

SECTION 6.1

Comprehension Questions

1. Define the following terms:
   a. activation energy
   b. $\Delta G^{\ddagger}$
   c. free energy of activation
   d. active site
   e. effective concentrations
2. Define the following terms:
   a. activity coefficient
   b. solute activity
   c. substrate
   d. transition state
   e. catalyst
3. Define the following terms:
   a. PTGS
   b. COX
   c. nociceptors
   d. lock-and-key model
   e. induced-fit model

Fill in the Blanks

4. Each type of enzyme contains a unique, intricately shaped binding surface called a(n) _____.
5. The _____ model states that substrate binding causes enzymes to undergo conformational change.
6. The _____ is the structural form of the substrate that has the highest energy.
7. An enzyme like all catalysts cannot alter the _____ of the reaction.
8. _____ such as dextran are used to simulate intracellular conditions in enzyme studies.
9. The _____ are a class of enzymes that catalyze the ATP-dependent phosphorylation of hexose sugars.

Short-Answer Questions

10. List four important properties of enzymes.
11. Briefly describe the induced-fit conformational change when hexokinase binds its substrate.
12. List two means by which enzyme activity is directly regulated.
13. How do reactions differ in ideal solutions and within living organisms?
14. What is meant by the term *activation energy*?
15. In many cases, an enzyme will attack one particular stereoisomer but not its enantiomer. Explain.

Critical-Thinking Questions

16. What possible effect would macromolecular crowding have on activity coefficients? Explain how macromolecular crowding can increase the effective concentration of a substance.
17. The mechanism that allows the exothermic hydrolysis of t-butyl chloride to form t-butyl alcohol and chloride ion is as follows:

$(CH_3)_3CCl \rightarrow (CH_3)_3C+ + Cl^-$

$(CH_3)_3C+ + H_2O \rightarrow (CH_3)_3COH + H^+$

Draw the transition state for each step.

SECTION 6.2

Comprehension Questions

18. Define the following terms:
    a. oxidoreductase
    b. lyase
    c. ligase
    d. transferase
    e. isomerase
19. Define the following terms:
    a. oxygenase
    b. epimerase
    c. protease
    d. hydroxylase
    e. oxidase

Fill in the Blanks

20. _____ are enzymes that catalyze oxidation-reduction reactions.
21. Hexokinase is an example of a general class of enzyme known as the _____.

SECTION 6.3

Comprehension Questions

22. Define the following terms:
    a. reaction mechanism
    b. carbocation
    c. enediol
    d. general acid
    e. general base
23. Define the following terms:
    a. catalytic dyad
    b. transition state theory
    c. hydroxide ion catalysis
    d. reaction intermediate
    e. proximity and orientation effects
24. Define the following terms:
    a. cytochrome P_{450}
    b. enzyme cofactor
    c. coenzyme

d. Lewis acid
e. pH optimum

25. Define the following terms:

a. one-carbon transfers
b. ceruloplasmin
c. Menkes syndrome
d. Wilson's disease
e. vitamins

Fill in the Blanks

26. A(n) _____ is a species that exists for a finite period of time during a reaction before it is converted to product.
27. _____ are organic molecules that provide enzymes with chemical versatility.
28. A(n) _____ is a molecular species in the chymotrypsin mechanism with a negatively charged oxygen.

Short-Answer Questions

29. Describe three functions of coenzymes in enzyme activity.
30. What special properties of transition metals make them especially useful in enzyme catalysis?
31. What is the structure of the intermediate formed during the conversion of dihydroxyacetone phosphate to glyceraldehyde-3-phosphate?
32. Quantum tunneling appears to play a significant role in the facilitation of efficient enzyme catalysis in reactions involving the transfer of protons and hydride ions. Provide an energy diagram for a hypothetical reaction in the presence and absence of an enzyme that illustrates the tunneling process. (*Hint:* Refer to **Figure 6.1** and the online Biochemistry in Perspective essay Quantum Tunneling and Catalysis.)
33. Review the structure of the standard amino acids and list those with side chains that are capable of acting as acids or bases in enzyme catalysis. Are there any that can function as both an acid and a base?

Critical-Thinking Questions

34. Review the electronic configuration of the magnesium ion and explain why Mg^{2+} can function as a Lewis acid.
35. Provide the reaction between the residue Ser 195 of chymotrypsin and diisopropylfluorophosphate. Why is this particular serine so reactive?
36. In the serine protease triad, the proximity of an aspartate carboxylate group of histidine raises the latter's pK_a. Explain.
37. When free water molecules are excluded from an enzyme's catalytic pocket, is a catalytic $-OH$ group on an amino acid side chain a stronger or weaker nucleophile? Explain.
38. Show how the arginine–histidine dyad within an active site could catalyze the dehydration of t-butyl alcohol.
39. Within an enzyme active site, how does the aspartate carboxyl group activate the histidine nitrogen and convert it into a stronger base?
40. Describe the possible effect of replacing the histidine in the triose phosphate isomerase with lysine.

SECTION 6.4

Comprehension Questions

41. Define the following terms:
    a. velocity
    b. kinetics
    c. half-life
    d. first-order reactions
    e. pseudo-first-order reaction
42. Define the following terms:
    a. Lineweaver–Burk plot
    b. allosteric enzyme
    c. macromolecular crowding
    d. zymogen
    e. metabolon
43. Define the following terms:
    a. reaction order
    b. turnover number
    c. double-displacement reactions
    d. inhibitor
    e. sigmoidal binding curve
44. Define the following terms:
    a. competitive inhibitor
    b. uncompetitive inhibitor
    c. noncompetitive inhibitor
    d. reversible inhibitor
    e. irreversible inhibitor

Fill in the Blanks

45. Molecules that reduce an enzyme's activity are called _____.
46. Reversible inhibition will be _____ if the inhibitor and substrate bind to the same active site.
47. Chemical species with unpaired electrons are called _____.
48. In _____ inhibition, the EI complex readily dissociates and the enzyme is again available for substrate binding.
49. _____ are molecules used in enzyme studies that resemble the structure of a substrate's transition state.

Short-Answer Questions

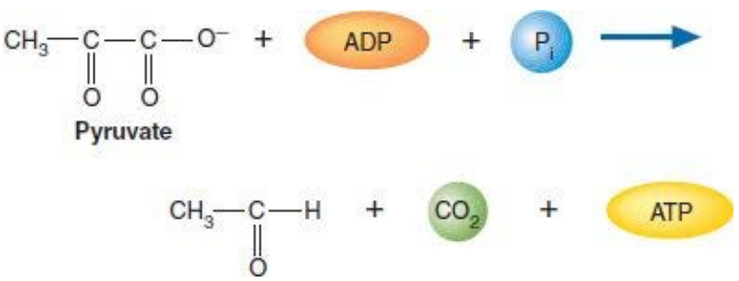
50. Suicide substrates are molecules that resemble the substrate. When these molecules enter into the active site, they form a covalent bond with the enzyme, thereby permanently inhibiting it. These molecules have been used to label the amino acid side chains that are involved in the catalytic reaction occurring within the active site. This is usually accomplished by hydrolyzing the inhibited enzyme. How could an investigator identify the amino acid bound to the suicide

substrate molecule?

51. Describe negative feedback inhibition and its significance in metabolism.
52. What properties of transition metals make them useful as enzyme cofactors?
53. In enzyme kinetics, why are measurements made at the start of a reaction?
54. What are two major types of enzyme inhibitors? Give an example of each.
55. What term describes k_{cat}/K_m? What is the maximum value that this term can attain? Explain.
56. List three effects of macromolecular crowding on the properties of enzymes and the reactions they catalyze.
57. When calculating the order of a reaction, the enzyme does not appear in the equation. Explain.
58. How does an enzyme attain catalytic perfection?
59. Mercuric ion and methanol are inhibitors of alcohol dehydrogenase. Explain.

Critical-Thinking Questions

60. Consider the following reaction:



Using the following data, determine the order of the reaction for each substrate and the overall order of the reaction.

	Concentration (mol/L)			Rate
Experiment	Pyruvate	ADP	P_1	(mol L^{-1} s^{-1})
1	0.1	0.1	0.1	8×10^{-4}
2	0.2	0.1	0.1	1.6×10^{-3}
3	0.2	0.2	0.1	3.2×10^{-3}
4	0.1	0.1	0.2	3.2×10^{-3}

61. Consider the following data for an enzyme-catalyzed hydrolysis reaction in the presence and absence of inhibitor I:

[Substrate] (M)	v_0 (μmol/min)	v_{0I} (μmol/min)
6×10^{-6}	20.8	4.2
1×10^{-5}	29	5.8
2×10^{-5}	45	9
6×10^{-5}	67.6	13.6
1.8×10^{-4}	87	16.2

Using a Michaelis–Menten plot, determine K_m for the uninhibited reaction and the inhibited reaction.

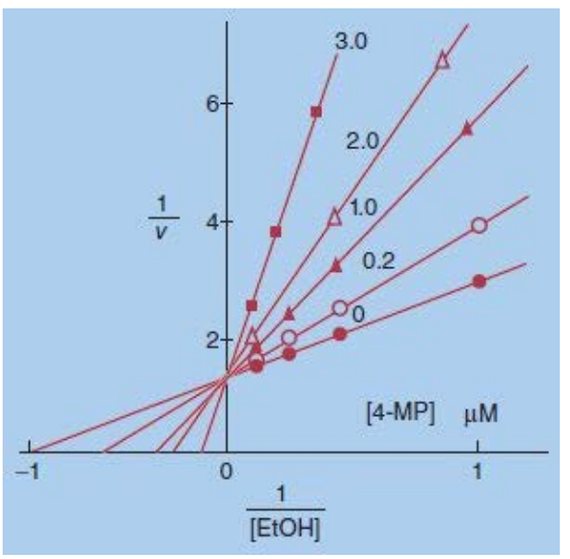
62. Use the data in Question 61 to do the following:

    a. Generate Lineweaver–Burk plots for the data.
    b. Explain the significance of the horizontal intercept, the vertical intercept, and the slope.
    c. Identify the type of inhibition being measured.

63. Two experiments were performed with the enzyme ribonuclease. In experiment 1, the effect of increasing substrate concentration on reaction velocity was measured. In experiment 2, the reaction mixtures were identical to those in experiment 1 except that 0.1 mg of an unknown compound was added to each tube. Plot the data according to the Lineweaver–Burk method. Determine the effect of the unknown compound on the enzyme's activity. (Substrate concentration is measured in millimoles per liter. Velocity is measured in the change in optical density per hour.)

Experiment 1		Experiment 2	
[S]	v	[S]	v
0.5	0.81	0.5	0.42
0.67	0.95	0.67	0.53
1	1.25	1	0.71
2	1.61	2	1.08

64. 4-methyl pyrazole (4-MP) has been developed as a long-acting and less toxic alternative to ethanol in the treatment of ethylene glycol poisoning. Shown is a Lineweaver–Burk plot of the inhibition of alcohol dehydrogenase by various concentrations of 4-methyl pyrazole. What type of inhibition does this molecule appear to exhibit?



65. The following table presents the rates of reaction at specific substrate concentrations for an enzyme that displays classical Michaelis–Menten kinetics. Two sets of inhibitor data are also included. Determine the K_m and V_{max} for the uninhibited enzyme.

[S] (mM)	**Without inhibitor**	**v_0 (mM/s) With inhibitor A**	**With inhibitor B**
1.3	2.50	1.17	0.62

2.6	4.00	2.10	1.42
6.5	6.30	4.00	2.65
13.0	7.60	5.70	3.12
26.0	9.00	7.20	3.58

66. Determine the type of inhibition exhibited by each inhibitor in Question 65. What kind of affinity would these inhibitors have for ES versus E?
67. Catalase has a K_m of 25 mM and a k_{cat} of 4.0 x 10^7 s^{-1} with H_2O_2 as a substrate. Carbonic anhydrase has a K_m of 26 mM and a k_{cat} of 4.0×10^5 s^{-1}. What do these data tell you about these two enzymes?
68. The K_m and k_{cat} for fumarase with fumarate as a substrate are 5×10^{-6} M and 8×10^2 s^{-1}, respectively. When malate is the substrate, the K_m and k_{cat} are 2.5×10^{-5} M and 9×10^2 s^{-1}, respectively. What do these data tell you about the operation of this enzyme in the citric acid cycle?
69. Given the equation for competitive inhibition:

$$\nu = \frac{V_{max}[S]}{\alpha K_m + [S]}$$

a. Can the value of α ever be less than 1?
b. What would happen if the value of α were less than 1?

70. Consider the following reaction along with its rate information.

$$A + B \rightarrow C$$

[A] (mM)	[B] (mM)	Rate (mM/s)
0.05	0.05	2×10^7
0.10	0.05	4×10^7
0.05	0.1	4×10^7
0.1	0.1	8×10^7

What is the overall rate expression for the reaction? What is the order of the reaction?

71. Review the following approximate K_m values for ethanol and acetaldehyde for ADH1 in *Saccharomyces cerevisiae* and explain why these data support the observed role of this enzyme in the organism's ethanol metabolism.

	K_m (ethanol)	K_m (acetaldehyde)
ADH1	20,000 μM	1500 μM

72. Ethylene glycol (HO—CH_2—CH_2—OH) is frequently used as antifreeze in automobile engines. Every year children and pets are poisoned because they tasted this sweet-tasting material. Ethylene glycol is metabolized in the liver by alcohol dehydrogenase. Suggest a possible medical treatment for ethylene glycol intoxication.
73. Given the following rate expression, complete the following table.

[A] (mM)	Rate = k[A]²[B] [B] (mM)	Rate (mM/s)

0.1	0.01	1×10^6
0.1	0.02	_____
0.2	0.01	_____
0.2	0.02	_____

SECTION 6.5

Comprehension Questions

74. Define the following terms:
    a. genetic control
    b. proenzyme
    c. zymogen
    d. positive cooperativity
    e. negative cooperativity
75. Define the following terms:
    a. allosteric regulation
    b. homotropic effects
    c. heterotropic effects
    d. concerted model
    e. sequential model
76. Define the following terms:
    a. enzyme induction
    b. covalent modification
    c. nucleotidylation
    d. allosteric site
    e. compartmentation

Fill in the Blanks

77. The _____ and _____ models are theoretical descriptions of allosteric enzymes.
78. _____ allosteric effects of enzymes involve ligands that are different from substrate molecules.
79. CTP inhibition of aspartate transcarbamoylase (ATCase) is an example of _____ inhibition.
80. The binding of the first O_2 to hemoglobin initiates a concerted _____ transition.

Short-Answer Questions

81. Describe the role that compartmentation plays in the regulation of metabolic pathways. Provide several examples.
82. What is the significance of metabolic flux?
83. List three types of chemical reactions that prepare drug molecules for excretion from the body.

Critical-Thinking Questions

84. Describe the alcohol dehydrogenases of humans and *S. cerevisiae*. How do they differ?
85. Alcohol dehydrogenase (ADH) is inhibited by numerous alcohols. Using the data given in the following table, calculate the k_{cat}/K_m values for each of the alcohols. Which of the listed alcohols is most easily metabolized by ADH?

Kinetic Parameters for Hamster Testes ADH

Substrate	K_m (μM)	k_{cat} (min^{-1})
Ethanol	960	480
1-Butanol	440	450
1-Hexanol	69	182
12-Hydroxydodecanoate	50	146
All-*trans*-retinol	20	78
Benzyl alcohol	410	82
2-Butanol	250,000	285
Cyclohexanol	31,000	122

86. In the blood coagulation pathway, several proenzymes are activated by unique proteolytic cleavage reactions. Describe the significance of this phenomenon.
87. Venomous snakes produce a wide variety of enzymes that digest their victims. Provide several examples of tissue molecules these enzymes digest. What protects snakes from these enzymes?

MCAT Study Questions

88. The conversion of hydrogen peroxide (H_2O_2) to water and molecular oxygen is catalyzed by catalase. To which of the following enzyme categories does it belong?

    a. hydroxylase
    b. lyase
    c. oxidoreducase
    d. transferase

89. Which of the following is *least* likely to be found in the functioning of a five-reaction metabolic pathway?

    a. A T- to R-transition in the first enzyme of the pathway
    b. Sodium ion as an enzyme cofactor
    c. A competitive inhibitor from a parallel pathway binding to an allosteric enzyme
    d. An isomerase enzyme

90. Enzymes increase a reaction's rate by

    a. changing its equilibrium constant to a more favorable value.
    b. using specific prosthetic groups
    c. decreasing the reaction's activation energy
    d. decreasing the reaction's ΔG value

91. Organic molecules required by enzymes to enhance catalysis are called

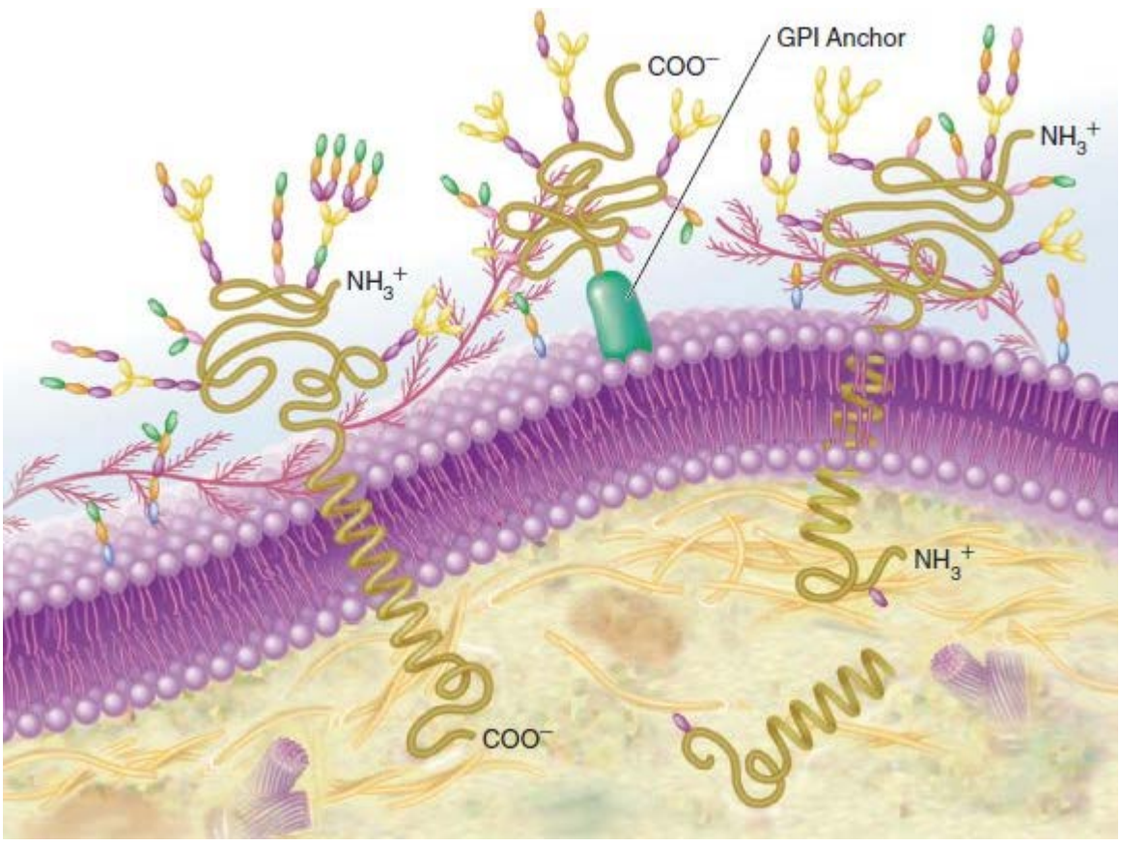
    a. coenzymes

b. zymogens
c. proenzymes
d. None of the above

92. The K_m value of an enzyme-catalyzed reaction and its V_{max} is 70 mmol/min. What is the rate of the reaction of the reaction when the substrate concentration is 7×10^{-2} mmol/min?

a. 35 mmol/min
b. 50 mmol/min
c. 60 mmol/min
d. 70 mmol/min

CHAPTER 7 Carbohydrates



Sugars and Cells Sugar molecules shape the molecular landscape of living organisms. Carbohydrates linked to membrane proteins and lipids are especially prominent on the external surface of cells.

OUTLINE

SWEET AND BITTER TASTE: THE ROLES OF SUGAR MOLECULES

7.1 MONOSACCHARIDES

Monosaccharide Stereoisomers
Cyclic Structure of Monosaccharides
Reactions of Monosaccharides
Important Monosaccharides
Monosaccharide Derivatives

7.2 DISACCHARIDES

7.3 POLYSACCHARIDES
Homoglycans
Heteroglycans

7.4 GLYCOCONJUGATES
Proteoglycans
Glycoproteins

7.5 THE SUGAR CODE
Lectins: Translators of the Sugar Code
The Glycome

AVAILABLE ONLINE

Biochemistry in Perspective
Scurvy and Ascorbic Acid

Biochemistry in Perspective
Sweet Medicine

Biochemistry in Perspective
Conversion of Fischer Structures into Haworth Structures

Biochemistry in the Lab
Glycomics

Sweet and Bitter Taste: The Roles of Sugar Molecules

For several hundred million years, plants and plant-eating animals (called herbivores) have been engaged in a struggle for survival. Herbivores derive energy and nutrients from plant biomass, the product of photosynthesis. To protect themselves, many plant species synthesize toxic molecules that either deter or kill herbivores. Examples include oxalate ($C_2O_4^{2-}$) crystals, which cause intense burning and irritation in the mouth, and strychnine, which causes muscular convulsions and asphyxiation by blocking a chloride channel in the spinal cord and brain. Although many animals have evolved several types of defenses against plant toxins (e.g., inactivating biotransformation reactions), their principal means of protection is the sense of taste. Taste, mediated by taste receptor cells within taste buds, allows an animal to quickly evaluate what it is eating. The binding of food molecules to taste receptors initiates the depolarization of taste receptor cell membranes. In humans, taste signals, carried via action potentials in the sensory nerve fibers in close proximity to taste receptor cells, are ultimately transmitted to the gustatory cortex in the frontal lobe of the brain. Chemosensory neurons within the cerebral cortex are responsible for taste perception.

Most animals have five primary types of taste perception: sweet, savory, bitter, sour, and salty. Each allows animals to identify specific nutrients or physiological threat. Sweet and savory tastes, for example, indicate that a food is a rich source of sugars or amino acids, respectively. These tastes are pleasurable and promote feeding behavior. Avoidance behavior is triggered by bitter taste, which indicates toxicity, and high levels of salty and sour tastes, which indicate the presence of electrolytes and acids in food. Of the five taste types, sweet and bitter are the most important in evaluating food energy content and safety. Sweet receptor cells detect sugars at 0.01 M or above so that only energy-rich foods will elicit eating behavior. In contrast, the threshold for bitter taste is very low. Quinine, an example of the alkaloids (a group of nitrogen-containing plant molecules with potent physiological properties) from the Cinchona tree, is detected at 0.8 μM.

One of the more intriguing features of sweet and bitter taste perception is that sugar molecules are not only sweet, energy-rich nutrients; they can also be components of bitter, toxic molecules. In two prominent classes of potentially lethal plant toxins, the cyanogenic and cardiac glycosides, one or more sugar molecules are covalently linked to a toxic aglycone (the noncarbohydrate portion of a molecule). (The sugar residues improve the water solubility of the aglycone.) *Cyanogenic glycosides* are rapidly hydrolyzed to form hydrogen cyanide (HCN; a potent inhibitor of aerobic respiration; see p. 373) when consumed by animals. *Cardiac glycosides* inhibit the Na^+/K^+-ATPase (p. 429) in the membrane of cells, resulting in nausea, dizziness, confusion, and cardiac arrest. Humans have learned how to render some plants safe to eat by inactivating toxins. For example, roasting the roots of the cassava plant, an important food source for many humans, destroys cyanogenic glycosides such as linamarin. Some plant toxins have been used for their medicinal properties. In small and carefully regulated amounts, digitoxin, a cardiac glycoside from the foxglove plant, has been used for hundreds of years to treat congestive heart failure.

Not all bitter plant molecules are toxic. In fact, numerous bitter molecules in various plant-derived foods and beverages have beneficial effects on human health. Examples include bioflavonoids (polyhydroxypolyphenols) in tea and citrus fruits that protect against cell-damaging free radicals and organosulfur compounds in cruciferous vegetables such as broccoli that have anticancer properties.

Overview

CARBOHYDRATES ARE NOT JUST AN IMPORTANT SOURCE OF RAPID ENERGY PRODUCTION FOR LIVING CELLS. THEY ARE ALSO STRUCTURAL building blocks of cells and components of numerous metabolic pathways. Sugar polymers linked to proteins and lipids are now recognized as a high-density coding system. Their vast structural diversity is exploited by living organisms to produce the immense informational capacity required for living processes. Chapter 7 describes the structures and chemistry of typical carbohydrate molecules found in living organisms. An introduction to glycomics, the investigation of the sugar code, is available online.

Carbohydrates, the most abundant biomolecules in nature, are a direct link between solar energy and the chemical bond energy of living organisms. (More than half of all “organic” carbon is found in carbohydrates.) They are formed during *photosynthesis* (Chapter 13), a biochemical process in which light energy is captured and used to drive the biosynthesis of energy-rich organic molecules from the energy-poor inorganic molecules CO_2 and H_2O. Most carbohydrates contain carbon, hydrogen, and oxygen in the ratio $(CH_2O)_n$—hence the name “hydrate of carbon.” They have been adapted for a wide variety of biological functions, which include energy sources (e.g., glucose), structural elements (e.g., cellulose and chitin in plants and insects, respectively), cellular communication and identity, and precursors in the production of other biomolecules (e.g., amino acids, lipids, purines, and pyrimidines). Carbohydrates are classified as monosaccharides, disaccharides, oligosaccharides, and polysaccharides according to the number of simple sugar units they contain. Carbohydrate moieties also occur as components of other biomolecules. A vast array of *glycoconjugates* (protein and lipid molecules with covalently linked carbohydrate groups) is distributed among all living species, most notably among the eukaryotes. The sugar molecules ribose and deoxyribose are structural elements of nucleotides and nucleic acids.

Chapter 7 provides a foundation for understanding the complex processes in living organisms by reviewing the structure and function of the most common carbohydrates and glycoconjugates. The chapter ends with a discussion of the *sugar code*, the mechanism by which carbohydrate structure is used to encode biological information.

7.1 MONOSACCHARIDES

Monosaccharides, or simple sugars, are polyhydroxy aldehydes or ketones. Recall from Chapter 1 that monosaccharides with an aldehyde functional group are called **aldoses**, whereas those with a ketone group are called **ketoses** (**Figure 7.1**). The simplest aldose and ketose are glyceraldehyde and dihydroxyacetone, respectively (**Figure 7.2**). Sugars are also classified according to the number of carbon atoms they contain. For example, the smallest sugars, called *trioses*, contain three carbon atoms. Four-, five-, and six-carbon sugars are called *tetroses*, *pentoses*, and *hexoses*, respectively. The most abundant monosaccharides found in living cells are the pentoses and hexoses. Often, class names such as aldohexoses and ketopentoses, which combine information about carbon number and functional groups, describe monosaccharides. For example, glucose, a six-carbon, aldehyde-containing sugar, is an aldohexose.

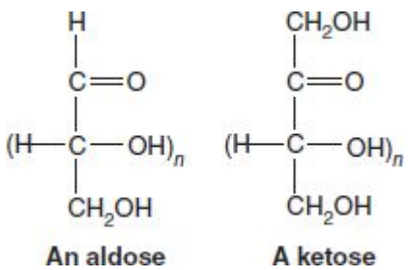


FIGURE 7.1

General Formulas for the Aldose and Ketose Forms of Monosaccharide

3D animation of aldose

3D animation of ketose

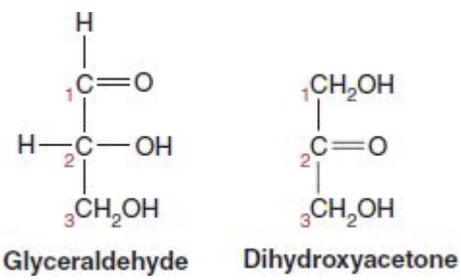


FIGURE 7.2

Glyceraldehyde (an Aldotriose) and Dihydroxyacetone (a Ketotriose)

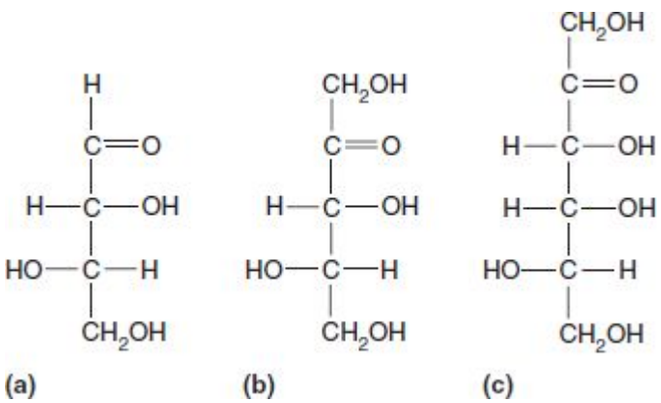
3D animation of glyceraldehyde

3D animation of dihydroxyacetone

The sugar structures shown in Figures 7.1 and 7.2 are known as Fischer projections (in honor of the German chemist Emil Fischer, 1852–1919). In these structures, the carbohydrate backbone is drawn vertically, with the most highly oxidized carbon usually shown at the top. The horizontal lines are understood to project toward the viewer; the vertical lines recede from the viewer.

QUESTION 7.1

Identify the class of each of the following sugars. For example, glucose is an aldohexose.



Monosaccharide Stereoisomers

When the number of chiral carbon atoms increases in optically active compounds, the number of possible optical isomers also increases. The total number of possible isomers can be determined using the van't Hoff's rule: A compound with n chiral carbon atoms has a maximum of 2^n possible stereoisomers. For example, when n is 4, there are 2^4 or 16 stereoisomers (8 D-stereoisomers and 8 L-stereoisomers).

In optical isomers, the reference carbon is the asymmetric carbon that is most remote from the carbonyl carbon. Its configuration is similar to that of the asymmetric carbon in either D- or L-glyceraldehyde. Almost all naturally occurring sugars have the D-configuration. They can be considered to be derived from either the triose D-glyceraldehyde (the aldoses) or the triose dihydroxyacetone (the ketoses). (Note that although dihydroxyacetone does not have an asymmetric carbon, it clearly is the parent compound for the ketoses.) In the D-aldose family of sugars (**Figure 7.3**), which contains most biologically important monosaccharides, the hydroxyl group is to the right on the chiral carbon atom in the Fischer model farthest from the most oxidized carbon (in this case, the aldehyde group) in the molecule (e.g., carbon 5 in a six-carbon sugar).

Stereoisomers that are not enantiomers (mirror-image isomers; see p. 141) are called

diastereomers. For example, the aldopentoses D-ribose and L-ribose are enantiomers, as are D-arabinose and L-arabinose (**Figure 7.4**). The sugars D-ribose and D-arabinose, which are isomers but not mirror images, are diastereomers. Diastereomers that differ in the configuration at a single asymmetric carbon atom are called **epimers**. For example, D-glucose and D-galactose are epimers because their structures differ only in the configuration of the OH group at carbon 4 (**Figure 7.3**). D-Mannose and D-galactose are not epimers because their configurations differ at more than one carbon.

Cyclic Structure of Monosaccharides

Sugars that contain four or more carbons exist primarily in cyclic forms. Ring formation occurs in aqueous solution when aldehyde and ketone groups react reversibly with hydroxyl groups present in the sugar to form cyclic **hemiacetals** and **hemiketals**, respectively. Ordinary hemiacetals and hemiketals, which form when molecules containing an aldehyde or ketone functional group react with an alcohol (p. P-27), are unstable and easily revert to the aldehyde or ketone forms (**Figure 7.5**). When the aldehyde or ketone group and the alcohol functional group are part of the same molecule, however, an intramolecular cyclization reaction occurs that can form stable products. The most stable cyclic hemiacetal and hemiketal rings contain five or six atoms. As cyclization occurs, the carbonyl carbon becomes a new chiral center. This carbon is called the *anomeric carbon atom*. The two possible diastereomers that may form during the cyclization reaction are called **anomers**.

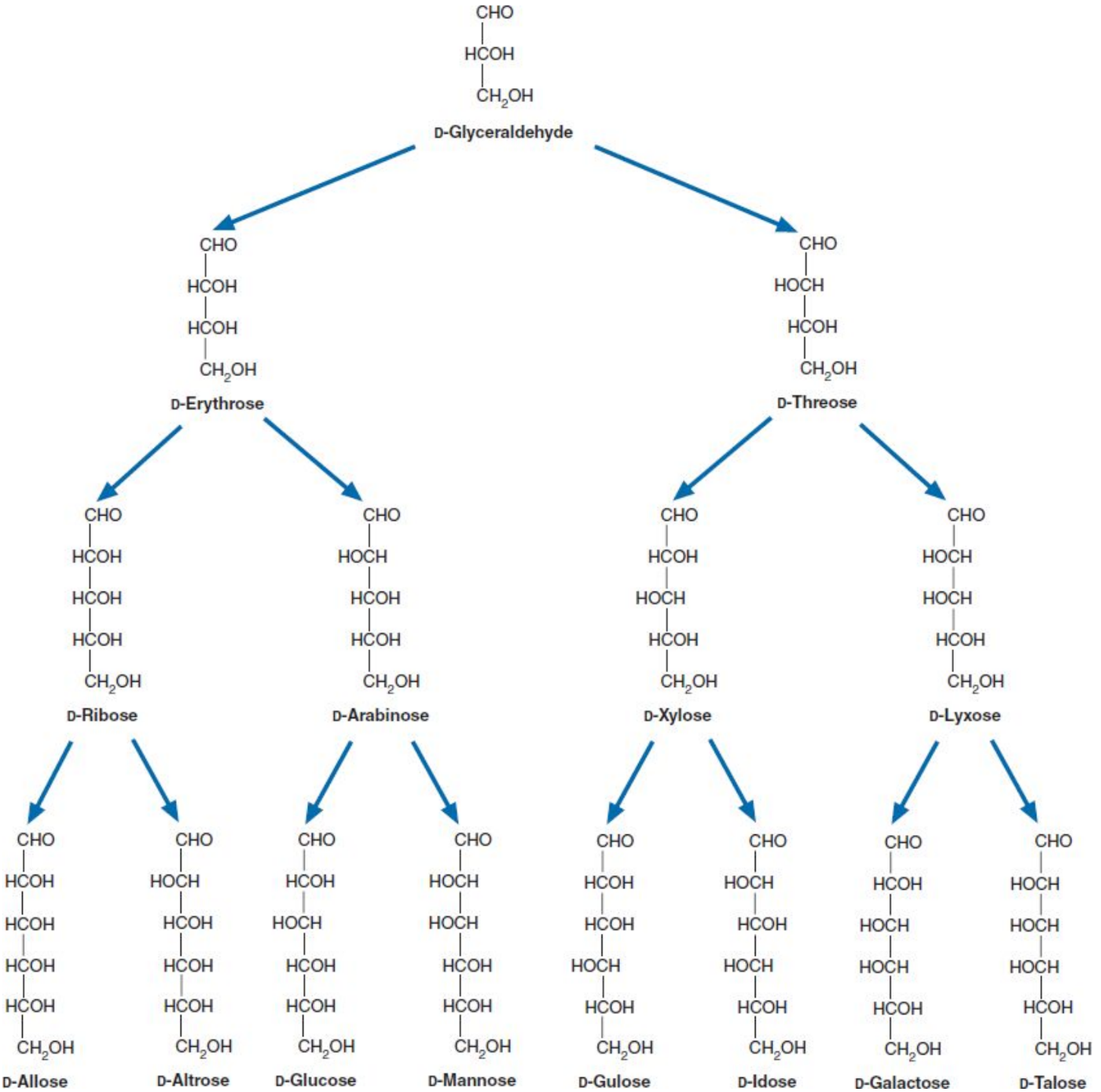


FIGURE 7.3

The D Family of Aldoses

3D animation of D-glyceraldehyde

3D animation of D-Erythrose

3D animation of ribose

In aldose sugars, the hydroxyl group of the newly formed hemiacetal occurs on carbon 1 (the anomeric carbon). Because the anomeric carbon is chiral, two stereoisomers of the aldose can form: either the α-anomer or the β-anomer. In Fischer projections, the α-anomeric hydroxyl occurs on the right and the β-hydroxyl on the left (**Figure 7.6**). It is important to note that the anomers are defined relative to the D- and L-classification of sugars. These rules apply only to D-sugars, the most common ones found in nature. In the L-sugars, the α-anomeric OH group occurs on the left. The cyclization of sugars is more easily visualized using Haworth structures.

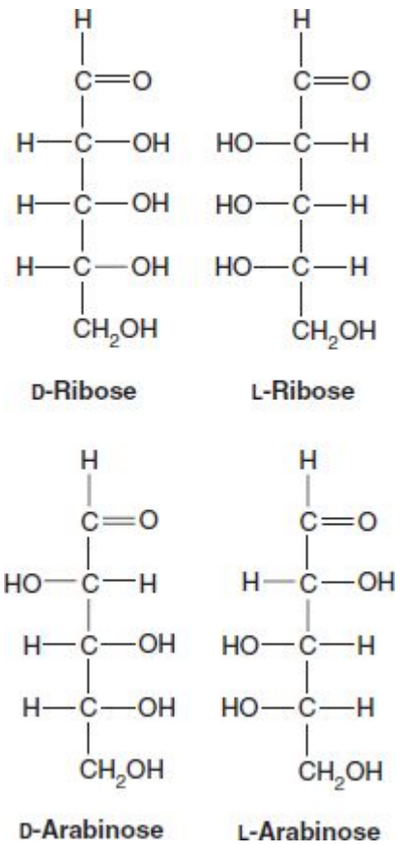


FIGURE 7.4

The Optical Isomers D- and L-Ribose and D- and L-Arabinose

D-Ribose and D-arabinose are diastereomers; that is, they are not mirror images.

3D animation of D-Arabinose

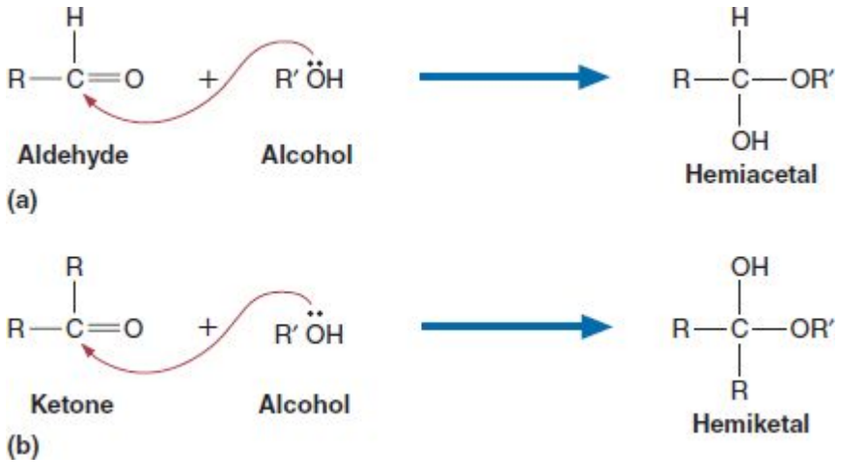


FIGURE 7.5

Formation of Hemiacetals and Hemiketals

(a) From an aldehyde. (b) From a ketone.

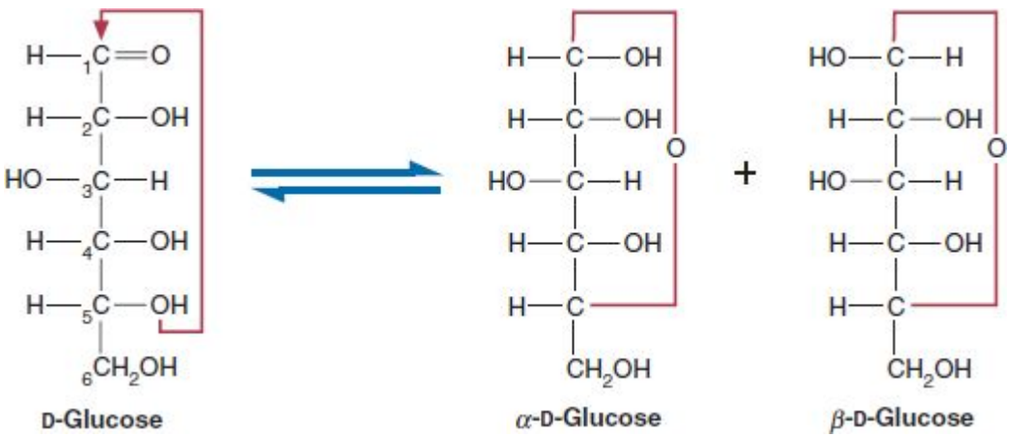


FIGURE 7.6

Monosaccharide Structure

Both the α- and the β-anomers may result when glucose forms a hemiacetal. Note that the right angles in the hemiacetal linkage in Fischer models of monosaccharides do not represent methylene groups.

KEY CONCEPTS

- Monosaccharides, which may be polyhydroxy aldehydes or ketones, are either aldoses or ketoses.
- Sugars that contain four or more carbons primarily have cyclic forms.
- Cyclic aldoses or ketoses are hemiacetals and hemiketals, respectively.

HAWORTH STRUCTURES Carbohydrate structure illustrations, developed by the English chemist W. N. Haworth (**Figure 7.7**), more closely depict proper bond angles and lengths than do Fischer representations. In Haworth depictions of aldoses, the hydroxyl group on the anomeric carbon occurs either above the ring on the same side as the CH_2OH group (the "up" position) or below the ring on the side opposite the CH_2OH group (the "down" position). For D-sugars, when the hydroxyl is down, the structure is in the α-anomeric form; if the hydroxyl is up, the structure is in the β-anomeric form. In the L-sugars, this rule is reversed: the α-anomeric OH group is above the ring, and the β-anomeric OH group is below the ring. (Refer to the online essay that describes

how Fischer projections can be converted to Haworth structures.)

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay on conversion of Fischer structures into Haworth structures.

Five-membered hemiacetal rings are called *furanoses* because of their structural similarity to furan (**Figure 7.8**). For example, the cyclic form of fructose depicted in **Figure 7.9** is called fructofuranose. Six-membered rings are called *pyranoses* because of their similarity to pyran. Glucose, in the pyranose form, is called glucopyranose.

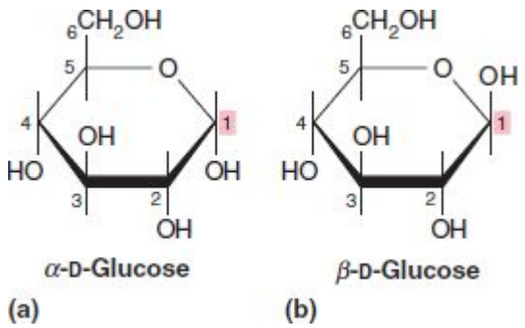


FIGURE 7.7

Haworth Structures of the Anomers of D-Glucose

(a) α-D-Glucose. (b) β-D-Glucose. Note that in carbohydrate chemistry, the hydrogens bonded to carbons in sugar rings can be represented by single lines.

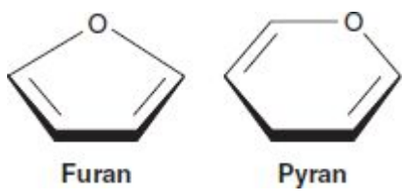


FIGURE 7.8

Furan and Pyran Structure

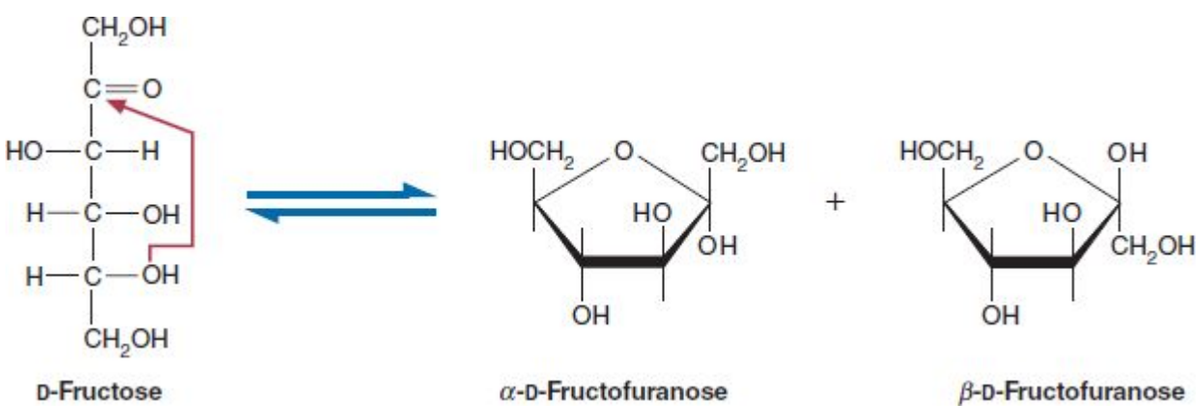


FIGURE 7.9

Fischer and Haworth Representations of D-Fructose Structure

3D animation of fructose, d-

CONFORMATIONAL STRUCTURES Although Haworth projection formulas are often used to represent carbohydrate structure, they are oversimplifications. Bond angle analysis and X-ray analysis demonstrate that *conformational formulas* are more accurate representations of monosaccharide structure (**Figure 7.10**) because they illustrate the puckered nature of sugar rings.

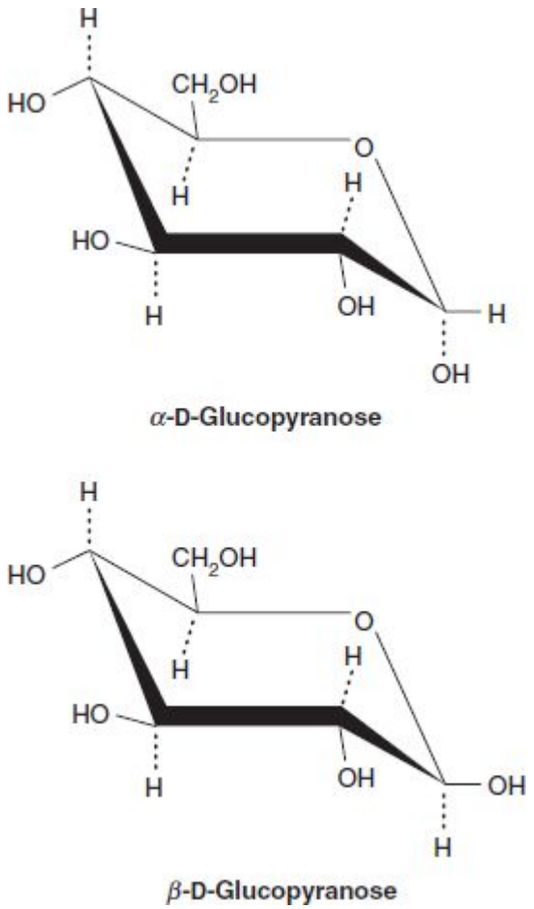


FIGURE 7.10

Conformational Representations of *α*- and *β*-D-Glucopyranose

Space-filling models, whose dimensions are proportional to the van der Waals radius of the atoms, also give useful structural information (see later: Figures 7.21, 7.22, and 7.23).

MUTAROTATION The *α* and *β* forms of monosaccharides are readily interconverted when dissolved in water. This spontaneous process, called **mutarotation**, produces an equilibrium mixture of *α* and *β* forms in both furanose and pyranose ring structures. The proportion of each form differs according to sugar type. Glucose, for example, exists primarily as a mixture of *α* (38%) and *β* (62%) pyranose forms (**Figure 7.11**). Fructose is predominantly found in the *α*- and *β*-furanose forms. The open chain formed during mutarotation can participate in oxidation-reduction reactions.

Reactions of Monosaccharides

The carbonyl and hydroxyl groups of sugars can undergo reactions that are typical of aldehydes, ketones, and alcohols. Among the most important are oxidation, reduction, isomerization, esterification, glycosylation reactions.

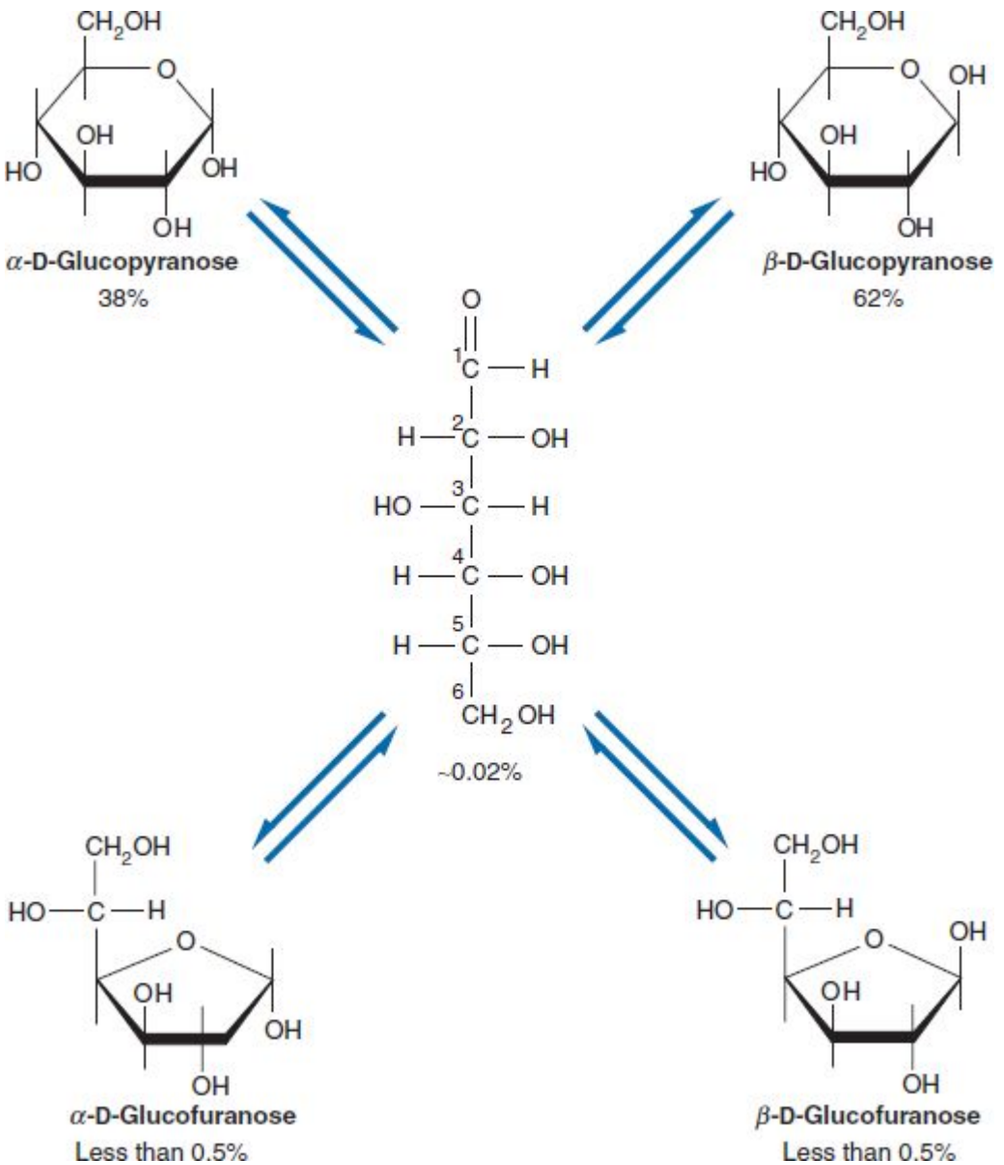
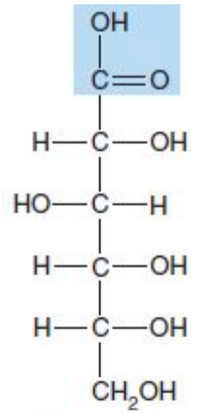


FIGURE 7.11

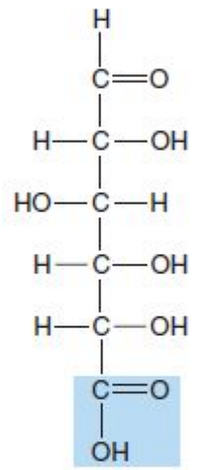
Equilibrium Mixture of D-Glucose

When glucose is dissolved in water at 25°C, the anomeric forms of the sugar undergo very rapid interconversions. When equilibrium is reached (i.e., there is no net change in the occurrence of each form), the glucose solution contains the percentages shown.

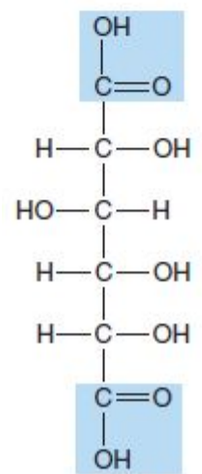
OXIDATION In the presence of oxidizing agents, metal ions such as Cu^{2+}, and certain enzymes, monosaccharides readily undergo several oxidation reactions. Oxidation of an aldehyde group yields an **aldonic acid**, whereas oxidation of a terminal CH_2OH group (but not the aldehyde group) gives a **uronic acid**. Oxidation of both the aldehyde and CH_2OH gives an **aldaric acid** (**Figure 7.12**).



D-Gluconic acid



D-Glucuronic acid



D-Glucaric acid

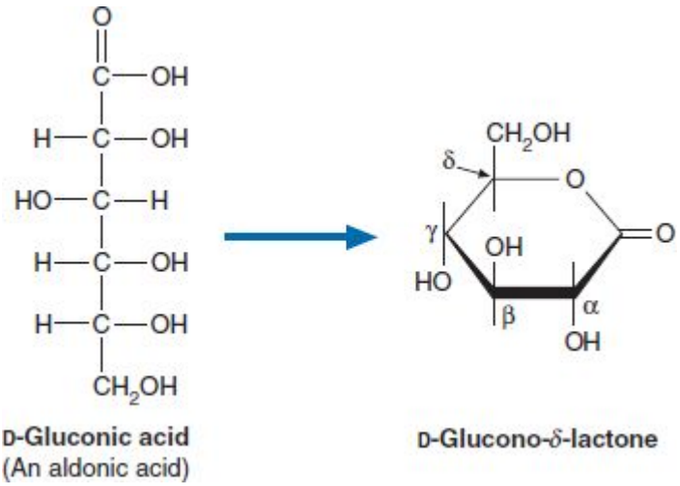
FIGURE 7.12

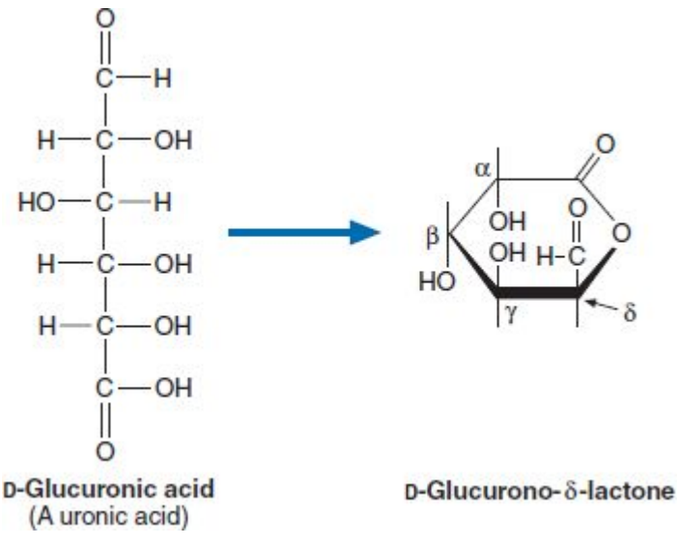
Oxidation Products of Glucose

The newly oxidized groups are highlighted.

The carbonyl groups in both aldonic and uronic acids can react with an OH group in the same

molecule to form a cyclic ester known as a **lactone**:





Lactones are commonly found in nature. For example, L-ascorbic acid, also known as vitamin C (**Figure 7.13**), is a lactone derivative of D-glucuronic acid. It is synthesized by all mammals except guinea pigs, apes, fruit-eating bats, and, of course, humans. These species must obtain ascorbic acid in their diet, hence the name vitamin C. Ascorbic acid is a powerful reducing agent; that is, it protects cells from highly reactive oxygen and nitrogen species (see p. 385). In addition, it is an enzyme cofactor required in the hydroxylation of proline and lysine residues in collagen, and in the synthesis of carnitine (p. 452) and the neurotransmitter dopamine (p. 542).

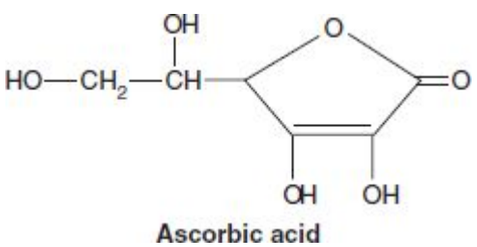


FIGURE 7.13

Structure of Ascorbic Acid

Humans and guinea pigs cannot synthesize ascorbic acid because they lack gluconolactone oxidase, one of the

three enzymes required to synthesize the acid from its precursor glucuronate.

Sugars that can be oxidized by weak oxidizing agents such as Benedict's reagent are called **reducing sugars** (**Figure 7.14**). The reaction occurs only with sugars that re-form an aldehyde group when they revert to the open-chain form. All aldoses are, therefore, reducing sugars. Ketoses such as fructose are reducing sugars because they convert to aldoses via isomerization reactions (see below).

REDUCTION Reduction of the aldehyde and ketone groups of monosaccharides yields the sugar alcohols (**alditols**). Reduction of D-glucose, for example, yields D-glucitol, also known as D-sorbitol (**Figure 7.15**). Sugar alcohols are used commercially in processing foods and pharmaceuticals. Sorbitol, for example, improves the shelf life of candy because it helps prevent moisture loss. Adding sorbitol syrup to artificially sweetened canned fruit reduces the unpleasant aftertaste of the artificial sweetener saccharin. Sorbitol is converted into fructose in the liver.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay on scurvy and ascorbic acid.

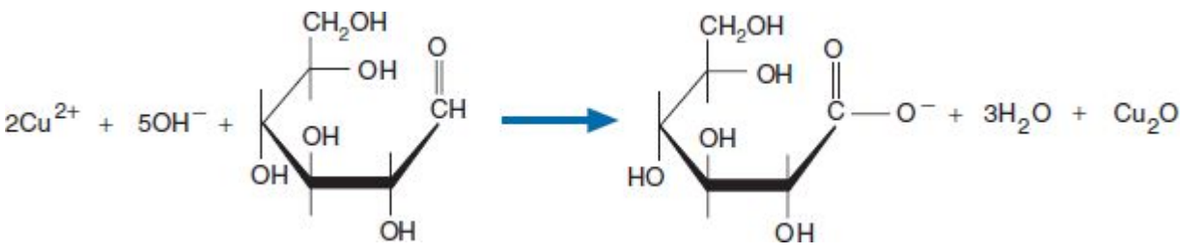
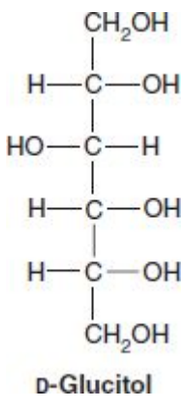


FIGURE 7.14

Reaction of Glucose with Benedict's Reagent

Benedict's reagent, copper (II) sulfate in a solution of sodium carbonate and sodium citrate, is reduced by the monosaccharide glucose. Glucose is oxidized to form the salt of gluconic acid. The reaction also forms the reddish-brown precipitate Cu_2O and other oxidation products (not shown).



D-Glucitol

FIGURE 7.15

Structure of D-Glucitol (Sorbitol)

3D animation of D-Glucitol

ISOMERIZATION Monosaccharides undergo several types of isomerization. For example, after several hours an alkaline solution of D-glucose also contains D-mannose and D-fructose. Both isomerizations involve an intramolecular shift of a hydrogen atom and a relocation of a double bond (**Figure 7.16**). The intermediate formed is an **enediol**. The reversible transformation of glucose to fructose is an aldose–ketose interconversion. Because the configuration at a single asymmetric carbon changes, the conversion of glucose to mannose is referred to as an **epimerization**. Several enzyme-catalyzed reactions involving enediols occur in carbohydrate metabolism (Chapter 8).

ESTERIFICATION Like all free OH groups, those of carbohydrates can be converted to esters by reactions with acids. Esterification often dramatically changes a sugar's chemical and physical properties. Phosphate and sulfate esters of carbohydrate molecules are among the most common ones found in nature.

Phosphorylated derivatives of certain monosaccharides are metabolic components of living cells that are frequently formed during reactions with ATP. They are important because many biochemical transformations use nucleophilic substitution reactions. Such reactions require a leaving group. In a carbohydrate molecule, this group is most likely to be an OH group. However, because OH groups are poor leaving groups, any substitution reaction is unlikely. Converting an appropriate OH group to a phosphate ester, which can then be displaced by an incoming nucleophile, solves the problem. As a consequence, a slow reaction now occurs much more rapidly.

Sulfate esters of carbohydrate molecules are found predominantly in the proteoglycan components of connective tissue (pp. 266–68). Sulfate esters are charged so that they bind large amounts of water and small ions. They also participate in forming salt bridges between carbohydrate chains.

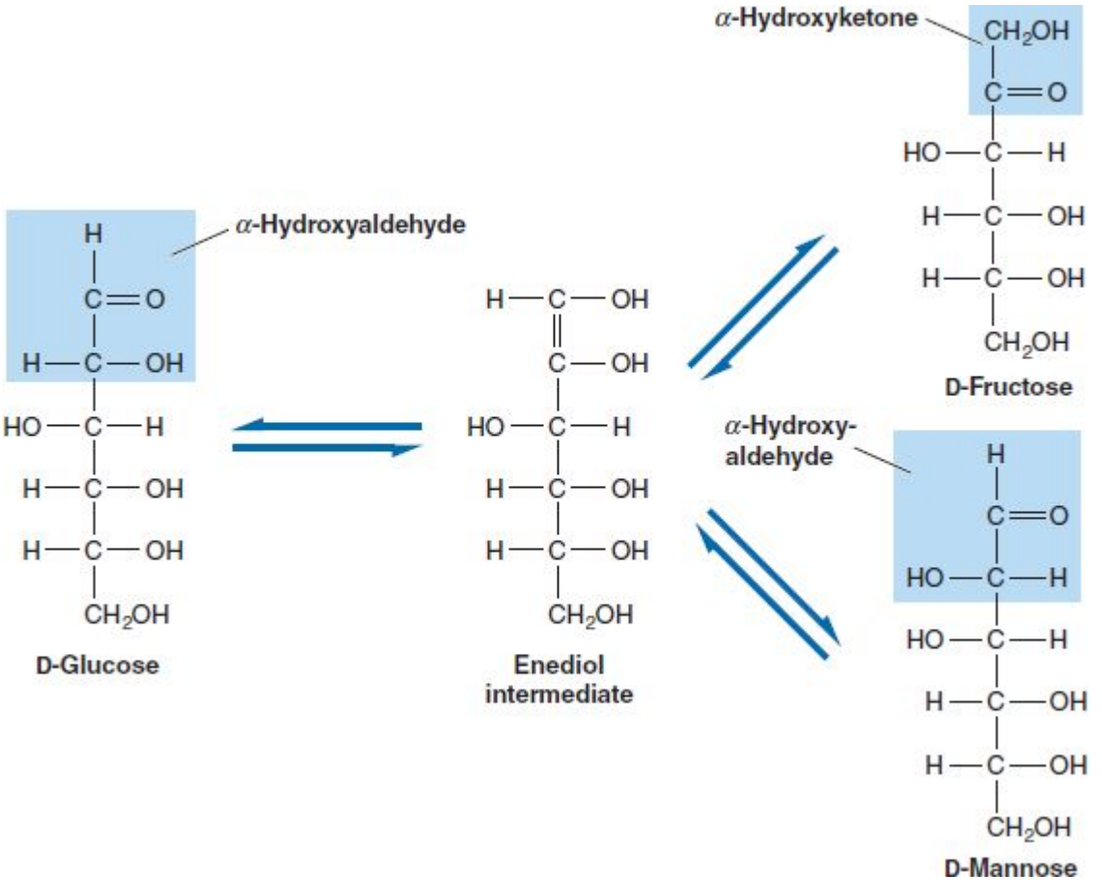


FIGURE 7.16

Isomerization of D-Glucose to Form D-Mannose and D-Fructose

An enediol intermediate is formed in this process.

QUESTION 7.2

Draw the following compounds:

a. α- and β-anomers of D-galactose
b. aldonic acid, uronic acid, and aldaric acid derivatives of galactose
c. galactitol
d. δ-lactone of galactonic acid

GLYCOSIDE FORMATION Hemiacetals and hemiketals react with alcohols to form the corresponding **acetal** or **ketal** (**Figure 7.17**). When the cyclic hemiacetal or hemiketal form of the monosaccharide reacts with an alcohol, the new linkage is called a **glycosidic linkage** and the compound is called a **glycoside**. The name of the glycoside specifies the sugar component. For example, the acetals of glucose and the ketals of fructose are called *glucoside* and *fructoside*, respectively. Additionally, glycosides derived from sugars with five-membered rings are called *furanosides*; those from six-membered rings are called *pyranosides*. A relatively simple example shown in **Figure 7.18** illustrates the reaction of glucose with methanol to form two anomeric types of methyl glucosides. Because glycosides are acetals, they are stable in basic solutions. Carbohydrate molecules that contain only acetal groups do not test positive with Benedict's reagent. (Acetal formation "locks" a ring so that it cannot undergo oxidation or mutarotation.)

Only hemiacetals act as reducing agents.

If an acetal linkage is formed between the hemiacetal hydroxyl group of one monosaccharide and a hydroxyl group of another monosaccharide, the resulting glycoside is called a **disaccharide**. A molecule containing a large number of monosaccharides linked by glycosidic linkages is called a **polysaccharide**.

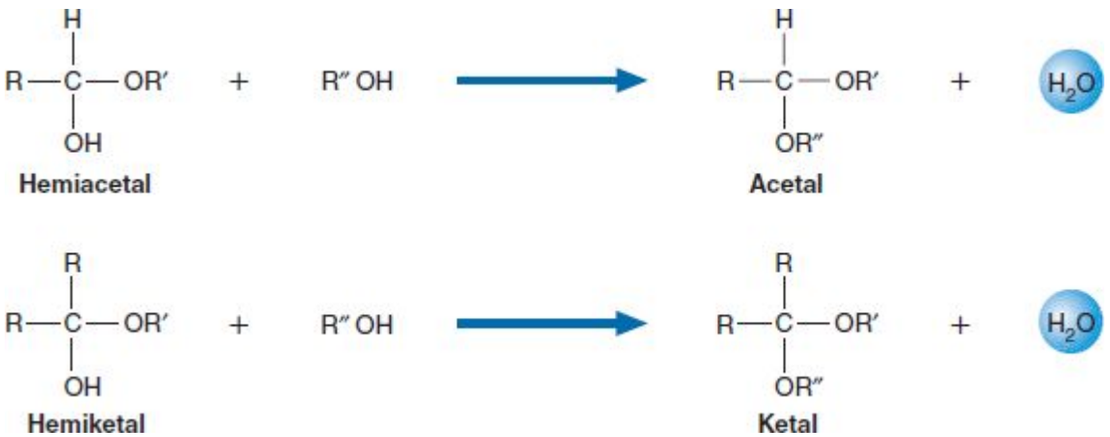


FIGURE 7.17

Formation of Acetals and Ketals

3D animation of hemiacetal

3D animation of acetal

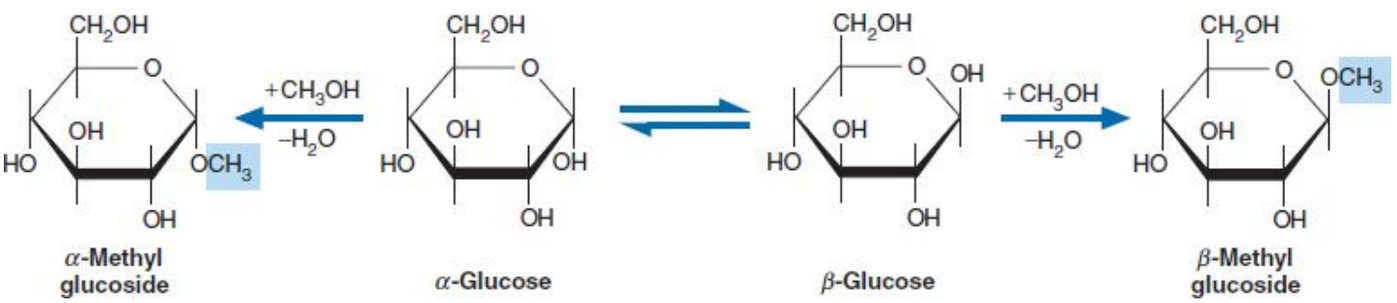


FIGURE 7.18

Methyl Glucoside Formation

Noncarbohydrate components of glycosides are called aglycones. The highlighted methyl groups are aglycones.

QUESTION 7.3

Draw the structure of a D-glucose molecule linked to threonine via a β-glycosidic linkage.

QUESTION 7.4

Glycosides are commonly found in nature. One example is salicin (Figure 7.19), a compound found in willow tree bark that has antipyretic (fever-reducing) and analgesic (pain-relieving) properties. Can you identify the carbohydrate and aglycone (noncarbohydrate) components of salicin?

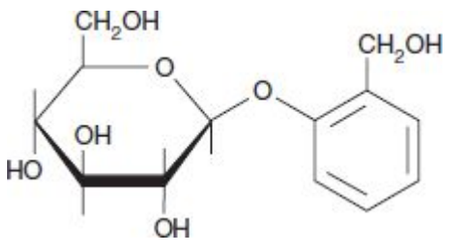


FIGURE 7.19
Salicin Structure

3D animation of salicin

GLYCOSYLATION REACTIONS Glycosylation reactions attach sugars or glycans (sugar polymers) to proteins or lipids. Analogous to glycoside formation between sugar molecules, the glycosylation reactions, catalyzed by the glycosyl transferases, form glycosidic bonds between anomeric carbons in certain sugars and nitrogen or oxygen atoms in other types of molecules. For example, both N- and O-glycosidic bonds are prominent structural features of glycoproteins.

The importance of glycosylation reactions is indicated by the fact that at least 2% of the human polypeptide-coding genome is involved in glycosylation processes. In addition to over 200 glycosylation enzymes, these gene products consist of enzymes involved in the synthesis and transport of donor substrate molecules (e.g., UDP-glucose and GDP [guanosine diphosphate]-mannose), the remodeling of oligosaccharides, and molecules that facilitate the precise localization of all these molecules within the ER and Golgi apparatus. At least 70 inherited human disorders are caused by defects in glycosylation proteins.

GLYCATION Reducing sugars can also react with nucleophilic nitrogen atoms in nonenzymatic reactions. These so-called *glycation* reactions occur rapidly in the presence of heat (e.g., during the cooking or baking of sugar-containing food) or slowly within the body when excess sugar molecules are present. The best-researched example is the reaction of glucose with the side chain amino nitrogen of lysine residues in proteins. The nonenzymatic glycation of protein, called the Maillard reaction (named for the French chemist Louis-Camille Maillard, who discovered it in 1912), begins with the nucleophilic attack of the amino nitrogen on the anomeric carbon of the reducing sugar (**Figure 7.20**). The Schiff base that forms rearranges to yield a stable ketoamine called the *Amadori product*. Both the protein-bound Schiff base and the Amadori product can undergo further reactions (e.g., oxidations, rearrangements, and dehydrations) to produce additional protein-bound products, referred to collectively as advanced glycation end products (AGEs). Reactive carbonyl-containing products such as the dicarbonyl compound glyoxal (CHOCHO) cause rapid protein cross-linkage and adduct formation. (An **adduct** is the product of

an addition reaction, that is, the reaction of two molecules to form a third molecule.)

Glycation alters the structural and functional properties of proteins. For example, the glycation of long-lived proteins such as collagen and elastin disrupts the structure of vascular and connective tissues. In addition, AGEs trigger the production of molecules such as the cytokines that promote inflammatory processes. The accumulation of AGEs has been linked to such age-related conditions as vascular and neurodegenerative diseases and arthritis. In one vascular disease, *atherosclerosis*, cells lining the arterial blood vessels are damaged by AGE formation. AGE-mediated damage initiates a repair process involving macrophages and growth factors that triggers an inflammatory process leading to the formation of artery-clogging deposits called *plaque*. The capacity of affected blood vessels to nourish nearby tissue is eventually compromised. The excessively high blood glucose levels that occur in diabetes mellitus (see the Biochemistry in Perspective essay, Diabetes Mellitus, in Chapter 16) cause an accelerated form of atherosclerosis as well as numerous other AGE-related pathological changes.

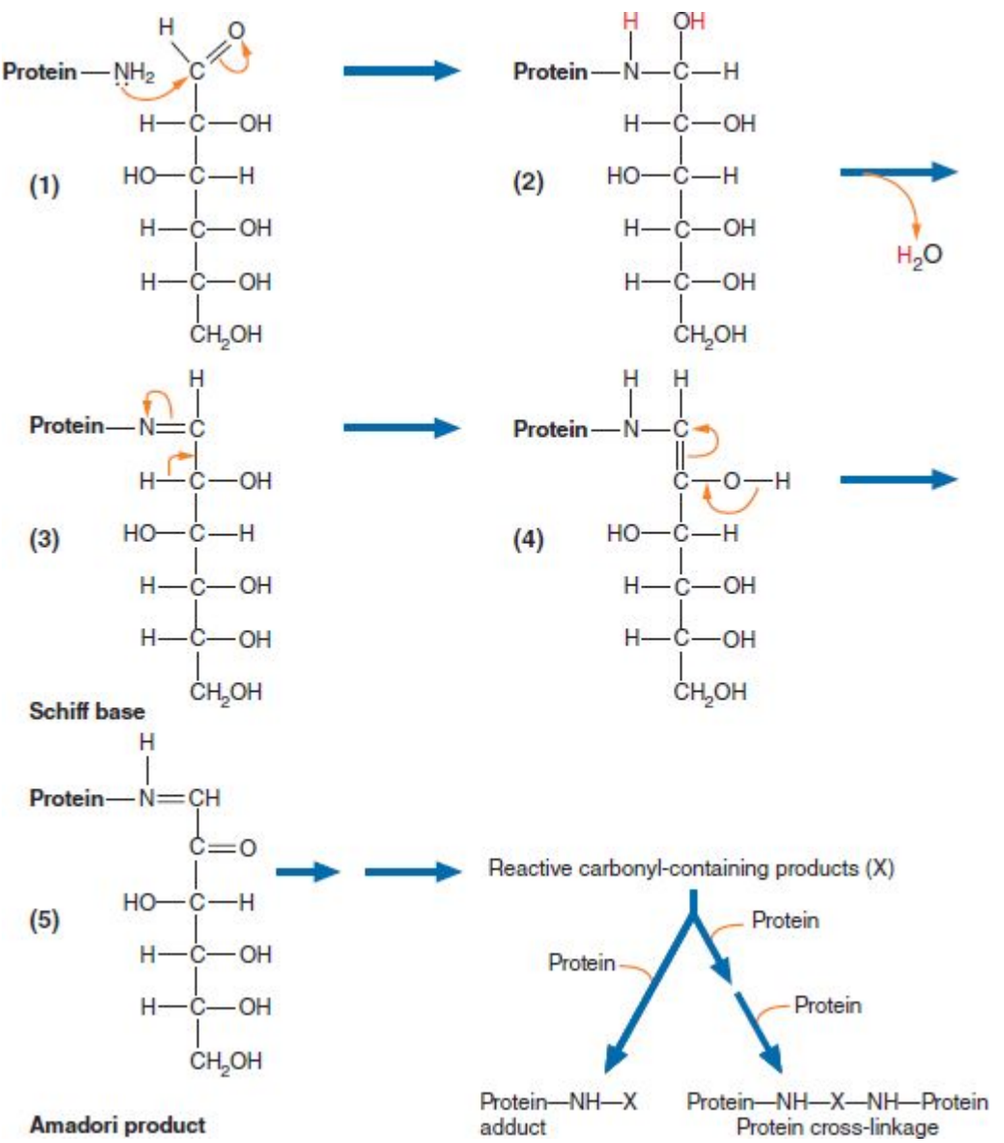


FIGURE 7.20

The Maillard Reaction

Any molecule that contains an amino group can undergo the Maillard reaction, so nucleotides and amines also react with glucose molecules. Since proteins have greater exposure to elevated circulating simple sugars, they

are more compromised by the process. The amino nitrogen of a protein side chain reacts with the carbonyl carbon of the open chain form of an aldose or ketose (1) to produce an intermediate (2) that then undergoes dehydration to form an imine bond (Schiff base) (3). The imine undergoes a tautomerization (p. 291) to yield an intermediate (4), which then undergoes a second tautomerization to yield the Amadori product (5) with a ketone and an amine bond. Amadori products undergo further reactions (oxidation and cleavage) to yield highly reactive carbonyl-containing products that form adducts with the amine groups of other proteins.

Atherosclerosis and Plaque

Important Monosaccharides

Among the most important monosaccharides found in living organisms are glucose, fructose, and galactose. The principal functional roles of these molecules are briefly described.

GLUCOSE D-Glucose, originally called dextrose, is found in large quantities throughout the living world (**Figure 7.21**). It is the primary fuel for living cells. In animals, glucose is the preferred energy source of brain cells and cells that have few or no mitochondria, such as erythrocytes. Cells that have a limited oxygen supply, such as those in the eyeball, also use large amounts of glucose to generate energy. Dietary sources include plant starch and the disaccharides lactose, maltose, and sucrose.

FRUCTOSE D-Fructose, originally called levulose, is often referred to as fruit sugar because of its high content in fruit. It is also found in some vegetables and in honey (**Figure 7.22**). This molecule is an important member of the ketose family of sugars. On a per-gram basis, fructose is twice as sweet as sucrose. It can therefore be used in smaller amounts. For this reason, fructose is often used as a sweetening agent in processed food products. Large amounts of fructose are used in the male reproductive tract. It is synthesized in the seminal vesicles and then incorporated into semen. Sperm use the sugar as an energy source. Fructose is seven times as likely as glucose to undergo glycation reactions. As a result, dietary fructose, especially in large amounts, is a serious risk factor for AGE-related diseases (p. 257).

GALACTOSE Galactose is necessary to synthesize a variety of biomolecules (**Figure 7.23**), including lactose (in lactating mammary glands), glycolipids, certain phospholipids, proteoglycans, and glycoproteins. Synthesis of these substances is not diminished by diets that lack galactose or the disaccharide lactose (the principal dietary source of galactose) because the sugar is readily synthesized from glucose-1-phosphate.

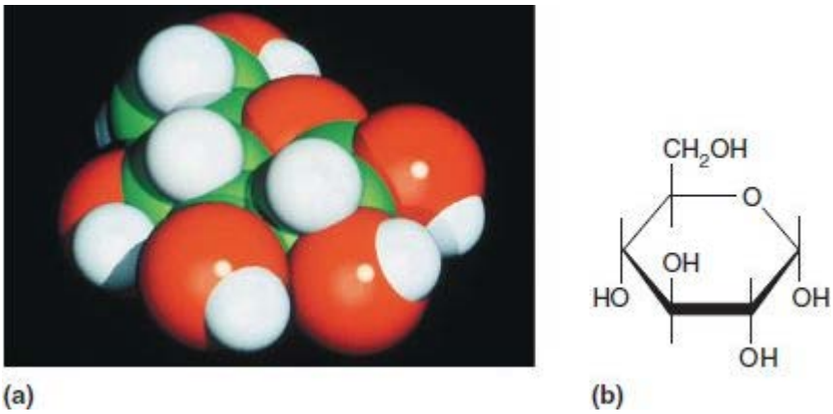


FIGURE 7.21

***α*-D-Glucopyranose**

Compare the information provided by these two representations. (a) The space-filling model, with carbon, oxygen, and hydrogen atoms in green, red, and white, respectively, and (b) the Haworth structure.

3D animation of B-d-Glucopyranose

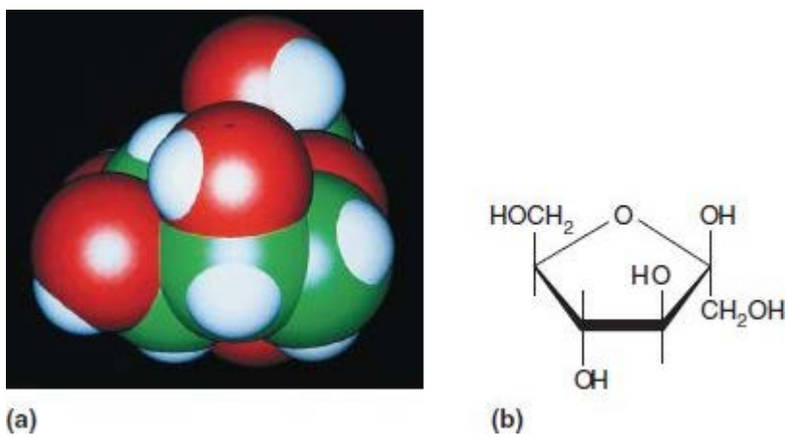


FIGURE 7.22

***β*-D-Fructofuranose**

(a) Space-filling model and (b) Haworth structure.

3D animation of Beta-d-fuctofuranose

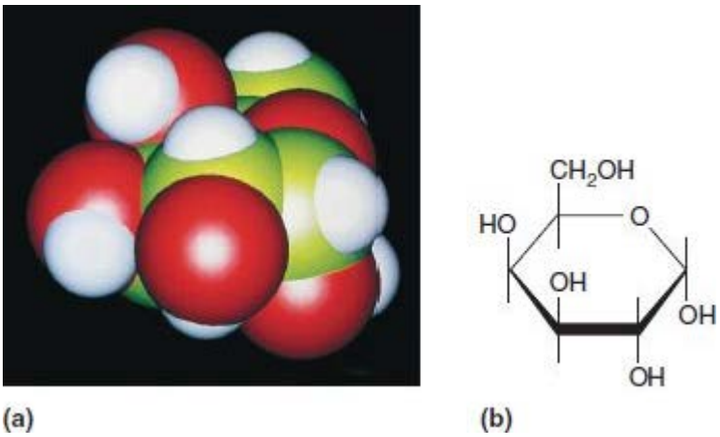


FIGURE 7.23

***α*-D-Galactopyranose**

(a) Space-filling model and (b) Haworth structure.

3D animation of Galactopyranose, beta-d

In *galactosemia*, a genetic disorder, an enzyme required to metabolize galactose is missing. Galactose, galactose-1-phosphate, and galactitol (a sugar alcohol derivative) accumulate and cause liver damage, cataracts, and severe mental retardation. The only effective treatment is early diagnosis and a diet free of galactose.

KEY CONCEPTS

Glucose, fructose, and galactose are among the most important monosaccharides in living organisms.

Monosaccharide Derivatives

Simple sugars may be converted to closely related chemical compounds. Several of these compounds are important metabolic and structural components of living organisms.

Galactosemia

URONIC ACIDS Recall that uronic acids are formed when the terminal CH_2OH group of a monosaccharide is oxidized. Two uronic acids are important in animals: D-glucuronic acid and its epimer, L-iduronic acid (*α*-D-glucuronate and *β*-L-iduronate in **Figure 7.24**). In liver cells, glucuronic acid is combined with molecules such as steroids, certain drugs, and bilirubin (a degradation product of heme in the oxygen-carrying protein hemoglobin) to improve water solubility. This process helps remove waste products from the body. Both D-glucuronic acid and L-iduronic acid are abundant in connective tissue carbohydrate components.

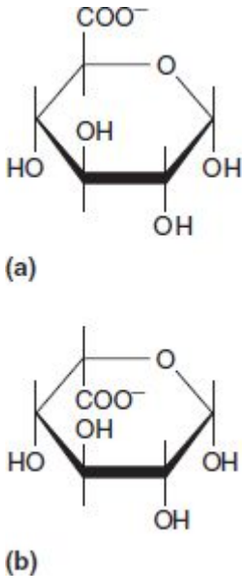


FIGURE 7.24

Uronic Acids

(a) α-D-Glucuronate and (b) β-L-iduronate.

AMINO SUGARS Amino sugars are formed when a hydroxyl group (most commonly on carbon 2) is replaced by an amino group (**Figure 7.25**). They are often constituents of the complex carbohydrate molecules attached to cellular proteins and lipids. The most common amino sugars of animal cells are D-glucosamine and D-galactosamine. Amino sugars are often acetylated. One such molecule is *N*-acetylglucosamine. *N*-Acetylneuraminic acid (the most common form of sialic acid) is a condensation product of D-mannosamine and pyruvic acid, a 2-ketocarboxylic acid. Sialic acids are ketoses containing nine carbon atoms that may be amidated with acetic or glycolic acid (hydroxyacetic acid). They are common components of glycoproteins and glycolipids.

DEOXY SUGARS Monosaccharides in which an —H has replaced an —OH group are known as *deoxy sugars*. Two important deoxy sugars found in cells are L-fucose (formed from D-mannose by reduction reactions) and 2-deoxy-D-ribose (**Figure 7.26**). Fucose is often found among the carbohydrate components of glycoproteins, such as those of the ABO blood group determinants on the surface of red blood cells. 2-Deoxyribose, the pentose sugar component of DNA, was shown earlier (**Figure 1.8**).

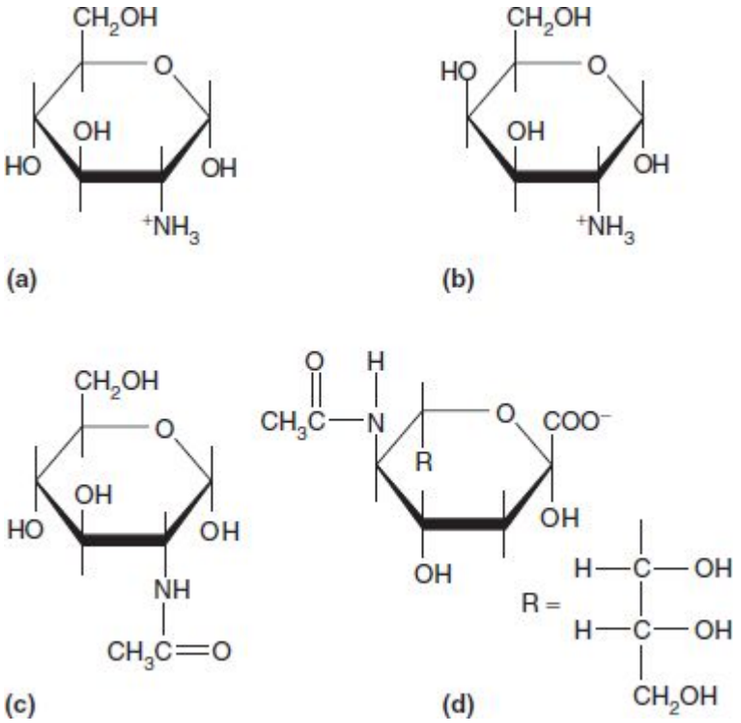


FIGURE 7.25

Amino Sugars

(a) *α*-D-Glucosamine, (b) *α*-D-galactosamine, (c) *N*-acetyl-*α*-D-glucosamine, and (d) *N*-acetylneuraminic acid (sialic acid).

3D animation of N-acetylneuramic acid

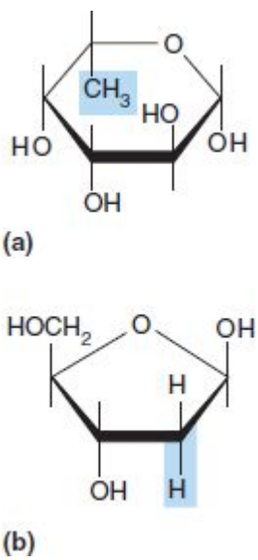


FIGURE 7.26

Deoxy Sugars

(a) *β*-L-Fucose (6-deoxygalactose) and (b) 2-deoxy-*β*-D-ribose. The carbon atoms that have —OH groups replaced by —H are highlighted.

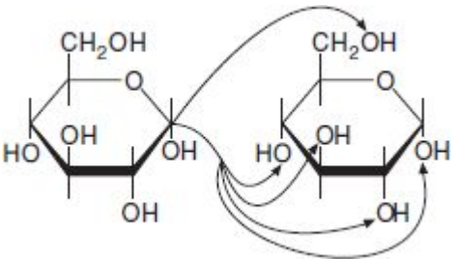


FIGURE 7.27

Glycosidic Bonds

Several types of glycosidic bonds can form between monosaccharides. The sugar α-D-glucopyranose (left) can theoretically form glycosidic linkages with any of the alcoholic functional groups of another monosaccharide, in this case another molecule of α-D-glucopyranose.

7.2 DISACCHARIDES

Disaccharides are molecules composed of two monosaccharides that are linked by a glycosidic bond. If one monosaccharide molecule is linked through its anomeric carbon atom to the hydroxyl group on carbon 4 of another monosaccharide, the glycosidic linkage is designated as 1,4. Because the anomeric hydroxyl group may be in either the α- or the β-configuration, two possible disaccharides may form when two sugar molecules are linked: α(1,4) or β(1,4). Other varieties of glycosidic linkages [i.e., α or β(1,1), (1,2), (1,3), and (1,6) linkages] also occur (**Figure 7.27**).

Lactose (milk sugar) is a disaccharide found in milk. It is composed of one molecule of galactose linked through the hydroxyl group on carbon 1 in a β-glycosidic linkage to the hydroxyl group of carbon 4 of a molecule of glucose (**Figure 7.28**). The anomeric carbon of galactose is in the β-configuration. The linkage between the two monosaccharides is, therefore, designated as β(1,4). Lactose is a reducing sugar because the glucose component contains a hemiacetal group.

Maltose, also known as malt sugar, is an intermediate product of starch hydrolysis and does not appear to exist freely in nature. Maltose is a disaccharide with an α(1,4) glycosidic linkage between two D-glucose molecules. In solution, the free anomeric carbon undergoes mutarotation, which results in an equilibrium mixture of α- and β-maltoses (**Figure 7.29**).

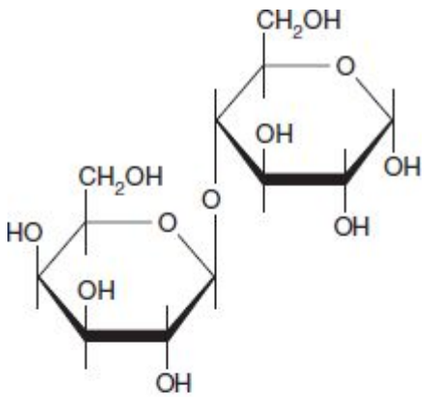
Visit the companion website at www.oup.com/us/mckee to read the Chapter 8 Biochemistry in Perspective essay Fermentation: An Ancient Heritage, which describes the use of maltose in beer brewing.

Cellobiose, a degradation product of cellulose, contains two molecules of glucose linked by a β(1,4) glycosidic bond (**Figure 7.30**). Like maltose, whose structure is identical except for the direction of the glycosidic bond, cellobiose does not occur freely in nature.

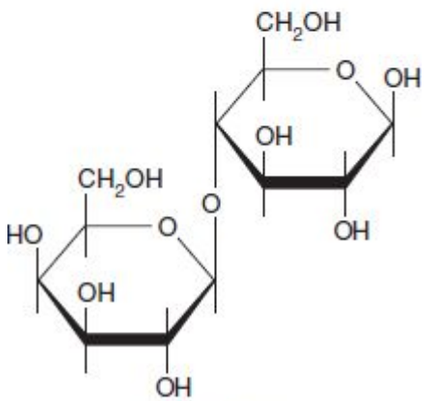
Sucrose (common table sugar: cane sugar or beet sugar) is produced in the leaves and stems of plants. It is a transportable energy source throughout the entire plant. Containing α-glucose and β-fructose residues, sucrose differs from the previously described disaccharides in that the monosaccharides are linked through a glycosidic bond between both anomeric carbons (**Figure 7.31**). Because neither monosaccharide ring can revert to the open-chain form, sucrose is a nonreducing sugar.

Enzymes synthesized by cells lining the small intestine mediate digestion of disaccharides and other carbohydrates. Deficiency of any of these enzymes causes unpleasant symptoms when the

indigestible disaccharide sugar is ingested. Because carbohydrates are absorbed principally as monosaccharides, any undigested disaccharide molecules pass into the large intestine, where osmotic pressure draws water from the surrounding tissues (diarrhea). Bacteria in the colon digest the disaccharides (fermentation), thus producing gas (bloating and cramps). The most commonly known deficiency is *lactose intolerance*, which may occur in most human adults except those with ancestors from northern Europe and/or certain African groups. Caused by the greatly reduced synthesis of the enzyme lactase following childhood, lactose intolerance is treated by eliminating the sugar from the diet or (in some cases) by treating food with the enzyme lactase.



α-Lactose



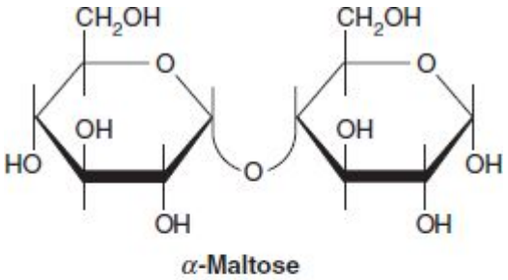
β-Lactose

FIGURE 7.28

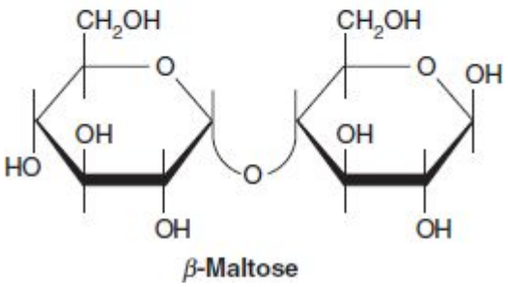
***α*- and *β*-Lactose**

3D animation of a-lactose

3D animation of beta-lactose



α-Maltose



β-Maltose

FIGURE 7.29

***α*- and *β*-Maltose**

3D animation of maltose

Lactose Intolerance

Both sucrose and fructose contribute to dental caries (tooth decay). (Other factors include genetic predisposition to bacteria-mediated tooth decay, e.g., genes that result in decreased antibodies in saliva or inadequate saliva production). Tooth decay is a chronic disease process that is associated with high sugar diets. Bacteria such as *Streptococcus mutans* within the biofilm (p. 43) that coats teeth metabolize sugar molecules to yield lactic, formic, and acetic acids. As the pH in the biofilm drops below 5.5, demineralization of the tooth begins and then accelerates as other acid-tolerant bacteria become more prevalent. Prevention of tooth decay involves good oral hygiene (brushing several times a day) and healthy diets with little or no processed sugar.

KEY CONCEPTS

- Disaccharides are glycosides composed of two monosaccharide units.
- Maltose, lactose, cellobiose, and sucrose are disaccharides.

QUESTION 7.5

Which of the following sugars or sugar derivatives are reducing sugars?

a. maltose

b. fructose
c. α-methyl-D-glucoside
d. sucrose

Which of these compounds are capable of mutarotation?

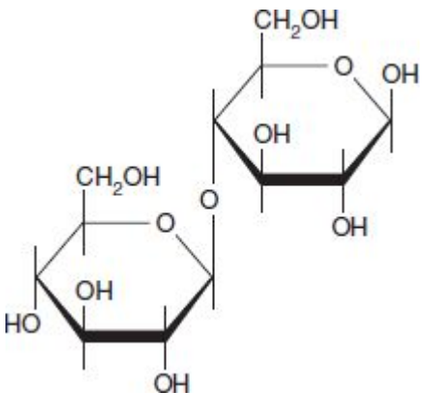


FIGURE 7.30
***β*-Cellobiose**

7.3 POLYSACCHARIDES

Polysaccharides, also referred to as **glycans**, are composed of large numbers of monosaccharide units connected by glycosidic linkages. Smaller glycans, called **oligosaccharides**, are polymers containing up to about 10 or 15 monomers, most often attached to polypeptides in glycoproteins (p. 268) and some glycolipids (p. 413). Larger glycans may contain hundreds to thousands of sugar units. These molecules may have a linear structure, or they may have branched shapes. Polysaccharides may be divided into two classes: **homoglycans**, which are composed of one type of monosaccharide, and **heteroglycans**, which contain two or more types of monosaccharides.

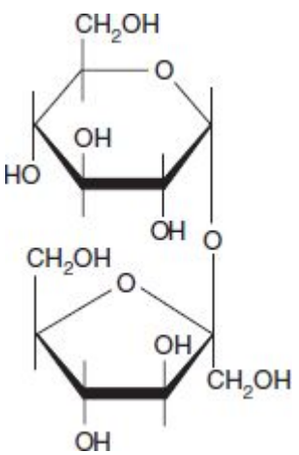


FIGURE 7.31
Sucrose

The glucose and fructose residues are linked by an α, β(1,2) glycosidic bond.

3D animation of sucrose

Homoglycans

The **homoglycans** found in abundance in nature are starch, glycogen, cellulose, and chitin. Starch, glycogen, and cellulose all yield D-glucose when they are hydrolyzed. Starch and glycogen are energy storage molecules in plants and animals, respectively. Cellulose is the primary structural component of plant cells. **Chitin**, the principal structural component of the exoskeletons of arthropods such as insects and crustaceans and the cell walls of many fungi, yields the glucose derivative *N*-acetylglucosamine when it is hydrolyzed.

Polysaccharides such as starch and glycogen, unlike proteins and nucleic acids, have no fixed molecular weights. The size of such molecules reflects the metabolic state of the cell producing them. For example, when blood sugar levels are high (e.g., after a meal), the liver synthesizes glycogen. Glycogen molecules in a well-fed animal may have molecular weights as high as 2×10^7 Da. When blood sugar levels fall, the liver enzymes begin breaking down the glycogen molecules, releasing glucose into the bloodstream. If the animal continues to fast, the process continues until glycogen reserves are almost used up.

STARCH Starch, the energy reservoir of plant cells, is a significant source of carbohydrate in the human diet. Much of the nutritional value of the world's major foodstuffs (e.g., potatoes, rice, corn, and wheat) comes from starch. Two polysaccharides occur together in starch: amylose and amylopectin.

Amylose is composed of long, unbranched chains of D-glucose residues that are linked with α(1,4) glycosidic bonds (**Figure 7.32**). A number of polysaccharides, including both types of starch, have one *reducing end* in which the ring can open to form a free aldehyde group with reducing properties. The internal anomeric carbons in these molecules are involved in acetal linkages and are not free to act as reducing agents.

Amylose molecules, which typically contain several thousand glucose residues, vary in molecular weight from 150,000 to 600,000 Da. Because the linear amylose molecule forms long, tight helices, its compact shape is ideal for its storage function. The common iodine test for starch works because molecular iodine inserts itself into these helices. (The intense blue color of a positive test comes from electronic interactions between iodine molecules and the helically arranged glucose residues of the amylose.)

The other form of starch, **amylopectin**, is a branched polymer containing both α(1,4) and α(1,6) glycosidic linkages. The α(1,6) branch points may occur every 20 to 25 glucose residues and prevent helix formation (**Figure 7.33a**). The number of glucose units in amylopectin may vary from a few thousand to a million.

Starch digestion begins in the mouth, where the salivary enzyme α-amylase initiates hydrolysis of the glycosidic linkages. Digestion continues in the small intestine, where pancreatic α-amylase randomly hydrolyzes all the α(1,4) glycosidic bonds except those next to the branch points. The products of α-amylase are maltose, the trisaccharide maltotriose, and the α-limit dextrins [oligosaccharides that typically contain eight glucose units with one or more α(1,6) branch points]. Several enzymes secreted by cells that line the small intestine convert these intermediate products into glucose. Glucose molecules are then absorbed into enterocytes, the cells that line the small intestine. After passage into the bloodstream, they are transported to the liver and then to the rest

of the body.

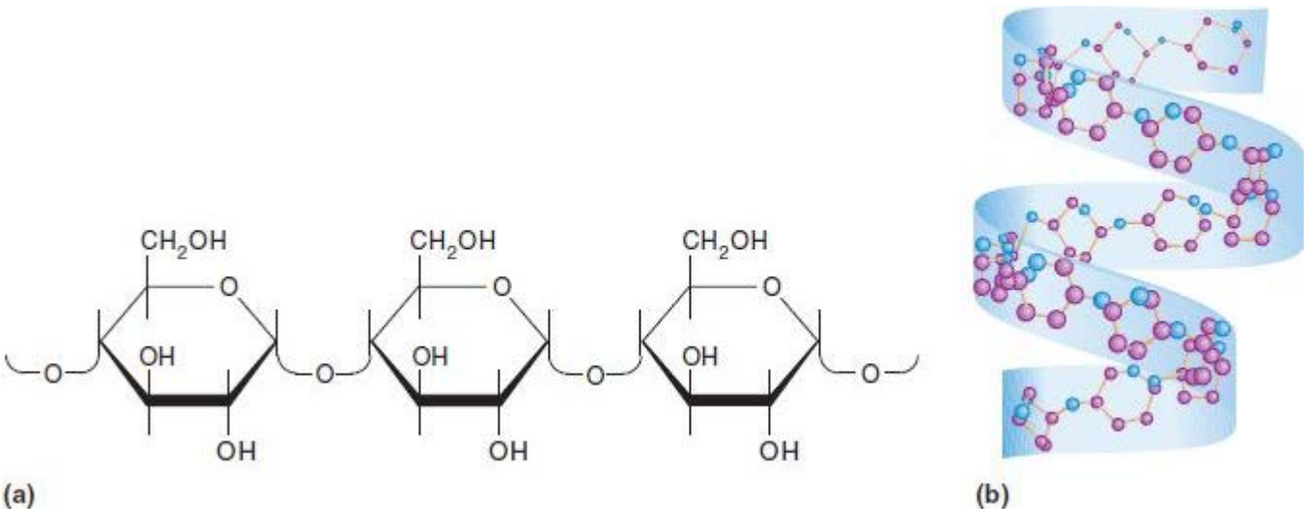


FIGURE 7.32

Amylose

(a) The D-glucose residues of amylose are linked through α(1,4) glycosidic bonds. (b) The amylose polymer forms a left-handed helix.

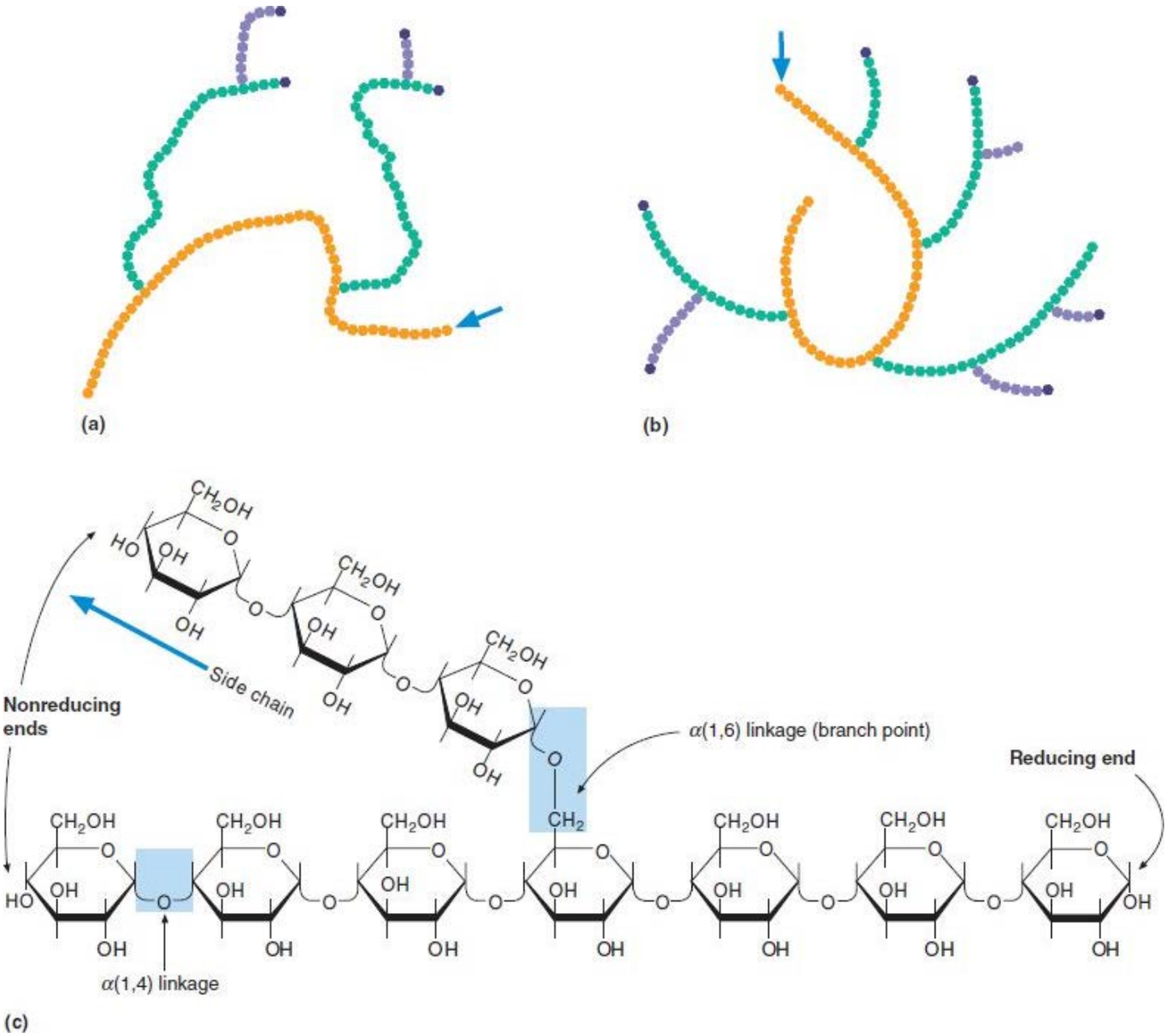


FIGURE 7.33

(a) Amylopectin and (b) Glycogen

Each hexagon represents a glucose molecule. Notice that each molecule has only one reducing end (arrow) and numerous nonreducing ends. (c) Detail from (a) or (b).

3D animation of amylopectin

3D animation of glycogen

GLYCOGEN **Glycogen** is the carbohydrate storage molecule in vertebrates. It is found in greatest abundance in liver and muscle cells. (Glycogen may make up as much as 8 to 10% of the wet weight of liver cells and 2 to 3% of that of muscle cells.) Glycogen (**Figure 7.33b**) is similar in

structure to amylopectin except that it has more branch points, possibly as close as every fourth glucose residue in the core of the molecule. In the outer regions of glycogen molecules, branch points are not so close together (approximately every 8–12 residues). The glycogen molecule is more compact than other polysaccharides, so it takes up little space, which is an important consideration in mobile animal bodies. Because hydrolysis occurs from the many nonreducing ends of the glycogen molecule, energy mobilization can be rapid.

QUESTION 7.6

It has been estimated that two high-energy phosphate bonds must be expended to incorporate one glucose molecule into glycogen. Why is glucose stored in muscle and liver in the form of glycogen, and not as individual glucose molecules? In other words, why is it advantageous for a cell to expend metabolic energy to polymerize glucose molecules? [*Hint*: Besides the reasons given in Section 7.3, refer to Chapter 3 for another problem that glucose polymerization solves.]

CELLULOSE **Cellulose** is a polymer composed of D-glucopyranose residues linked by β(1,4) glycosidic bonds (**Figure 7.34**). It is the most important structural polysaccharide of plants. Because cellulose comprises about one-third of plant biomass, it is the most abundant organic substance on Earth. Approximately 100 trillion kilograms of cellulose are produced each year.

Unbranched cellulose molecules, each of which may contain as many as 12,000 glucose units, are held together by hydrogen bonding to form tough and inflexible sheet-like strips called *microfibrils* (**Figure 7.35**). With a tensile strength comparable to that of steel wire, cellulose microfibrils are components of both plant primary and secondary cell walls, where they provide a structural framework that both protects and supports cells.

The ability to digest cellulose is found only in microorganisms that possess the enzyme cellulase. Certain animal species (e.g., termites and cows) use such organisms in their digestive tracts to digest cellulose. The breakdown of the cellulose makes glucose available to both the microorganisms and their host. Although many animals cannot digest cellulose-containing plant materials, these substances play a vital role in nutrition. Cellulose is one of several plant products that make up the dietary fiber that is now believed to be important for good health.

Because of its structural properties, cellulose has enormous economic importance. Products such as wood, paper, and textiles (e.g., cotton, linen, and ramie) owe many of their unique characteristics to their cellulose content.

Heteroglycans

Heteroglycans are high-molecular-weight carbohydrate polymers that contain more than one kind of monosaccharide. The major classes of heteroglycans found most often in mammals are N- and O-linked heteropolysaccharides (N- and O-glycans) attached to membrane and secretory proteins, the glycosaminoglycans of the extracellular matrix, and the glycan components of glycolipids and GPI (glycosylphosphatidylinositol) membrane anchors. The structure and properties of the N- and O-glycans and the glycosaminoglycans are described next. Discussion of glycolipids and GPI anchors, a means for attaching peripheral proteins to membrane, is deferred to Chapter 11.

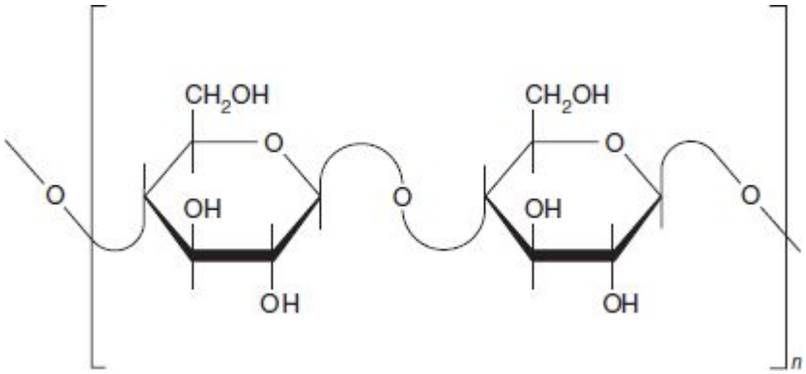


FIGURE 7.34

The Disaccharide Repeating Unit of Cellulose

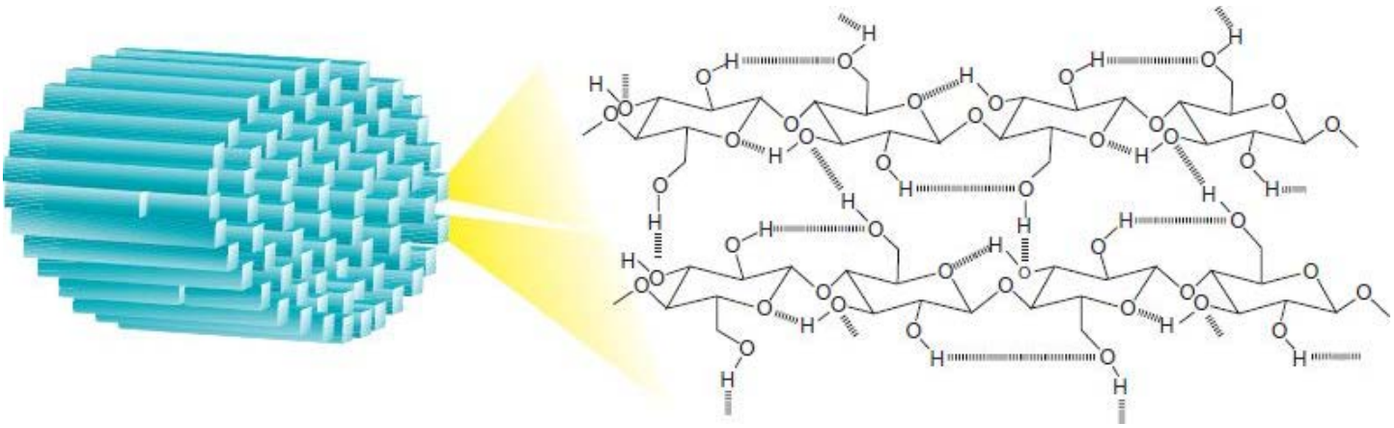


FIGURE 7.35

Cellulose Microfibrils

Intermolecular hydrogen bonds between adjacent cellulose molecules are largely responsible for the great strength and water insolubility of cellulose.

N- AND O-GLYCANS Many proteins have both N- and O-linked heterooligosaccharides that can comprise a significant proportion of the molecule's molecular weight. The N-linked heterooligosaccharides (**N-heteroglycans**) are linked via a β-glycosidic bond between the core *N*-acetylglucosamine anomeric carbon and a side chain amide nitrogen of an asparagine residue. In addition to *N*-acetylglucosamine, the most commonly observed sugars among the N-glycans include mannose, galactose, *N*-acetylneuraminic acid, and glucose. There are three major types of asparagine-linked glycans: high-mannose, complex, and hybrid (**Figure 7.36**).

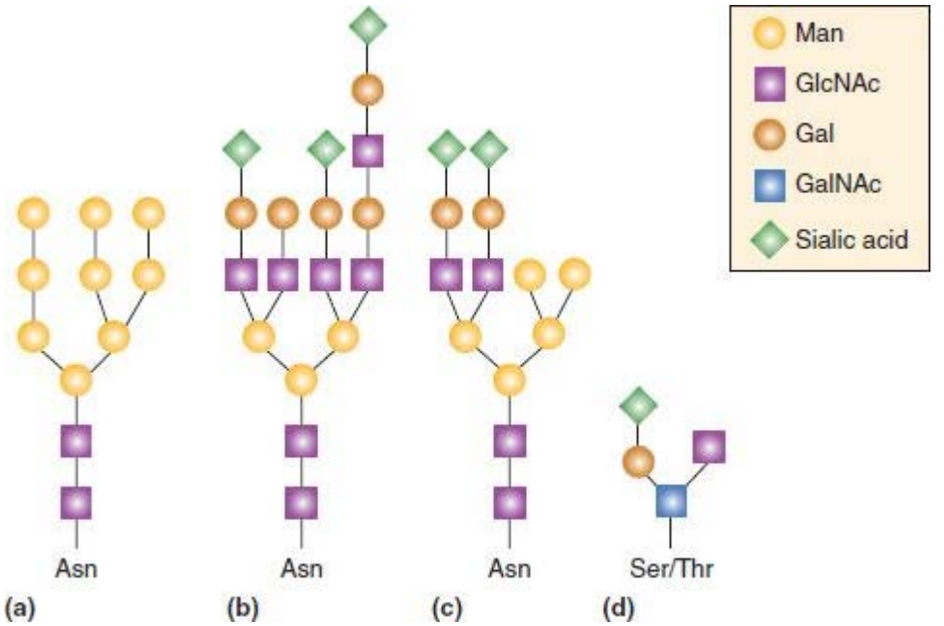


FIGURE 7.36

Heteroglycans Linked to Glycoproteins

Two major classes of branched heteroglycans linked to polypeptides in glycoproteins are N-linked and O-linked. Three classes of N-linked heteroglycan precursors, formed by a glycosidic linkage between the amide nitrogen of asparagine and the heteroglycan, are (a) high-mannose, (b) complex, and (c) hybrid. Notice that the core structure, consisting of two N-acetylglucosamine residues and three mannose residues, is the same in all three N-glycan classes. The O-linked heteroglycan precursor (d) is formed by a glycosidic linkage between the hydroxyl group of serine or threonine and the heteroglycan. After their transport to the Golgi apparatus, these glycans are processed further by selective removal of some sugars and the addition of others, resulting in the final form of the glycoprotein.

The O-linked heterooligosaccharides **(O-heteroglycans)** have a disaccharide core of galactosyl-β-(1,3)-*N*-acetylgalactosamine linked to the protein via an α-glycosidic bond to the hydroxyl oxygen of serine or threonine residues. In the collagens, the core β-linked disaccharide may be Gal-Gal or Glc-Gal and is linked to the side chain hydroxyl oxygen of 5-hydroxylysine. Other sugars found in O-heteroglycans are *N*-acetylneuraminic acid and other sialic acids.

GLYCOSAMINOGLYCANS **Glycosaminoglycans** (GAGs) are linear polymers with disaccharide repeating units. Many of the sugar residues are amino derivatives. There are five GAG classes: hyaluronic acid, chondroitin sulfate, dermatan sulfate, heparin and heparan sulfate, and keratan sulfate. The repeating units contain a hexuronic acid (a uronic acid containing six carbon atoms), except for keratan sulfate, which contains galactose. Usually, an *N*-acetylhexosamine sulfate is also present, except in hyaluronic acid, which contains *N*-acetylglucosamine. Many disaccharide units contain both carboxyl and sulfate groups. GAGs are classified according to their sugar residues, the linkages between these residues, and the presence and location of sulfate groups.

KEY CONCEPTS

Polysaccharide molecules, composed of large numbers of monosaccharide units, are used in energy storage and as structural materials.

GAGs have many negative charges at physiological pH. The charge repulsion keeps GAGs separated from each other. Additionally, the relatively inflexible polysaccharide chains are

strongly hydrophilic. GAGs occupy a huge volume relative to their mass because they attract large volumes of water. For example, hydrated hyaluronic acid may occupy a volume 1000 times greater than its dry state.

7.4 GLYCOCONJUGATES

The molecules that result from the covalent linkages of carbohydrate molecules to both proteins and lipids are collectively known as the **glycoconjugates**. There are two classes of carbohydrate-protein conjugate: proteoglycans and glycoproteins. Although both molecular types contain carbohydrate and protein, their structures and functions appear, in general, to be substantially different. The *glycolipids* (Chapter 11), which are sugar-containing lipid molecules, are found predominantly on the outer surface of plasma membranes.

The *extracellular matrix* (ECM; p. 45) is a dynamic and gelatinous network of macromolecules, secreted by cells such as fibroblasts. A large proportion of these molecules are glycoconjugates: the GAG-containing proteoglycans and several glycoproteins (most notably laminin and fibronectin). The glycoconjugates in combination with fibrous proteins such as collagen (p. 174) and elastin are responsible for the ECM's major influence on processes such as cellular development, differentiation, and mobility, and wound healing.

Proteoglycans

Proteoglycans are distinguished from the more commonly known glycoproteins by their extremely high carbohydrate content, which may constitute as much as 95% of the dry weight of such molecules. Most of these molecules occur on cell surfaces or are secreted into the extracellular matrix. All proteoglycans contain GAG chains that are linked to protein molecules (known as *core proteins*) by N- and O-glycosidic linkages. Proteoglycans are produced in the Golgi apparatus where GAG chains are synthesized and then covalently linked to a core protein previously synthesized in the RER. The diversity of proteoglycans results from both the number of different core proteins and the large variety of classes and lengths of the carbohydrate chains. Examples include the syndecans, glypicans, and aggrecan, which occur on cell surfaces.

Both syndecans and glypicans (**Figure 7.37**) are referred to as heparan sulfate proteoglycans (HSPGs) because of their heparan sulfate GAG components. Syndecan core proteins possess a transmembrane domain. They also have domains linked to both heparan sulfate and chondroitin sulfate GAGs. The glypicans are proteoglycans that contain heparan sulfate and are linked to the cell membrane by GPI anchors (p. 411). Aggrecan, a proteoglycan found in abundance in cartilage and intervertebral discs, consists of a core protein to which are attached more than 100 GAGs that are separated into chondroitin sulfate and keratan sulfate domains. Up to 100 aggrecan monomers are in turn attached noncovalently to hyaluronic acid by link proteins to form a proteoglycan aggregate (**Figure 7.38**).

In addition to their roles in organizing extracellular matrices, proteoglycans participate in all cellular processes that involve events at cell surfaces. For example, the membrane-bound syndecans and glypicans that bind to specific signal molecules (e.g., growth factors) are components in several signal transduction pathways that regulate the cell cycle. The vast numbers of polyionic GAG chains in the aggrecans allow them to trap large volumes of water. Consequently, these molecules occupy thousands of times as much space as a densely packed molecule of the same mass. The strength, flexibility, and resilience of cartilage are made possible by the combination of the compressive stiffness contributed by repulsion between the negatively

charged GAGs and the tensile strength of collagen fibers.

Proteoglycans also have intracellular roles. For example, serglycin is found in the secretory granules of several cell types. In mast cells (tissue cells that mediate inflammatory responses such as allergic reactions), the serglycin core protein is linked to highly sulfated heparin GAG chains. These negatively charged molecules form an anionic gel, which binds to and protects positively charged molecules (histamine and numerous proteolytic enzymes) stored within the granules.

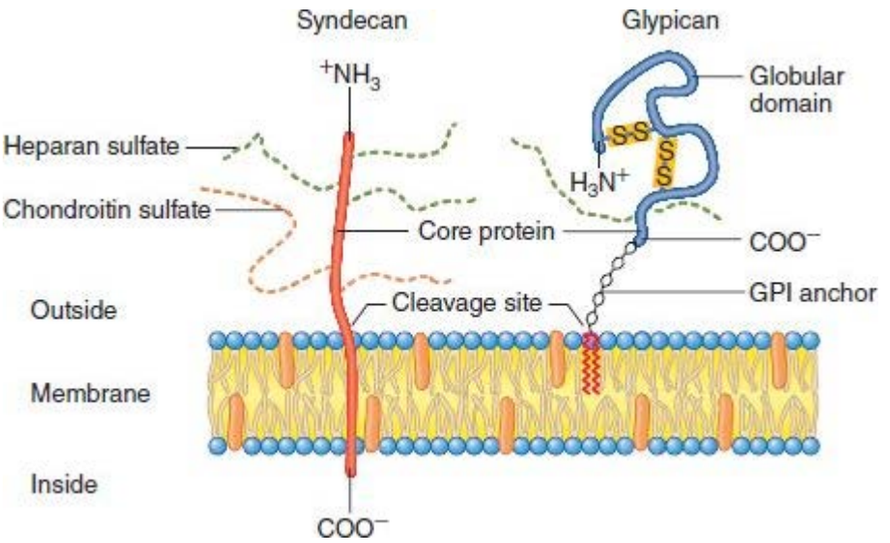


FIGURE 7.37

Structure of the Syndecans and Glypicans

A syndecan is a single-pass transmembrane core polypeptide that is linked to heparan sulfate and chondroitin sulfate GAG chains. Each type of glypican consists of a core polypeptide bound to heparan sulfate chains that in turn is bound to the cell membrane by a GPI linkage. Several disulfide bridges maintain the three-dimensional structure of the polypeptide's globular domain. Certain proteases and phospholipases can cleave the syndecans and glypicans, respectively, from the cell membrane, thus releasing them into the extracellular matrix. This activity, called shedding, often occurs during inflammatory processes such as bacterial infections and cancer.

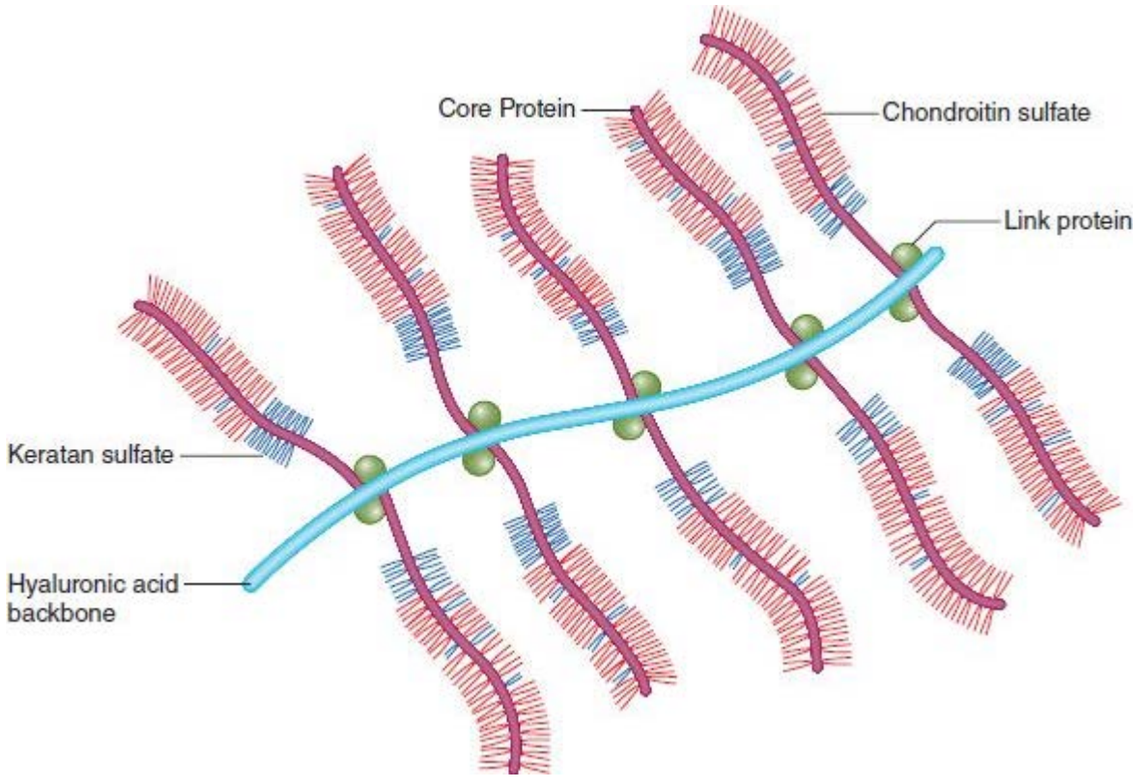


FIGURE 7.38

Proteoglycan Aggregate Structure

Aggrecan, the core protein of aggrecan proteoglycans, is attached to GAG chains, in this case chondroitin sulfate and keratan sulfate, by an O-glycosidic bond formed between a polypeptide serine residue and a terminal sugar residue in the GAG chain. Aggrecan aggregates are formed by link protein, a glycoprotein that stabilizes the noncovalent interaction of aggrecan proteoglycans with hyaluronic acid.

Hurler's Syndrome

A number of genetic diseases associated with proteoglycan metabolism, known as *mucopolysaccharidoses*, have been identified. For example, in *Hurler's syndrome*, deficiency of a specific lysosomal enzyme causes excessive accumulation of dermatan sulfate. Symptoms include mental retardation, skeletal deformity, and death in early childhood. Like Tay–Sachs disease (p. 54), Hurler's syndrome is an autosomal recessive disorder (i.e., one copy of the defective gene is inherited from each parent).

Glycoproteins

Glycoproteins are defined as proteins that are covalently linked to carbohydrate through N- or O-linkages. The carbohydrate composition of glycoprotein varies from 1% to more than 85% of total weight. The carbohydrates found include monosaccharides and disaccharides such as those attached to the structural protein collagen and the branched oligosaccharides on plasma glycoproteins. Although the glycoproteins are sometimes considered to include the proteoglycans, for structural reasons they are examined separately. There is a relative absence in glycoproteins of uronic acids, sulfate groups, and the disaccharide repeating units that are typical of proteoglycans.

The carbohydrate groups of glycoproteins are linked to the polypeptide by either N-glycosidic or O-glycosidic linkages (see **Figure 7.36**). The former glycoprotein class is sometimes referred to as *asparagine-linked*; the latter is often called *mucin-type*, in reference to the mucins, a glycoprotein constituent of mucus (a slippery secretion that protects the surface of some cell types). The asparagine-linked oligosaccharides are constructed on a membrane-bound lipid molecule and are then covalently linked to asparagine residues during ongoing protein synthesis (Chapter 19). Several additional reactions, in the lumen of the endoplasmic reticulum and the Golgi complex, form the final N-linked oligosaccharide structures. Examples of proteins with asparagine-linked oligosaccharides include the iron transport protein transferrin and ovalbumin, a nutritional storage protein found in chicken eggs. Mucin-type carbohydrate (O-glycan) units vary considerably in size and structure from disaccharides such as Gal-1,3-GalNAc, found in the antifreeze glycoprotein of Antarctic fish, to the complex oligosaccharides of blood groups such as those of the ABO system.

GLYCOPROTEIN FUNCTIONS Glycoproteins are a diverse group of molecules that are ubiquitous constituents of most living organisms (**Table 7.1**). They occur in cells, in both soluble and membrane-bound form, and in extracellular fluids. Examples in vertebrates include the metal-transport proteins transferrin and ceruloplasmin, the blood-clotting factors, and many of the components of complement (proteins involved in cell destruction during immune reactions). A number of hormones are glycoproteins. Consider, for example, follicle-stimulating hormone produced by the anterior pituitary gland, which stimulates the development of both eggs and sperm. Many enzymes are glycoproteins. Ribonuclease (RNase), the enzyme that degrades ribonucleic acid, is a well-researched example. Other glycoproteins are integral membrane proteins (Chapter 11). Of these, Na^+-K^+-ATPase (an ion pump found in the plasma membrane of animal cells) and the major histocompatibility antigens (cell-surface markers used to cross-match organ donors and recipients) are especially interesting examples.

Recent research has focused on how carbohydrate groups assist in protein folding and function in recognition processes in multicellular organisms. The presence of carbohydrate on protein molecules also protects them from denaturation. For example, bovine RNase A is more susceptible to heat denaturation than its glycosylated counterpart RNase B. Several other studies have shown that sugar-rich glycoproteins are relatively resistant to proteolysis (enzyme-catalyzed hydrolysis of polypeptides). Carbohydrate on the molecule's surface shields the polypeptide chain from proteolytic enzymes.

The carbohydrates in glycoproteins also affect biological function. For example, a large content of sialic acid residues is responsible for the high viscosity of salivary mucins (the lubricating glycoproteins of saliva). The disaccharide residues of the antifreeze glycoproteins of Antarctic fish form hydrogen bonds with water molecules, thereby retarding the growth of ice crystals.

Glycoproteins, as components of the glycocalyx (p. 45), are now known to be important in complex recognition phenomena. A prime example is the insulin receptor, whose binding to insulin facilitates the transport of glucose into numerous cell types. It does so, in part, by recruiting glucose transporters to the plasma membrane. A variety of cell-surface glycoproteins are also involved in cellular adhesion, a critical event in the cell–cell interactions of growth and differentiation (**Figure 7.39**). The best characterized of these molecules are called cell adhesion molecules (CAMs). Examples include the *selectins* (transient cell–cell interactions), the *integrins* (cell attachment to the components of the extracellular matrix), and the *cadherins* (Ca^{2+}-dependent binding of cells to each other within a tissue). The roles of glycoconjugates in living processes are explored further in the next section.

KEY CONCEPTS

- Glycoconjugates are biomolecules in which carbohydrate is covalently linked to either proteins or lipids.
- Proteoglycans are composed of relatively large amounts of carbohydrate (GAG units) covalently linked to small polypeptide components.
- Glycoproteins are proteins covalently linked to carbohydrate through N- or O-linkages.

TABLE 7.1 Glycoproteins

Type	Example	Source	Molecular Mass (Da)
Enzyme	Ribonuclease B	Bovine	14,700
Immunoglobulin	Immunoglobulin A Immunoglobulin M	Human Human	160,000 950,000
Hormone	Chorionic gonadotropin Follicle-stimulating hormone	Human placenta Human	38,000 34,000
Membrane protein	Glycophorin	Human red blood cells	31,000
Lectin (carbohydrate-building proteins)	Potato lectin Soybean agglutinin Ricinus lectin	Potato Soybean Castor bean	50,000 120,000 120,000
Viral envelope protein	gp120	HIV	120,000

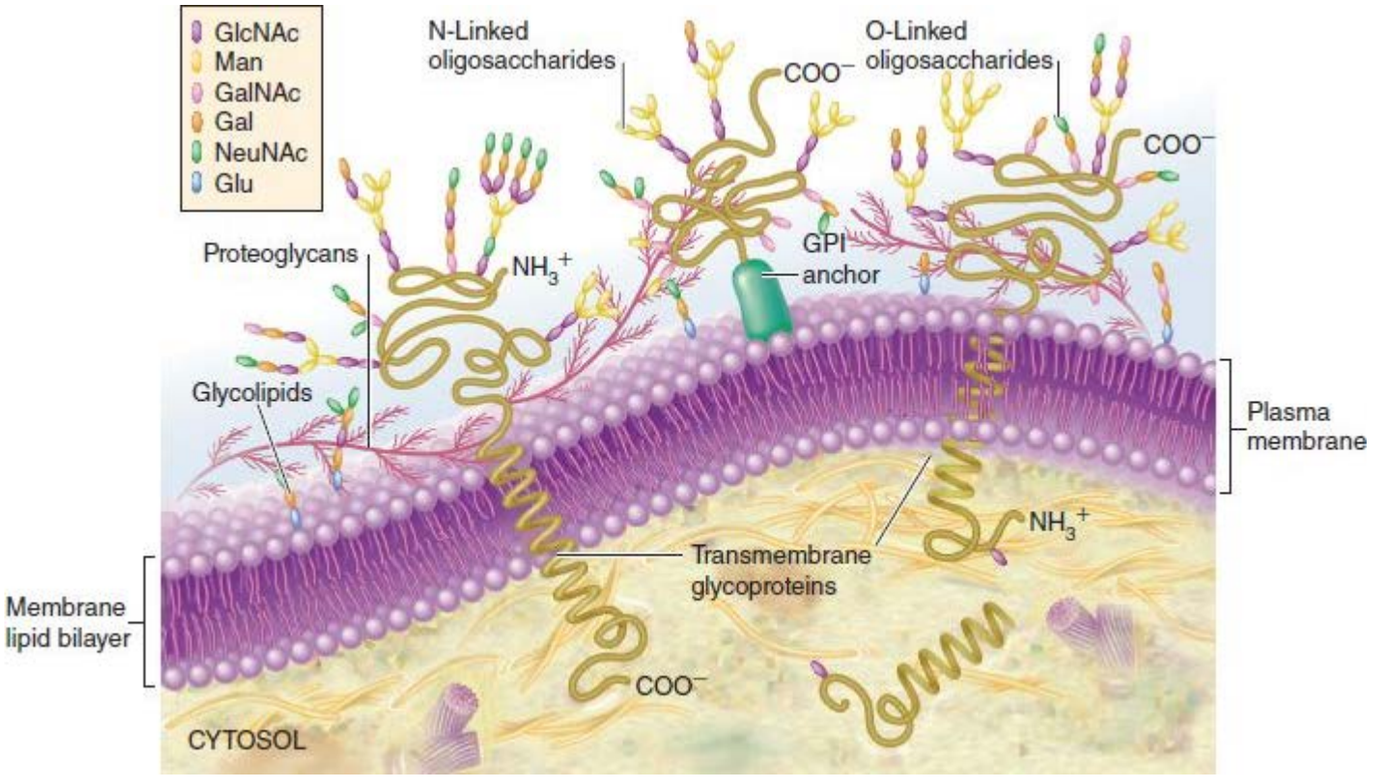


FIGURE 7.39

The Glycocalyx

The glycocalyx is made up of the carbohydrate groups attached to the glycoprotein, proteoglycan, and glycolipid components on the external surface of eukaryotic cells. A glycosylphosphatidylinositol (GPI) anchor (p. 411) is a specialized oligosaccharide-containing structure that attaches several types of protein to the plasma membrane of some cells.

7.5 THE SUGAR CODE

Living organisms require extraordinarily large coding capacities. Each information transfer event, whether it is the binding of a substrate within an enzyme's active site, the transduction of a hormonal signal, or the engulfment of a bacterial cell by a macrophage, is initiated by the specific binding of one unique molecule by another that has been selected from millions of other nearby molecules. In other words, the functioning of systems as profoundly complicated as living organisms require a correspondingly large repertoire of molecular codes. To succeed as a coding mechanism, a class of molecules must provide a large capacity for variations in shape because the number of different messages that must be quickly and unambiguously deciphered is tremendous.

For more than 65 years, research efforts to understand information flow in biosystems focused primarily on the nucleic acids DNA and RNA. As a result of this monumental work, life scientists fully expected to find that approximately 100,000 genes existed to code for proteins in humans. Instead, analysis of the data generated by the Human Genome Project (pp. 716–17) produced a much lower number of about 20,000 genes.

Living organisms have two strategies for expanding the coding capacity of their genes: alternative splicing and covalent modification. *Alternative splicing* (described in Chapter 18) is a mechanism whereby eukaryotes produce several polypeptides from the same gene by cutting RNA transcripts and then splicing together various combinations of the RNA fragments. Each type of spliced mRNA product is translated into a unique polypeptide. *Posttranslational modifications* (described in Chapter 19) are enzyme-catalyzed changes in a protein's structure that occur after its

synthesis.

Of all the types of posttranslational modification (e.g., phosphorylation, acetylation, and proteolytic cleavage), glycosylation is the most important in terms of coding capacity, as the following examples illustrate. Recall that just 20 amino acids account for the enormous diversity of proteins observed in living organisms. The total number of hexapeptides that can be synthesized from these amino acids is an impressive 20^6 (6.4×10^7). Carbohydrates, however, have structural properties (e.g., glycosidic linkage variations, branching, and anomeric isomers) that provide them with significantly higher coding capacity. In contrast to peptide linkages, the glycosidic linkages between monosaccharides can be considerably more variable. Consequently, the potential number of permutations in oligosaccharides is substantially higher than that of peptides. For example, the total number of possible linear and branched hexasaccharides that can form from 20 simple or modified monosaccharides is 1.44×10^{15}. Glycans, whether they are attached to proteins or lipids, have one other property: their relative inflexibility (in comparison to peptides), which allows them to bind more precisely with ligands.

Lectins: Translators of the Sugar Code

Once information has been encoded, it must be translated. Sugar code translation is accomplished by lectins. **Lectins** are carbohydrate-binding proteins (CBPs) that are not antibodies and have no enzymatic activity. Originally discovered in plants, they are now known to exist in all organisms. Lectins usually consist of two or four protein or glycoprotein subunits, each with one or more carbohydrate recognition domains (CRDs). Each type of CRD binds to specific carbohydrate groups via hydrogen bonds, van der Waals forces, and hydrophobic interactions. Lectins occur in cell organelles, cytoplasm, cell surfaces, and the extracellular matrix where they serve a variety of functions. Examples include lysosomal enzyme targeting and ER membrane chaperones. The uptake of lysosomal enzymes into lysosomes (p. 52) is mediated by the mannose-6-phosphate receptor, a lysosomal transmembrane lectin that binds to proteins with a mannose-6-phosphate containing N-linked glycan targeting sequence. Within the RER, the lectins calnexin and calreticulin (pp.776–77) promote the refolding of misfolded glycoproteins. Other processes that involve lectin binding (**Figure 7.40**) include infections by microorganisms, the mechanisms of many toxins, and physiological processes such as leukocyte rolling.

Gastritis and Stomach Ulcers

Infection by many bacteria is initiated when the microorganisms become firmly attached to host cells. Often, attachment is mediated by the binding of bacterial lectins to oligosaccharides on the cell's surface. *Helicobacter pylori*, the causative agent of gastritis and stomach ulcers, possesses several lectins that allow it to establish a chronic infection in the mucous lining of the stomach. One of these lectins binds with high affinity to a portion of the type O blood group determinant, an oligosaccharide. This circumstance explains the observation that humans with type O blood are at considerably greater risk of developing ulcers than those with other blood types. Individuals with type A or B blood are not immune to infection, however, since the bacterium can also use other lectins to achieve adhesion. Many viral infections begin with the binding of viral lectins with host

cell surface glycoconjugates. For example, the first step in the multistep infection of epithelial cells by HPV16 (human papilloma virus 16), the causative agent of cervical cancer, begins with viral lectin binding of the heparan sulfate chains of proteoglycans.

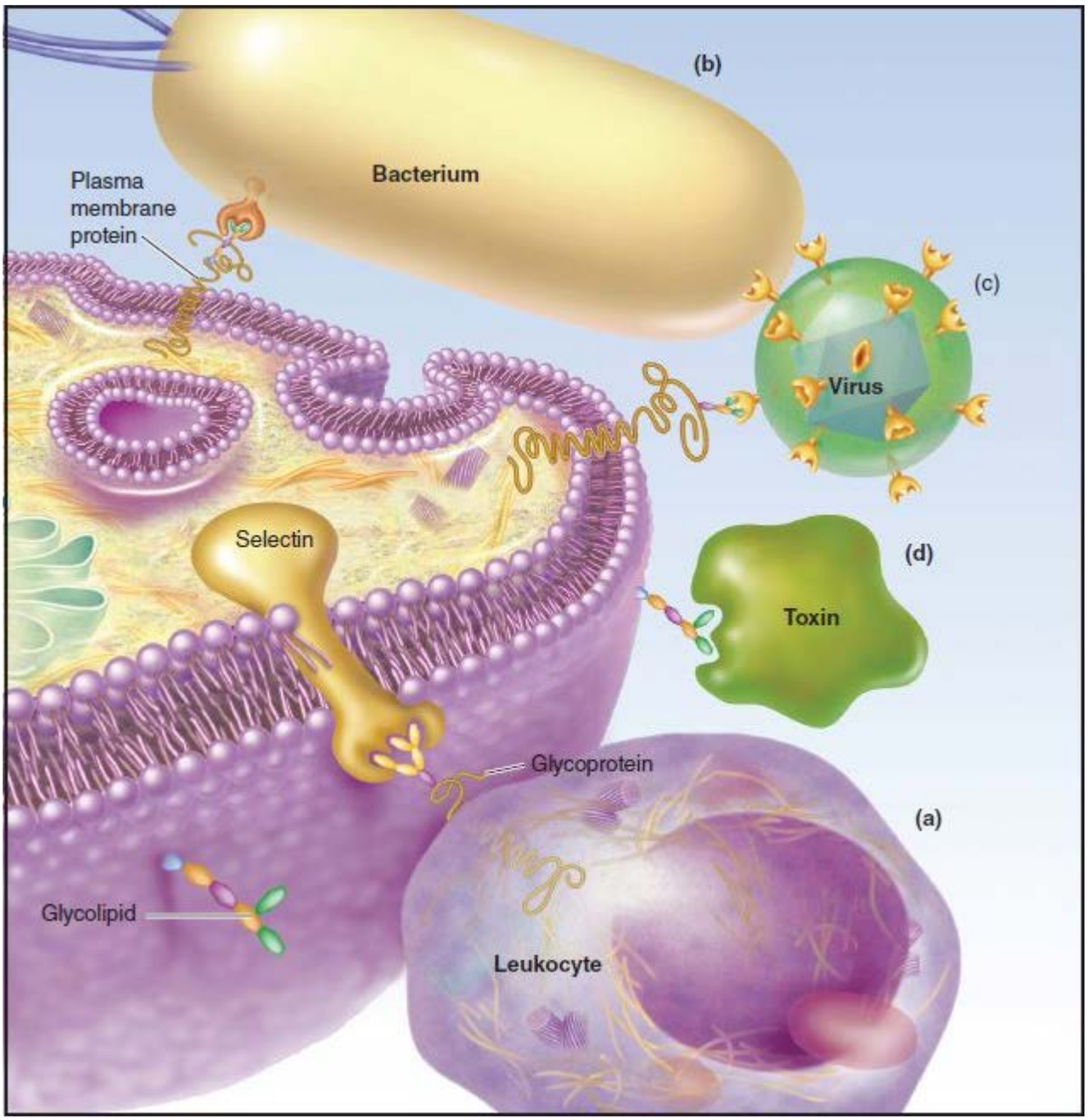


FIGURE 7.40

Role of Glycans in Biological Recognition

The specific binding of lectins (carbohydrate-binding proteins) to the glycan groups of glycoconjugate molecules is an essential feature of many biological phenomena. (a) Cell–cell interactions (e.g., leukocyte rolling), (b and c) cell–pathogen infections, and (d) the binding of toxins (e.g., cholera toxin) to cells. Note that this figure is a diagrammatic illustration of lectin–glycoconjugate binding events. Such binding events are typically multivalent; that is, lectin glycan-binding sites may have interactions with several cell-surface glycoconjugates.

The damaging effects of many bacterial toxins occur only after endocytosis into the host cell, a process that is initiated by lectin-ligand binding. The binding of the B subunit of cholera toxin to a glycolipid on the surface of intestinal cells results in the uptake of the toxic A subunit. Once

internalized, the A subunit proceeds to disrupt the mechanism that regulates chloride transport, a process that results in a life-threatening diarrhea.

Leukocyte rolling is a well-known example of cell–cell interaction mediated by lectin binding. When a tissue becomes damaged in an animal either by infection with a pathogenic organism or by physical trauma, it emits signal molecules that trigger an inflammatory process. In response, the endothelial cells that line nearby blood vessels produce and insert into their plasma membranes first P-selectins and then, within several hours, E-selectins. The selectins are a family of lectins that act as cell adhesion molecules. They bind transiently to the *selectin* ligand (an oligosaccharide) on the surface of white blood cells such as neutrophils. These relatively weak binding events serve to slow the rapid motion of the neutrophils as they flow in blood, causing them to appear to be rolling along the luminal surface of the blood vessel. Once rolling has been initiated and white blood cells approach the inflammation site, they encounter other signal molecules that cause them to express another lectin called *integrin* on their surfaces. The binding of integrin with its oligosaccharide ligand on the endothelial surface of the blood vessel causes the neutrophils to stop rolling. Subsequently, the neutrophils undergo changes that allow them to squeeze between the cells of the endothelium and migrate to the infected site, where they proceed to consume and degrade bacteria or cellular debris.

The Glycome

The term **glycome**, derived from *glyco* (sweet) and *-ome* (as in genome), was created to describe the total set of sugars and glycans that a cell or organism produces. The investigation of glycome structures is referred to as glycomics. Glycomes are constantly in flux because cells responding to environmental signals fine-tune biological responses by altering the glycan structures attached to proteins and lipids. This capacity exists in large part because there is no template for glycan biosynthesis. In contrast to nucleic acid and protein biosynthesis, which are template-driven processes (i.e., multiple identical copies are produced using a nucleotide base sequence), glycans are constructed stepwise, on an assembly line within the ER and Golgi complex. Factors such as variations in sugar nucleotide precursor concentrations and the localization of glycan-processing enzymes can result in variations in the glycan components of each type of glycoprotein, thereby producing a series of slightly different forms called **glycoforms**. This phenomenon, referred to as **microheterogeneity**, may be a means by which cells can generate cell- or tissue-specific signal transduction ligands and/or a mechanism whereby cells elude pathogens whose binding to certain glycan structures initiates an infective process.

KEY CONCEPTS

- The covalent modification of biomolecules such as proteins and lipids by glycosylation reactions provides living organisms with enormous coding capacity.
- The entire set of sugars and glycans produced by a cell or organism is referred to as the glycome.
- Glycoproteins are often produced in slightly different versions called glycoforms.

QUESTION 7.7

The sugar code, with its diverse and subtle nontemplate mechanism for encoding information, has been described as an “analog” system, whereas genetic information processing (DNA- and RNA-directed protein synthesis) is considered “digital.” Explain.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay on sweet medicine and the Biochemistry in the Lab essay on glycomics.

Chapter Summary

1. Carbohydrates, the most abundant organic molecules in nature, are classified as monosaccharides, disaccharides, oligosaccharides, and polysaccharides according to the number of simple sugar units they contain. Carbohydrate moieties also occur as components of other biomolecules. Glycoconjugates are protein and lipid molecules, with covalently linked carbohydrate groups. They include proteoglycans, glycoproteins, and glycolipids.
2. Monosaccharides with an aldehyde functional group are called aldoses; those with a ketone group are known as ketoses. Aldoses belong to either the D or the L family, according to whether the configuration of the asymmetric carbon farthest from the aldehyde group resembles the D- or L-isomer of glyceraldehyde. The D family of aldoses contains most biologically important sugars.
3. Monosaccharides containing five or six carbons exist in cyclic forms that result from reactions between hydroxyl groups and either aldehyde (hemiacetal product) or ketone groups (hemiketal product). In both five-membered rings (furanoses) and six-membered rings (pyranoses), the hydroxyl group attached to the anomeric carbon lies either below (α) or above (β) the plane of the ring for D-sugars. The spontaneous interconversion between α- and β-forms is called mutarotation.
4. Simple sugars undergo a variety of chemical reactions. Derivatives of these molecules, such as uronic acids, amino sugars, deoxy sugars, and phosphorylated sugars, have important roles in cellular metabolism. Glycosylation reactions attach sugars to proteins or lipids. Glycation reactions are nonenzymatic reactions in which reducing sugars react with nucleophilic nitrogens.
5. Hemiacetals and hemiketals react with alcohols to form acetals and ketals, respectively. When the cyclic hemiacetal or hemiketal form of a monosaccharide reacts with an alcohol, the new linkage is called a glycosidic linkage. Molecules in which a sugar is linked via a glycosidic linkage to a noncarbohydrate moiety (aglycone) are called glycosides.
6. Glycosidic bonds form between the anomeric carbon of one monosaccharide and one of the free hydroxyl groups of another monosaccharide. Disaccharides are carbohydrates composed of two monosaccharides. Oligosaccharides, carbohydrates that typically contain as many as 10 to 15 monosaccharide units, are often attached to proteins and lipids. Polysaccharide molecules, which are composed of large numbers of monosaccharide units, may have a linear structure like cellulose and amylose or a branched structure like glycogen and amylopectin. Oligosaccharides and polysaccharides are now referred to as glycans. Glycans may consist of only one sugar type (homoglycans) or multiple types (heteroglycans).
7. The three most common homoglycans found in nature (starch, glycogen, and cellulose) all yield D-glucose when hydrolyzed. Starch and glycogen are storage forms of glucose in plant and animal cells, respectively; cellulose is a plant structural material. Chitin, the principal structural material in insect exoskeletons, is composed of *N*-acetylglucosamine residues linked in unbranched chains. The major classes of heteroglycans, carbohydrate polymers that contain more than one kind of monosaccharide, are N- and O-glycans, glycosaminoglycans, and the glycan components of glycolipids and GPI anchors.
8. The enormous heterogeneity of proteoglycans, which are found predominantly in the extracellular matrix of animal tissues, allows them to play diverse, but as yet poorly understood, roles in living organisms. Glycoproteins occur in cells, in both soluble and membrane-bound forms, and in extracellular fluids. The diverse structures of the glycoconjugates, which include proteoglycans, glycoproteins, and glycolipids, allow them to play important roles in information transfer in living organisms. The glycome is the total set of sugars and glycans that a cell or organism produces.

www.oup.com/us/mckee where you can complete a multiple-choice quiz on carbohydrates to help you prepare for exams.

Chapter 7 Review Quiz

Suggested Readings

Andrews ZB, Horvath TL. 2008 Sep. Why calories taste delicious: eating and the brain. ScientificAmerican.com/Mind Matters.

Boger DL. 2015. When sugar is not so sweet. Science 350:275–6.

Corfield AP, Berry M. 2015. Glycan variation and evolution in the eukaryotes. Trends Biochem Sci 40(7):351–9.

Hennet T, Cabalzer J. 2015. Congenital disorders of glycosylation: a concise chart of glycocalyx dysfunction. Trends Biochem Sci 40(7):377–84.

Iozzo RV, Schaefer L. 2015. Proteoglycan form and function: a comprehensive nomenclature of proteoglycans. Matrix Biol 42:11–55.

Maeder T. 2002. Sweet medicines. Sci Am 287(1):40–7.

Meyers B, Brewer MS. 2008. Sweet taste in man: a review. J Food Sci 73(6):R81–90.

Reed DR, Tanaka T, McDaniel AH. 2006. Diverse tastes: genetics of sweet and bitter taste perception. Physiol Behav 88:215–26.

Rillahan CD, Paulson JC. 2011. Glycan microarrays for decoding the glycome. Ann Rev Biochem 89:797–823.

Seeberger PH. 2005. Exploring life's sweet spot. Nature 437:1239.

Solis D, et al. 2015. A guide into glycosciences: how chemistry, biochemistry and biology cooperate to crack the sugar code. Biochem Biophys Acta 1850:186–215.

Key Words

acetal, *255*

adduct, *256*

aldaric acid, *252*

alditol, *253*

aldonic acid, *252*

aldose, *247*

amylopectin, *262*

amylose, *262*

anomer, *249*

cellobiose, *260*

cellulose, *264*

chitin, *261*

diastereomer, *248*

disaccharide, *255*
enediol, *254*
epimer, *248*
epimerization, *254*
glycan, *261*
glycoconjugate, *266*
glycoform, *273*
glycogen, *263*
glycome, *273*
glycosaminoglycan, *266*
glycoside, *255*
glycosidic linkage, *255*
hemiacetal, *248*
hemiketal, *248*
heparan sulfate proteoglycans, *266*
heteroglycan, *261*
homoglycan, *261*
ketal, *255*
lactone, *252*
lactose, *260*
lectin, *271*
maltose, *260*
microheterogeneity, *273*
monosaccharide, *247*
mutarotation, *251*
N-glycan, *264*
O-glycan, *264*
oligosaccharide, *261*
polysaccharide, *255*
proteoglycan, *266*
proteoglycan aggregate, *267*
reducing sugar, *253*
sucrose, *260*
uronic acid, *252*

Review Questions

SECTION 7.1

Comprehension Questions

1. Define the following terms:

a. monosaccharide
b. aldose
c. ketose
d. epimer
e. diastereomer

2. Define the following terms:

a. pyran
b. furan
c. anomer
d. hemiacetal
e. hemiketal

3. Define the following terms:

a. reducing sugar
b. alditol
c. enediol
d. acetal
e. ketal

4. Define the following terms:

a. lactone
b. aldaric acid
c. aldonic acid
d. uronic acid
e. epimerization

5. Define the following terms:

a. glycoside
b. disaccharide
c. oligosaccharide
d. polysaccharide
e. glycosidic linkage

6. Give an example of each of the following:

a. epimer
b. acetal linkage
c. reducing sugar
d. monosaccharide
e. anomer
f. diastereomer

7. Define the following terms:

a. adduct
b. Schiff base
c. reactive carbonyl-containing product
d. galactosemia
e. Amadori product

Fill in the Blanks

8. _____ are stereoisomers of sugars that differ only in the configuration at an acetal or ketal carbon.
9. D-Erythrose and D-threose are referred to as _____ because their structures only differ in the configuration at one asymmetric carbon atom.
10. Oxidation of the aldehyde group of a sugar produces a(n) _____ acid.
11. The product of the reaction between an aldehyde or ketone and an amine is called a _____ base.
12. Oxidation of the aldehyde group of sugar produces a(n) _____ acid.
13. _____ are nonsuperimposable, non-mirror-image stereoisomers.
14. Sugars that react with Benedict's reagent are called _____.
15. Reduction of the aldehyde group of a monosaccharide yields a _____.

Short-Answer Questions

16. What structural relationship is indicated by the term D-sugar? Why are (+) glucose (shifts polarized light to the right) and (−) fructose (shifts polarized light to the left) both classified as D-sugars?
17. Define the term *reducing sugar*. What structural feature does a reducing sugar have?
18. Classify each of the following sugar pairs as enantiomers, diastereomers, epimers, or an aldose–ketose pair.

    a. D-erythrose and D-threose
    b. D-glucose and D-mannose
    c. D-ribose and L-ribose
    d. D-allose and D-galactose
    e. D-glyceraldehyde and dihydroxyacetone

19. Draw the structure of a disaccharide unit in a polysaccharide composed of D-glucose linked to α(1,4) to D-galactosamine.
20. Pure α-D-glucopyranose has a specific rotation in a polarimeter of +19°, whereas that of β-D-glucopyranose is 112°. The specific rotation of the equilibrating mixture of α- and β-D-glucopyranose produced during mutarotation is +53°. What is the composition of this equilibrating mixture?
21. Why can aldoses and ketoses both behave as reducing sugars?

Critical-Thinking Questions

22. Why is it advantageous for a plant toxin to elicit a bitter taste when an animal eats the plant rather than a bland or sweet taste?
23. Phosphate esters can form at carbons 2 to 6 of an aldohexose but not at carbon 1. Explain.
24. When glucose is reduced, only one alditol is produced. When fructose undergoes the same reaction, however, two diasteriometric sugars are produced. Draw their structures.
25. A newly isolated aldohexose is oxidized to produce the corresponding aldaric acid that has an internal plane of symmetry; that is, it is a symmetric molecule. What is the structure of the original aldohexose?
26. What structures are produced by the epimerization of galactose?
27. The carbohydrate molecule 3-ketoglucose can exist in several ring forms. Draw them and

determine which is the most stable.

28. For a sugar to behave as a reducing sugar, it must have a free aldehyde group. Fructose is a ketose, yet it behaves like a reducing sugar. Explain.

SECTIONS 7.2 & 7.3

Comprehension Questions

29. Define the following terms:

    a. glycan
    b. cellobiose
    c. chitin
    d. amylose
    e. amylopectin

30. Define the following terms:

    a. enterocyte
    b. N-glycan
    c. O-glycan
    d. glycogen
    e. glycosaminoglycan

31. Define the following terms:

    a. lactose
    b. maltose
    c. homoglycans
    d. heteroglycans
    e. starch

Fill in the Blanks

32. _____ is the structural glucose-containing polysaccharide of plants.
33. Polysaccharides that contain more than one type of monosaccharides are called _____.
34. The acetals of glucose and the ketals of fructose are called _____ and _____, respectively.
35. _____ is a disaccharide in which galactose is linked to glucose by a β-glycosidic linkage.
36. _____ are macromolecules with no fixed molecular weights.
37. Animals such as _____ can use cellulose as an energy source.

Short-Answer Questions

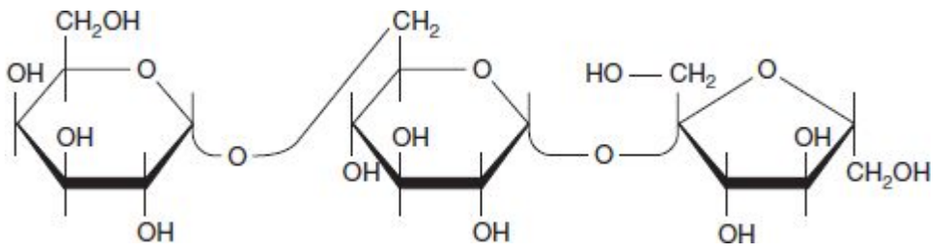
38. What is the difference between a heteroglycan and a homoglycan? Give examples.
39. What structural differences characterize starch, cellulose, and glycogen?
40. Raffinose, the most abundant trisaccharide found in nature, occurs in whole grains and numerous vegetables (e.g., asparagus, cabbage, and beans). Hydrolysis of raffinose yields

galactose and sucrose. Provide the systematic name for this trisaccharide. Is raffinose a reducing or nonreducing sugar? Is raffinose capable of mutarotation?

41. Which of the following carbohydrates are reducing and which are nonreducing?

    a. starch
    b. sucrose
    c. cellulose
    d. ribose
    e. fructose

42. The polymer chains of glycosaminoglycans are widely spread apart and bind large amounts of water. What are the two functional groups of the polymer that make the binding of water possible? What type of bonding is involved?



Critical-Thinking Questions

43. Olestra has been used in certain snack foods as an alternative to fats and oils. Its structure consists of a sucrose molecule in which all free hydroxyl groups have formed esters with oleic acid (an 18-carbon monounsaturated fatty acid). Olestra molecules have no caloric value because they are exceptionally large and cannot be digested. Draw the structure of olestra. Use R-COOH as an abbreviation for oleic acid.
44. Ripe fruit has a high-carbohydrate content and tastes sweet. In contrast, grain, which also has a high-carbohydrate content, does not elicit a sweet sensation when consumed. Suggest a reason for this disparity from the plant's point of view.
45. Steroids are cholesterol-derived, polycyclic, lipid-soluble molecules that are very insoluble in water. Reaction with glucuronic acid makes a steroid much more water-soluble and enables transport through the blood and excretion in urine. What structural feature of the glucuronic acid increases steroid solubility?
46. β-Galactosidase is an enzyme that hydrolyses only β(1,4) linkages of lactose. An unknown trisaccharide is converted by β-galactosidase into maltose and galactose. Draw the structure of the trisaccharide.
47. An oligosaccharide isolated from an organism is found to contain two glucose residues and one galactose residue. Exhaustive methylation followed by hydrolysis produced two glucoses with methoxy groups at positions 2, 3, and 6 and galactose with methoxy groups at positions 2, 3, 4, and 6. What is the structure of the original oligosaccharide?
48. Sucrose does not undergo mutarotation. Explain.

SECTION 7.4

Comprehension Questions

49. Define the following terms:

    a. GPI anchor
    b. glycolipid
    c. glycoprotein
    d. glycoconjugate
    e. proteoglycan

50. Define the following terms:

    a. heparan sulfate proteoglycans
    b. syndecan
    c. glypian
    d. serglycin
    e. mucin

51. Define the following terms:

    a. asparagine-linked oligosaccharide
    b. mucin-type oligosaccharide
    c. Na^+-K^+-ATPase
    d. major histocompatibility antigens
    e. Hurler's syndrome

52. Define the following terms:

    a. CAMs
    b. selectins
    c. cadherins
    d. antifreeze glycoprotein
    e. RNase

Fill in the Blanks

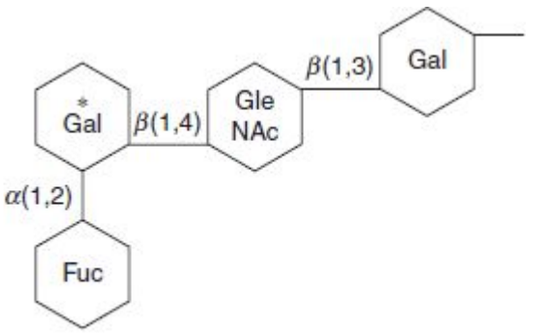
53. _____ is a glycoprotein component of _____, a slippery substance that protects the surfaces of several cell types.
54. _____ and _____ are examples of heparan sulfate proteoglycans.
55. _____ are genetic diseases associated with defects in glycoconjugate synthesis reactions.

Short-Answer Questions

56. In glycoproteins, what are the three amino acid residues to which carbohydrates are most frequently linked? To what functional group is the glycan linked in each case?
57. Chondroitin sulfate chains have been likened to a large fishnet, allowing small molecules to pass through but excluding large ones. What structural features of chondroitin sulfate and proteoglycans in general are responsible for this phenomenon?
58. Compare the structures of proteoglycans and glycoproteins. How are structural differences related to their functions?

Critical-Thinking Questions

59. Proteoglycan aggregates form hydrated, viscous gels in the extracellular matrix of tissues. Can you think of any obvious mechanical reason why their capacity to form gels is an important factor is tissue function?
60. The ABO blood group antigens are the terminal sugars covalently linked to the end of the glycolipid in the red blood cell membrane. The H antigen (seen below) is the precursor of the A and B antigens. Individuals with type A blood have a gene that codes for an enzyme that adds *N*-acetylgalactosamine in an $\alpha(1,3)$ linkage to the Gal* residue to the H antigen. Type B blood requires that an enzyme add a α-D-galactose in an $\alpha(1,3)$ linkage to the Gal*. Draw the structure of the A and B antigens.



H Antigen

61. What will happen to the H antigen precursor (see Question 60) if an individual has both A and B genes? Keep in mind that the substrate of the respective enzymes is the H antigen.
62. Congenital diseases of glycosylation (CDGs) are diseases in which the synthesis of the glycan components of glycoconjugates is impaired. In CDG type 1a, the most common form, the enzyme phosphomannomutase is defective. Symptoms of affected individuals range from early death (before or after birth), developmental delays, poor muscle tone, vision problems, liver, kidney, and heart disorders, and mental disabilities, among others. CDG 1a is an autosomal recessive disease; that is, affected individuals inherit one defective gene from each parent. These parents have no discernible symptoms, and the degree of impairment of affected children varies from mild to severe. Considering that the inheritance of completely nonfunctional variants of this gene is lethal in the early stages of embryonic development, explain these disparities in general terms.

SECTION 7.5

Comprehension Questions

63. Define the following terms:

    a. sugar code
    b. lectin
    c. glycoform
    d. glycome
    e. microheterogeneity

64. Define the following terms:

    a. Heliobacter pylori
    b. leukocyte rolling

c. cholera toxin
d. HPV16
e. glycomics

Fill in the Blanks

65. The _____ are a family of lectins on the surface of endothelial cells that act as cell adhesion molecules.
66. There are approximately _____ protein-coding genes in the human genome.
67. Slightly different versions of a glycan component of glycoprotein are called _____.

Short-Answer Questions

68. Describe the process of leucocyte rolling.
69. Describe the major features of how cholera toxin damages intestinal cells.
70. Why are individuals with type O blood at high risk for stomach ulcers?

Critical-Thinking Questions

71. Calculate how many hexapeptides modified with one hexasaccharide can be synthesized from 20 amino acids and 20 simple or modified monosaccharides.
72. There are approximately 200 cell types in the human body. Throughout development, the pattern of surface glycoconjugates on each cell type changes. Speculate about what factors are responsible for these changes.

MCAT Study Questions

73. Fructose is classified as a(n)

    a. aldotriose
    b. aldohexose
    c. ketopentose
    d. ketohexose

74. Amylopectin has which of the following structural features?

    a. only α(1,4) glycosidic linkages
    b. only β(1,4) glycosidic linkages
    c. both α(1,4) and α(1,6) glycosidic linkages
    d. both α(1,4) and β(1,6) glycosidic linkages

75. Which of the following is a nonreducing sugar?

    a. cellobiose
    b. maltose
    c. lactose
    d. sucrose

76. How many isomers of glucose can exist?

    a. 4
    b. 8

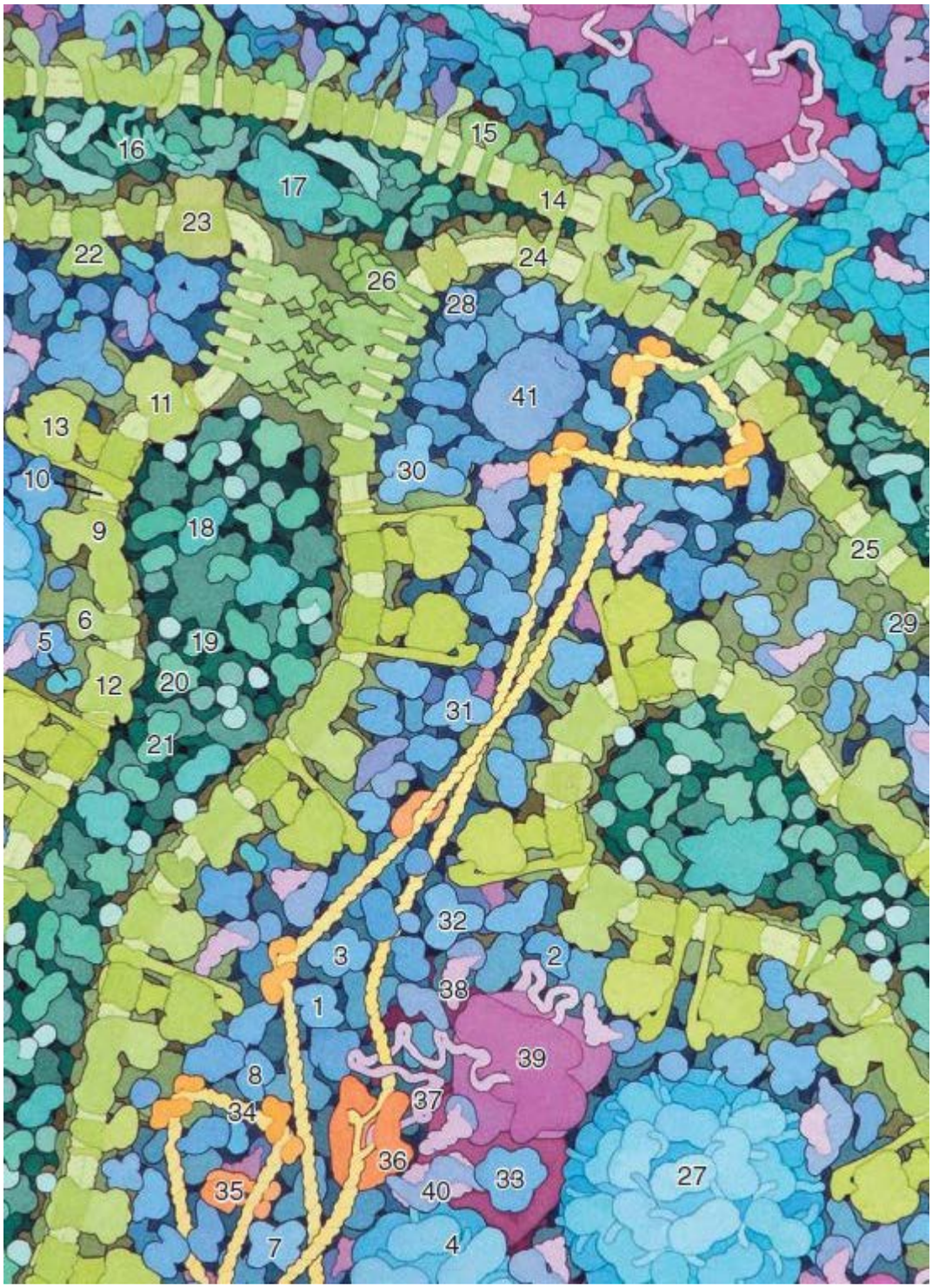
c. 16
d. 32

77. Which of the following molecules gives a positive Benedict’s test?

a. glycogen
b. sucrose
c. amylose
d. ribose

CHAPTER 9

Aerobic Metabolism I: The Citric Acid Cycle



Cross Section of a Mitochondrion This illustration, based on electron microscopy and X-ray crystallography, shows a cross section of a mitochondrion in magnified molecular detail (1×10^6). Each component is numbered. The citric acid cycle enzymes are (1) citrate synthase, (2) aconitase, (3) isocitrate dehydrogenase, (4) α-ketoglutarate dehydrogenase, (5) succinyl CoA synthetase, (6) succinate dehydrogenase, (7) fumarase, and (8) malate dehydrogenase. The inner membrane electron transport chain components are (9) NADH dehydrogenase, (6) succinate dehydrogenase, (10) coenzyme Q, (11) cytochrome bc1 reductase, (12) cytochrome oxidase, and (13) ATP synthase. Note that for clarity the citric acid cycle enzymes and the electron transport system molecules (Chapter 10) are shown as separate molecules. They are believed to function as metabolons. Other mitochondrion components mentioned elsewhere are (14) VDAC (p. 380), (15) monoamine oxidase, (16) TIM (p. 783), (17) creatine kinase (p. 378), (18) glycerol-3-phosphate dehydrogenase (p. 380), (19) adenylate kinase (p. 154), (20) nucleoside diphosphate kinase (p. 557), (21) sulfate oxidase (p. 587), (22) Mg^{2+} transporter, (23) calcium channel, (24) ADP/ATP translocator (p. 378), (25) potassium channel, (26) dynamin-like protein (regulates mitochondrial fusion and cristae structure), (27) pyruvate dehydrogenase (p.

342), (28) Mn-superoxide dismutase (p. 389), (29) Cu-Zn superoxide dismutase (p. 389), (30) pyruvate carboxylase (p. 302), (31) acyl CoA dehydrogenase (p.453), (32) ornithine transcarbamoylase (p. 578), (33) glutamate dehydrogenase (p. 536), (34) DNA, (35) steroid receptor (p. 614), (36) RNA polymerase (p. 725), (37) mRNA, (38) tRNA, (39) ribosome, (40) amino acyl-tRNA transferase (p. 757), (41) Hsp60 (p. 172).

OUTLINE

OXYGEN AND EVOLUTION: CHANCE AND NECESSITY

9.1 OXIDATION-REDUCTION REACTIONS
Redox Coenzymes
Aerobic Metabolism

9.2 CITRIC ACID CYCLE
Conversion of Pyruvate to Acetyl-CoA
Reactions of the Citric Acid Cycle
Fate of Carbon Atoms in the Citric Acid Cycle
The Amphibolic Citric Acid Cycle
Citric Acid Cycle Regulation
The Citric Acid Cycle and Human Disease
The Glyoxylate Cycle

Biochemistry in Perspective
Carcinogenesis: The Warburg Effect and Metabolic Reprogramming

AVAILABLE ONLINE
Biochemistry in Perspective
Hans Krebs and the Citric Acid Cycle
Biochemistry in Perspective
The Evolutionary History of the Citric Acid Cycle

Oxygen and Evolution: Chance and Necessity

Aerobic organisms use oxygen (O_2) to extract energy from nutrient molecules. *Aerobic respiration*, the oxygen-requiring energy-generating mechanism, generates significantly more energy than does fermentation, an anaerobic (without oxygen) process. When it first appeared about 2 billion years ago (bya), aerobic respiration marked a critical turning point in the evolution of life. The large increases in energy it supplied to organisms with the molecular equipment to exploit oxygen provided the resources for evolutionary innovation. It is not coincidental that eukaryotic cells (about 1.5 bya) and multicellular organisms (about 1 bya) originated after aerobic respiration became a common means of energy generation. Oxygen accumulated in the primordial atmosphere because it is a waste product of *oxygenic photosynthesis* (Chapter 13), a 3-billion-year-old process initiated by cyanobacteria that uses sunlight to drive the synthesis of biomolecules from CO_2 and the hydrogen atoms of water. Oxygenic photosynthesis is so complex and thermodynamically challenging, however, that it might never have come about at all.

Why Oxygen?

Why is oxygen so useful in energy generation? The answer lies in the nature of energy generation and the chemistry of oxygen and water. Energy is captured when electrons are transferred from an electron donor to an electron acceptor. In aerobic respiration, oxygen acts as the terminal acceptor of electrons removed from organic nutrients. As O_2 combines with these electrons as well as protons, water molecules are formed. Oxygen is an excellent electron acceptor for two major reasons. First, the element oxygen is abundant, which makes it far more useful as an oxidizing agent than less commonly available elements such as sulfur. Second, oxygen is a powerful oxidizing agent because it is highly electronegative (i.e., it has a considerable affinity for electrons). Consequently, there is a correspondingly large energy release with every electron that is transferred from carbon to oxygen.

Why Is Oxygenic Photosynthesis Such a Challenge?

The origin of oxygenic photosynthesis is improbable for two major reasons. First, because oxygen atoms are powerful electron acceptors, water is a very poor reducing agent. In contrast, nonoxygenic photosynthesis, the precursor of the oxygen-generating process, used more powerful reducing agents such as H_2 and H_2S. Significantly more light energy is required to oxidize water molecules than for H_2 and H_2S. As a result, the photosynthetic pigment molecules that absorb light energy had to be adapted to absorbing higher-energy photons (electromagnetic particles). Second, photosynthetic electron transport (one electron at a time) and water oxidation (a sequential and concerted four-electron process) are seemingly incompatible. The cyanobacteria solved this extraordinarily complex set of problems with a multisubunit pigment–enzyme complex now referred to as *photosystem II* (PSII).

The One and Only Oxygen-Evolving Complex

Within PSII, water molecules are split into electrons, protons, and O_2 by the *oxygen-evolving complex* (OEC), a protein complex that contains a unique inorganic cube-like cofactor: the Mn_3CaO_4 cubane. Of all of the problems solved by the cyanobacteria, it is the OEC that provides insight into the tenuous origin of oxygenic photosynthesis. Unlike all other biochemical devices, the OEC was invented only once. Consequently, all modern photosynthesizing organisms contain an identical water-splitting mechanism. The emergence of the water-splitting OEC of PSII was the most pivotal event in the history of life on Earth. Without the OEC, all modern organisms would be anaerobic prokaryotes.

Overview

MODERN AEROBIC ORGANISMS TRANSDUCE THE CHEMICAL BOND ENERGY OF FOOD MOLECULES INTO THE BOND ENERGY OF ATP. THEY PERFORM THIS feat because oxygen is used as the terminal acceptor of the electrons extracted from food molecules. The capacity to use oxygen to oxidize nutrients such as glucose and fatty acids yields a substantially greater amount of energy than does fermentation.

As atmospheric O_2 began accumulating on Earth, about 2 bya, existing organisms were confronted with a serious problem: molecular oxygen forms toxic oxygen ions and peroxides called *reactive oxygen species* (ROS). ROS react with and damage or destroy biomolecules. Consequently, exposure to O_2 acted as a severe selection pressure. Species in existence during this period either evolved a means of adapting to O_2 or became extinct. Modern

organisms are classified based on the strategies they use to cope with ROS or use O_2 in energy generation:

1. **Obligate anaerobes** are organisms that grow only in the absence of O_2 (i.e., they live in highly reduced environments such as soil) and use fermentation to generate energy.
2. **Aerotolerant anaerobes** depend on fermentation but possess enzymes and antioxidant molecules that detoxify ROS.
3. **Facultative anaerobes** not only possess the biochemical mechanisms required for detoxifying ROS but also can use O_2 as an electron acceptor when the gas is available.
4. **Obligate aerobes** are highly dependent on O_2 for energy production. They protect themselves from ROS with elaborate detoxifying mechanisms that are composed of enzymes and numerous endogenous and exogenous antioxidant molecules.

Aerobic metabolism consists of three biochemical processes: the citric acid cycle, the electron transport pathway, and oxidative phosphorylation (**Figure 9.1**). In eukaryotes, these processes occur within the mitochondrion (**Figure 9.2**). The citric acid cycle is a metabolic pathway in which two-carbon fragments derived from organic fuel molecules are oxidized to form CO_2 and the coenzymes NAD^+ and FAD are reduced to form NADH and $FADH_2$, respectively. The electron transport pathway, also called the electron transport chain (ETC), is a mechanism by which electrons are transferred from NADH and $FADH_2$ to a series of electron carriers that are sequentially reduced and then oxidized. The terminal electron acceptor is O_2. In oxidative phosphorylation, the energy released by electron transport is captured in the form of a proton gradient that drives the synthesis of ATP, the energy currency of living organisms.

Chapter 9 begins with a review of oxidation-reduction reactions and the relationship between electron flow and energy transduction. This review is followed by a detailed discussion of the citric acid cycle, the central pathway in aerobic metabolism, and its roles in energy generation and biosynthesis. In Chapter 10, the discussion of aerobic metabolism continues with an examination of electron transport and oxidative phosphorylation, the means by which aerobic organisms use oxygen to generate significant amounts of ATP. Chapter 10 ends with a review of *oxidative stress*, a series of reactions in which toxic oxygen species are created and subsequently damage cells. Chapter 10 also describes the principal mechanisms used by living organisms to protect themselves from oxidative stress.

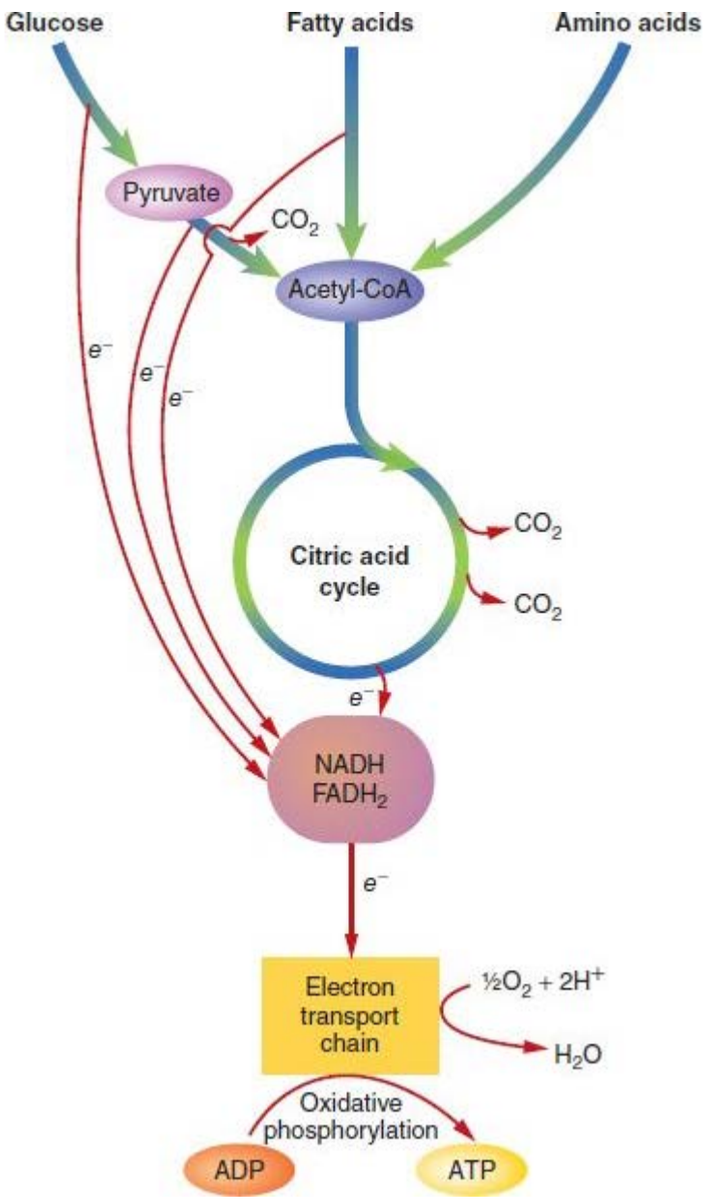


FIGURE 9.1

Overview of Aerobic Metabolism

In aerobic metabolism, the nutrient molecules glucose, fatty acids, and some amino acids are degraded to form acetyl-CoA. Acetyl-CoA then enters the citric acid cycle. Electron carriers (NADH and $FADH_2$) produced by glucose and fatty acid degradation and several citric acid cycle reactions donate electrons (e^-) to the electron transport chain. Energy captured by the electron transport chain is then used to synthesize ATP in a process referred to as oxidative phosphorylation. Note that O_2, the terminal electron acceptor in aerobic metabolism, combines with protons to form water molecules.

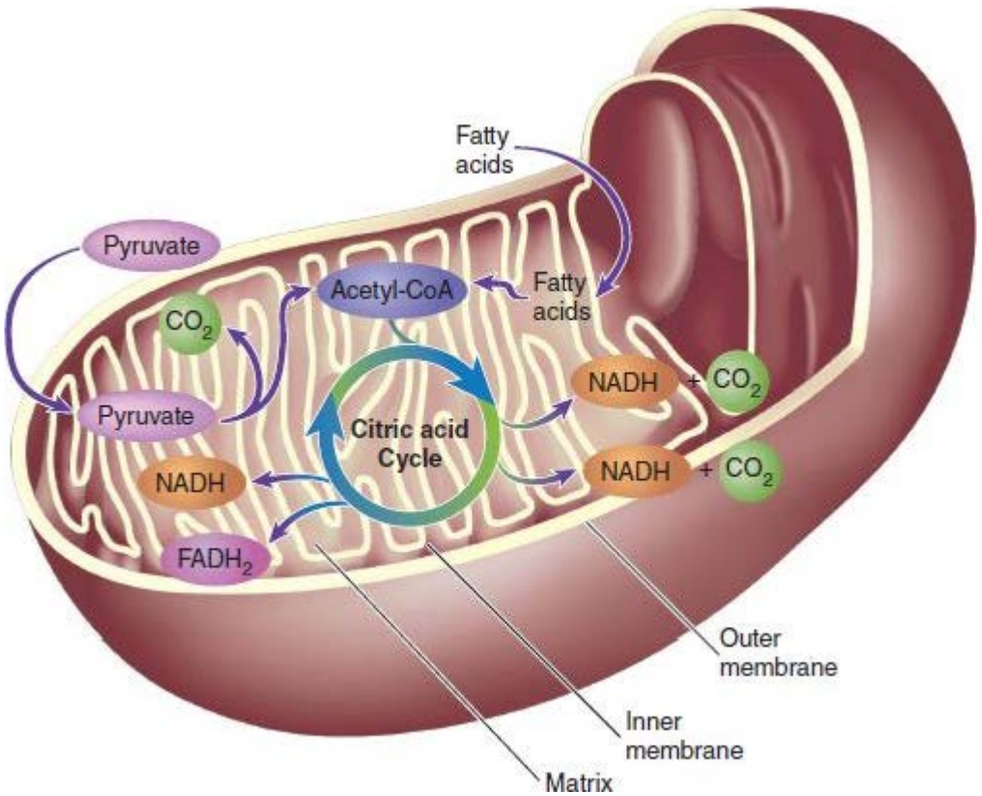


FIGURE 9.2

Aerobic Metabolism in the Mitochondrion

In eukaryotic cells, aerobic metabolism occurs within the mitochondrion. Acetyl-CoA, the oxidation product of pyruvate, fatty acids, and certain amino acids (not shown), is oxidized by the reactions of the citric acid cycle within the mitochondrial matrix. The principal products of the cycle are the reduced coenzymes, NADH and $FADH_2$, and CO_2. The high-energy electrons of NADH and $FADH_2$ are subsequently donated to the electron transport chain (ETC), a series of electron carriers in the inner membrane. The terminal electron acceptor for the ETC is O_2. The energy derived from the electron transport mechanism drives ATP synthesis by creating a proton gradient across the inner membrane. The large folded surface of the inner membrane is studded with ETC complexes, transport proteins of numerous types, and ATP synthase, the enzyme complex responsible for ATP synthesis. Small organic anions such as pyruvate, malate, and succinate, as well as ATP, ADP, NADH, and NAD^+, are exchanged across the outer mitochondrial membrane by VDAC (not shown). VDAC (voltage-dependent anion channels) allow the diffusion of small hydrophilic molecules as well as Ca^{2+} into and out of the intermembrane space. Proteins such as hexokinase, glucokinase (p. 297), glycerol kinase (p. 306), and creatine kinase (p. 380) can bind to VDAC.

9.1 OXIDATION-REDUCTION REACTIONS

In living organisms, both energy-capturing and energy-releasing processes consist largely of redox reactions. Recall that redox reactions occur when electrons are transferred between an electron donor (a reducing agent) and an electron acceptor (an oxidizing agent). In some redox reactions, only electrons are transferred. For example, in the reaction

$$Cu^+ + Fe^{3+} \rightleftharpoons Cu^{2+} + Fe^{2+}$$

an electron is transferred from Cu^+ to Fe^{3+}. Cu^+, the reducing agent, is oxidized to form Cu^{2+}. Meanwhile, Fe^{3+} is reduced to Fe^{2+}. In many reactions, however, both electrons and protons are transferred. For example, the reaction catalyzed by lactate dehydrogenase begins with the transfer

of a hydride ion ($H:^-$), that is, a hydrogen nucleus and two electrons, from NADH to pyruvate. A proton (H^+) is gained from the environment (**Figure 9.3**) to form the final products lactate and NAD^+.

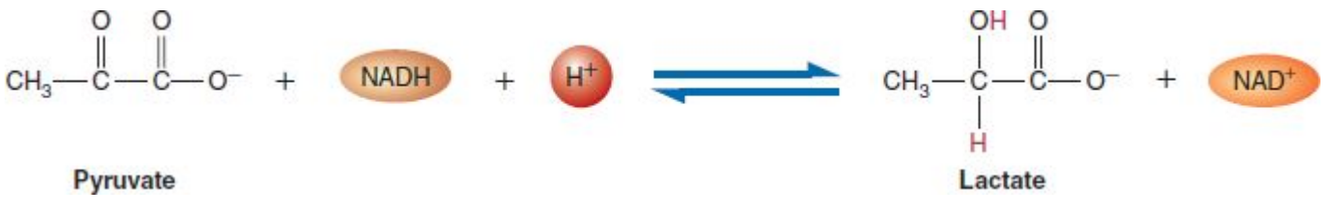


FIGURE 9.3

Reduction of Pyruvate by NADH

In this redox reaction, a hydride ion ($H:^-$) is transferred from NADH to pyruvate, and the product is protonated from the surrounding medium to form lactate.

Redox reactions are more easily understood if they are separated into half-reactions. In the reaction between copper and iron, the Cu^+ ion loses an electron to become Cu^{2+}:

$$Cu^+ \rightleftharpoons Cu^{2+} + e^-$$

This equation indicates that Cu^+ is the electron donor. (Together Cu^+ and Cu^{2+} constitute a **conjugate redox pair**.) As Cu^+ loses an electron to form Cu^{2+}, Fe^{3+} gains an electron to form Fe^{2+}:

$$Fe^{3+} + e^- \rightleftharpoons Fe^{2+}$$

In this half-reaction, Fe^{3+} is an electron acceptor. The separation of redox reactions emphasizes that electrons are always the common intermediates between half-reactions.

The constituents of half-reactions may be observed in an electrochemical cell (**Figure 9.4**). Each half-reaction takes place in a separate container, or *half-cell*. The movement of electrons generated in the half-cell undergoing oxidation (e.g., $Cu^+ \rightarrow Cu^{2+} + e^-$) creates a voltage (or potential difference) between the two half-cells. The sign of the voltage (measured by a voltmeter) is positive or negative according to the direction of the electron flow. The magnitude of the potential difference is a measure of the energy that drives the reaction.

The tendency for a specific substance to gain electrons is called its **reduction potential**. The **standard reduction potential** ($E°$) of a substance is measured in a galvanic cell relative to a standard hydrogen electrode. A standard cell has all solutes at 1.0 M concentration, all gases at 1 atm pressure, and the temperature at 25°C. The reduction potential for the half-reaction, $2H^+ + 2e^- \rightarrow H_2$ (g), against the standard hydrogen electrode is set at 0.00 V.

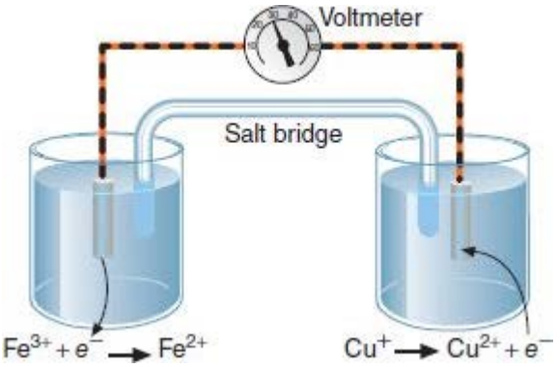


FIGURE 9.4

An Electrochemical Cell

Electrons flow from the Cu^{2+}/Cu^+ half-cell (cathode) through the voltmeter to the Fe^{3+}/Fe^{2+} half-cell (anode). The salt bridge containing KCl completes the electrical circuit. The voltmeter measures the electrical potential, which drives electrons from one half-cell to the other.

In biochemistry, the reference half-reaction is

$$2H^+ + 2e^- \rightleftharpoons H_2$$

when pH = 7
temperature = 25°
pressure = 1 atm

Under these conditions, the reduction potential ($E°'$) of the hydrogen electrode is –0.42 V when measured against the standard hydrogen electrode in which the hydrogen ion concentration is 1 M. Substances with reduction potentials lower than –0.42 V (i.e., those with more negative values) have a lower affinity for electrons than does H^+. Substances with higher reduction potentials (i.e., those with more positive values) have a greater affinity for electrons (**Table 9.1**). The pH in the test electrode is 7.0 for each of the redox half-reactions, and the pH of the reference standard electrode is 0 or the $[H^+]$ is 1.0 M.

A substance with a more negative (less positive) reduction potential will receive electrons from a substance with a more positive reduction potential, and the overall cell potential ($\Delta E°'$) will be positive. The relationship between $\Delta E°'$ and $\Delta G°'$ is

$$\Delta G°' = -nF\,\Delta E°'$$

TABLE 9.1 Standard Reduction Potentials*

Redox Half-Reaction	Standard Reduction Potentials ($E°'$) (V)
$2H^+ + 2e^- \rightarrow H_2$	−0.42
α-Ketoglutarate + CO_2 + $2H^+$ + $2e^-$ → isocitrate	−0.38
$NADP^+ + H^+ + 2e^- \rightarrow$ NADPH	−0.324
+ + −	

$NAD + H + 2e \rightarrow NADH$	−0.32
$S + 2H^+ + 2e^- \rightarrow H_2S$	−0.23
$FAD + 2H^+ + 2e^- \rightarrow FADH_2$	−0.22
Acetaldehyde + $2H^+ + 2e^- \rightarrow$ ethanol	−0.20
Pyruvate + $2H^+ + 2e^- \rightarrow$ lactate	−0.19
Oxaloacetate + $2H^+ + 2e^- \rightarrow$ malate	−0.166
$Cu^{2+} + e^- \rightarrow Cu^+$	+0.16
Fumarate + $2H^+ + 2e^- \rightarrow$ succinate	+0.031
Cytochrome b (Fe^{3+}) + $e^- \rightarrow$ cytochrome b (Fe^{2+})	+0.075
Cytochrome c_1 (Fe^{3+}) + $e^- \rightarrow$ cytochrome c_1 (Fe^{2+})	+0.22
Cytochrome c (Fe^{3+}) + $e^- \rightarrow$ cytochrome c (Fe^{2+})	+0.235
Cytochrome a (Fe^{3+}) + $e^- \rightarrow$ cytochrome a (Fe^{2+})	+0.29
$NO_3^- + 2H^+ + 2e^- \rightarrow NO_2^- + H_2O$	+0.42
$NO_2^- + 8H^+ + 6e^- \rightarrow NH_4^+ + 2H_2O$	+0.44
$Fe^{3+} + e^- \rightarrow Fe^{2+}$	+0.77
$^1/_2O_2 + 2H^+ + 2e^- \rightarrow H_2O$	+0.82

* By convention, redox reactions are written with the reducing agent to the right of the oxidizing agent and the number of electrons transferred. In this table, the redox pairs are listed in order of increasing $E°'$ values. The more negative the $E°'$ value is for a redox pair, the lower the affinity of the oxidized component for electrons. The more positive the $E°'$ value is, the greater the affinity of the oxidized component of the redox pair for electrons. Under appropriate conditions, a redox half-reaction reduces any of the half-reactions below it in the table.

where $\Delta G°'$ = the standard free energy at pH 7
n = the number of electrons transferred
F = the Faraday constant (96,485 J/V·mol)
$\Delta E°'$ = the difference in reduction potential between the electron donor and the electron acceptor under standard conditions

Living organisms utilize redox coenzymes as high-energy electron carriers. The most prominent examples are described next.

Redox Coenzymes

The coenzyme forms of the vitamin molecules nicotinic acid and riboflavin are universal electron carriers. Their structural and functional properties are as follows.

NICOTINIC ACID There are two coenzyme forms of nicotinic acid: **nicotinamide adenine dinucleotide (NAD)** and **nicotinamide adenine dinucleotide phosphate (NADP)**. These coenzymes occur in oxidized forms (NAD^+ and $NADP^+$) and reduced forms (NADH and NADPH). The structures of NAD^+ and $NADP^+$ both contain adenosine and the *N*-ribosyl derivative of nicotinamide (derived from the vitamin niacin), which are linked together through a pyrophosphate group (**Figure 9.5a**). $NADP^+$ has an additional phosphate attached to the 2′ OH group of adenosine. (The ring atoms of the sugar in a nucleotide are designated with a prime to distinguish them from atoms in the base.) Both NAD^+ and $NADP^+$ carry electrons for several enzymes in a group known as the dehydrogenases. (Dehydrogenases catalyze hydride transfer reactions. Many dehydrogenases that catalyze reactions involved in energy generation use the coenzyme NADH. The enzymes that require NADPH usually catalyze biosynthetic reactions (e.g., fatty acid synthesis, pp. 460–66, and quench [detoxify] ROS, pp. 389–90). A small number of dehydrogenases can use either NADH or NADPH.)

Recall that alcohol dehydrogenase catalyzes the reversible oxidation of ethanol to form acetaldehyde (p. 216).

During this reaction NAD^+ accepts a hydride ion from ethanol, the substrate molecule undergoing oxidation. The product deprotonates to form the acetaldehyde molecule. The reversible reduction of NAD^+ is illustrated in **Figure 9.5b**.

In most reactions catalyzed by dehydrogenases, the NAD^+ (or $NADP^+$) is bound only transiently to the enzyme. After the reduced version of the coenzyme is released from the enzyme, it donates the hydride ion to another molecule, called an *electron acceptor*, with a more positive reduction potential than NADH.

RIBOFLAVIN Riboflavin (vitamin B_2) is a component of two coenzymes: **flavin mononucleotide (FMN)** and **flavin adenine dinucleotide (FAD)** (**Figure 9.6**). FMN and FAD function as tightly bound prosthetic groups in a class of enzymes known as the **flavoproteins**. Flavoproteins are a diverse group of redox enzymes; they function as dehydrogenases, oxidases, and hydroxylases. These enzymes use the isoalloxazine group of FAD or FMN as a donor or acceptor of two hydrogen atoms. FMN plays a key role in the link between two-electron transfer reactions in the mitochondrial matrix and the one-electron transfer reactions of the electron transport chain because it can transfer one hydrogen atom at a time. Succinate dehydrogenase is a prominent example of a flavoprotein. It catalyzes the oxidation of succinate by FAD to form fumarate and $FADH_2$, an important reaction in the citric acid cycle.

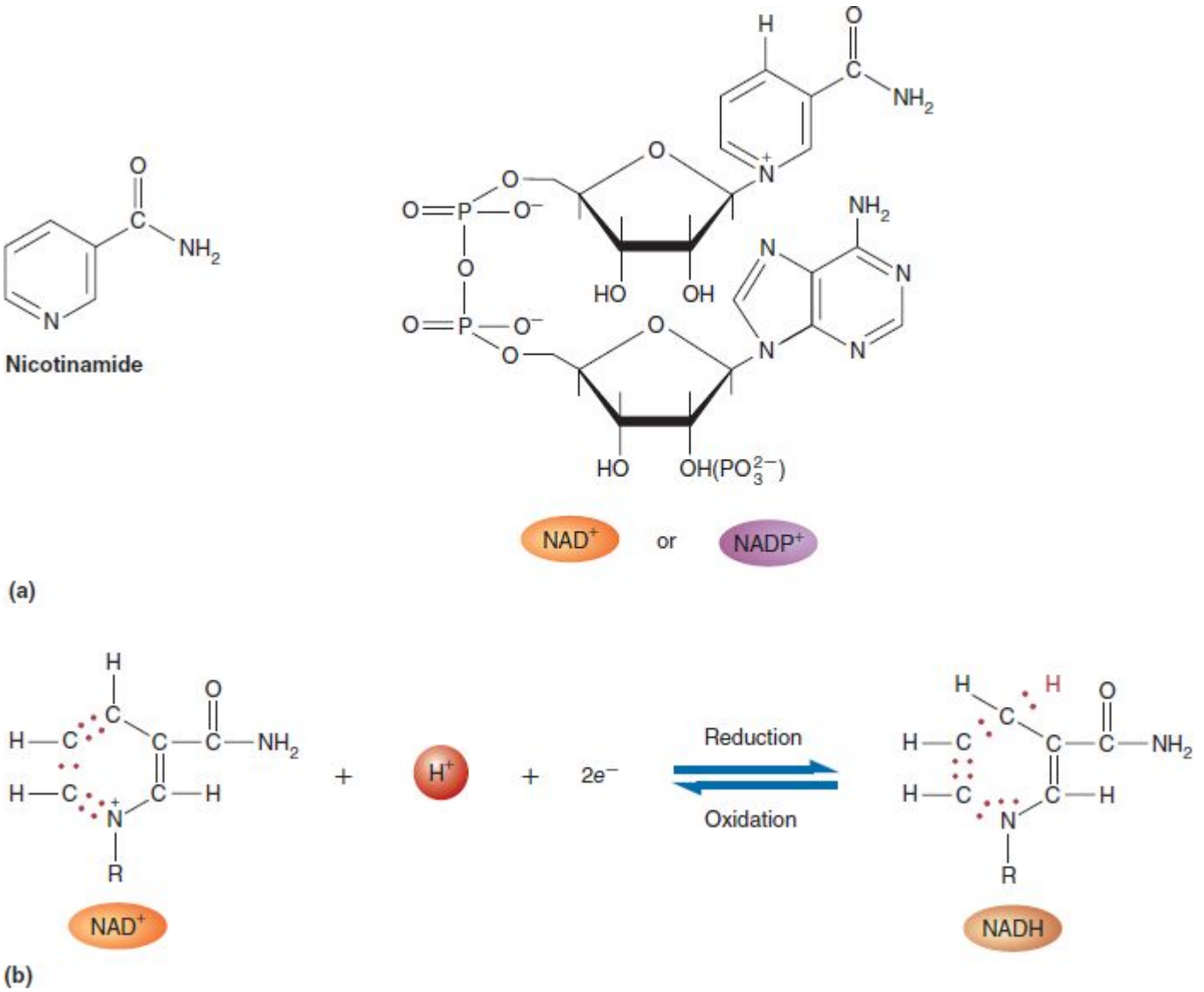


FIGURE 9.5

Nicotinamide Adenine Dinucleotide (NAD)

(a) Nicotinamide and $NAD(P)^+$. (b) Reversible reduction of NAD^+ to NADH. To simplify the equation, only the nicotinamide ring is shown. The rest of the molecule is designated R.

QUESTION 9.1

Use **Table 9.1** to determine which of the following reactions will proceed as written:

$$CH_3CH_2OH + 2\ \text{cyt b}\ (Fe^{3+}) \rightarrow CH_3CHO + 2\ \text{cyt b}\ (Fe^{2+}) + 2H^+$$

$$NO_2^- + H_2O + 2\ \text{cyt b}\ (Fe^{3+}) \rightarrow 2\ \text{cyt b}\ (Fe^{2+}) + NO_3^- + 2H^+$$

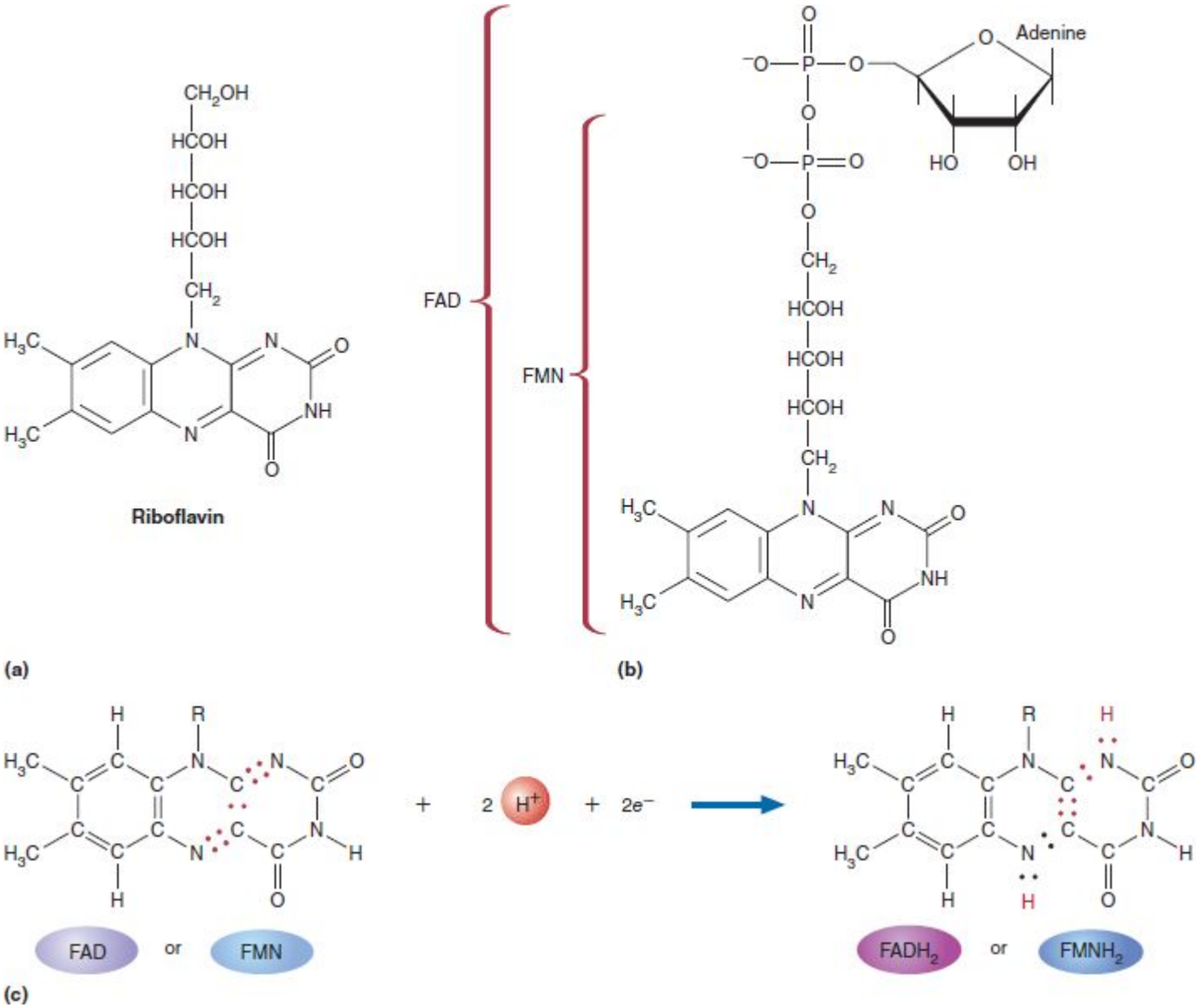


FIGURE 9.6

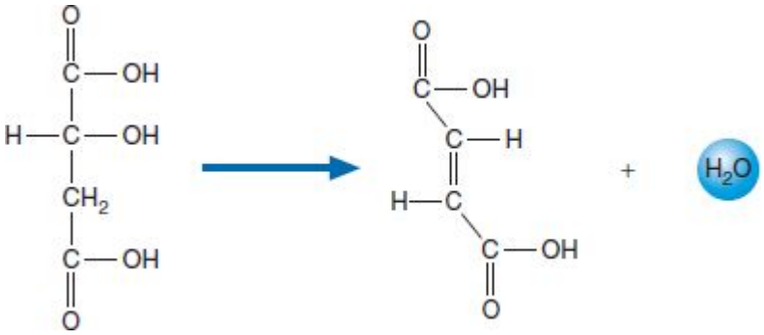
Flavin Coenzymes

(a) The vitamin riboflavin consists of an isoalloxazine ring system linked to ribitol (an alcohol formed by the reduction of ribose). (b) Structure of FAD and FMN. (c) Reversible reduction of flavin coenzymes: to simplify the equation, only the isoalloxazine ring system is shown. The rest of the coenzyme is designated R.

QUESTION 9.2

Which of the following reactions are redox reactions? For each redox reaction, identify the oxidizing and reducing agents.

1. Glucose + ATP → glucose-1-phosphate + ADP

2. 

3. Lactate + $NAD^+ \rightarrow$ pyruvate + NADH + H^+

4. $NO_2^- + 8H^+ + 6$ cyt b $(Fe^{2+}) \rightarrow NH_4^+ + 2H_2O + 6$ cyt b (Fe^{3+})

5. $CH_3CHO + NADH + H^+ \rightarrow CH_3CH_2OH + NAD^+$

WORKED PROBLEM 9.1

Use the following half-cell potentials to calculate (a) the overall cell potential and (b) $\Delta G^{\circ\prime}$.

$$\text{Succinate} + {}^1/_2O_2 \rightarrow \text{fumarate} + H_2O$$

The half-reactions are

$$\text{Succinate} \rightarrow \text{fumarate} + 2H^+ + 2e^- \qquad (E^{\circ\prime} = -0.031\ \text{V})$$
$$\tfrac{1}{2}\ O_2 + 2H^+ + 2e^- \rightarrow H_2O \qquad (E^{\circ\prime} = +0.82\ \text{V})$$

SOLUTION

Write the fumarate reaction as an oxidation (lower reduction potential), balance if the number of electrons transferred differs between the two half-reactions (not necessary here), and add the two reactions to get the net reaction.

$$\text{Succinate} + {}^1/_2O_2 \rightarrow \text{fumarate} + H_2O$$

a. The overall potential is defined by the following:

$$\Delta E^{\circ\prime} = E^{\circ\prime}\ (\text{electron acceptor}) - E^{\circ\prime}\ (\text{electron donor})$$
$$\Delta E^{\circ\prime} = (+0.82\ \text{V}) - (-0.031\ \text{V})$$
$$\Delta E^{\circ\prime} = +0.85\ \text{V}$$

b. Use the formula to find $\Delta G^{\circ\prime}$.

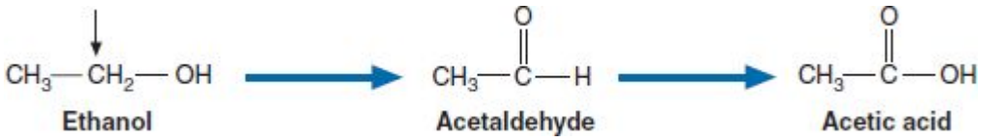
$$\Delta G^{\circ\prime} = -nF\Delta E^{\circ\prime}$$
$$= -(2)(96.5\ \text{kJ/V.mol})(0.85\ \text{V})$$
$$= -164.05\ \text{kJ/mol}$$
$$= -164\ \text{kJ/mol}$$

■

QUESTION 9.3

Because redox reactions play an important role in living processes, biochemists need to determine the oxidation state of the atoms in a molecule. In one method, the oxidation state of an atom is determined by assigning numbers to carbon atoms based on the type of groups attached to them. For example, a bond to a hydrogen is assigned the value −1. A bond to another carbon atom is valued at 0, and a bond to an electronegative atom such as oxygen or nitrogen is valued at +1. The values of a single carbon atom in a molecule may range from −4 (e.g., CH_4) to +4 (CO_2). Note that methane is a high-energy molecule and carbon dioxide is a low-energy molecule. As carbon changes its oxidation state from −4 to +4, a large amount of energy is released. This process is therefore highly exothermic.

Ethanol is degraded in the liver by a series of redox reactions. Identify the oxidation state of the indicated carbon atom in each molecule in the following reaction sequence:



QUESTION 9.4

As CO_2 is incorporated into a sugar molecule during photosynthesis, is it being oxidized or reduced?

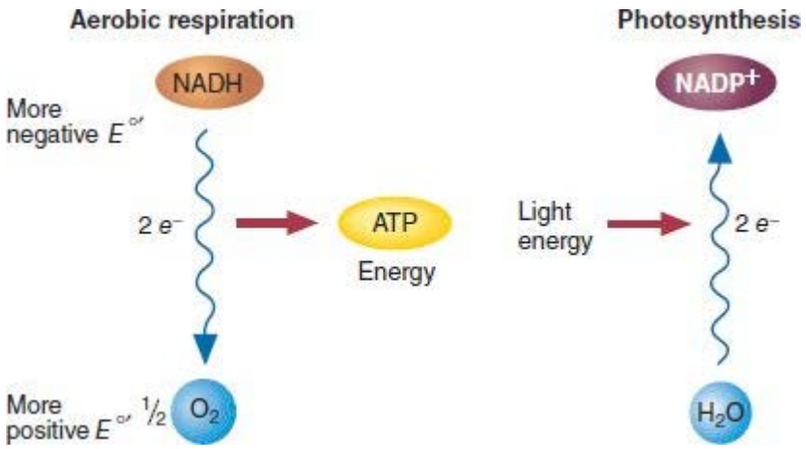


FIGURE 9.7

Electron Flow and Energy

Electron flow is used to generate and capture energy in aerobic respiration. Radiant (light) energy is used to drive electron flow in photosynthesis. Note that the energy captured by photosynthesis in the chemical bonds of sugars and other biomolecules is released by aerobic respiration and used to synthesize ATP.

Aerobic Metabolism

Most of the aerobic cell's free energy is captured by the mitochondrial ETC. During this process,

electrons are transferred from a redox pair with a more negative reduction potential ($NADH/NAD^+$) to those with more positive reduction potentials. The last component in the system is the $H_2O/½\ O_2$ pair:

$$½\ O_2 + NADH + H^+ \rightarrow H_2O + NAD^+$$

The free energy released as a pair of electrons passes from NADH to O_2 under standard conditions is calculated as follows:

$$\begin{aligned}\Delta G^{\circ\prime} &= -nF\Delta E^{\circ\prime} \\ &= -2(96.5\ kJ/V{\cdot}mol)[0.815 - (-0.32)] \\ &= -220\ kJ/mol\end{aligned}$$

A significant portion of the free energy generated as electrons move from NADH to O_2 in the ETC is used to drive ATP synthesis.

KEY CONCEPTS

- In living organisms, both energy-capturing and energy-releasing processes consist primarily of redox reactions.
- In redox reactions, electrons move between an electron donor and an electron acceptor.
- In many reactions, both electrons and protons are transferred.
- In biological systems, most redox reactions involve hydride ion transfer ($NAD(P)H/NAD(P)^+$) or hydrogen atom transfer ($FADH_2/FAD$).

In several metabolic processes, electrons move from redox pairs with more positive reduction potentials to those with more negative reduction potentials. Of course, energy is required. The most prominent example of this phenomenon is photosynthesis (Chapter 13). Photosynthetic organisms use captured light energy to drive electrons from electron donors, such as water, to electron acceptors with more negative reduction potentials (**Figure 9.7**). The energized electrons eventually flow back to acceptors with more positive reduction potentials, thereby providing energy for ATP synthesis and CO_2 reduction to form carbohydrate.

In Section 9.2, the citric acid cycle is examined. In this pathway, which is the first phase of aerobic metabolism, the energy released by the oxidation of two-carbon fragments derived from glucose, fatty acids, and some amino acids is captured by and carried in the reduced coenzymes NAD(P)H and $FADH_2$.

9.2 CITRIC ACID CYCLE

The citric acid cycle (**Figure 9.8**) is a series of biochemical reactions that aerobic organisms use to release chemical energy stored in the two-carbon acetyl group in acetyl-CoA. Acetyl-CoA is composed of an acetyl group derived from the breakdown of carbohydrates, lipids, and some amino acids that is linked to the acyl carrier molecule **coenzyme A** (**Figure 9.9**). Acetyl-CoA is synthesized from pyruvate and is also the product of fatty acid catabolism (described in Chapter 11) and certain reactions in amino acid metabolism (Chapter 15). In the citric acid cycle, the acetyl group's carbon atoms are eventually oxidized to CO_2, and the electrons are transferred to NAD^+ and FAD.

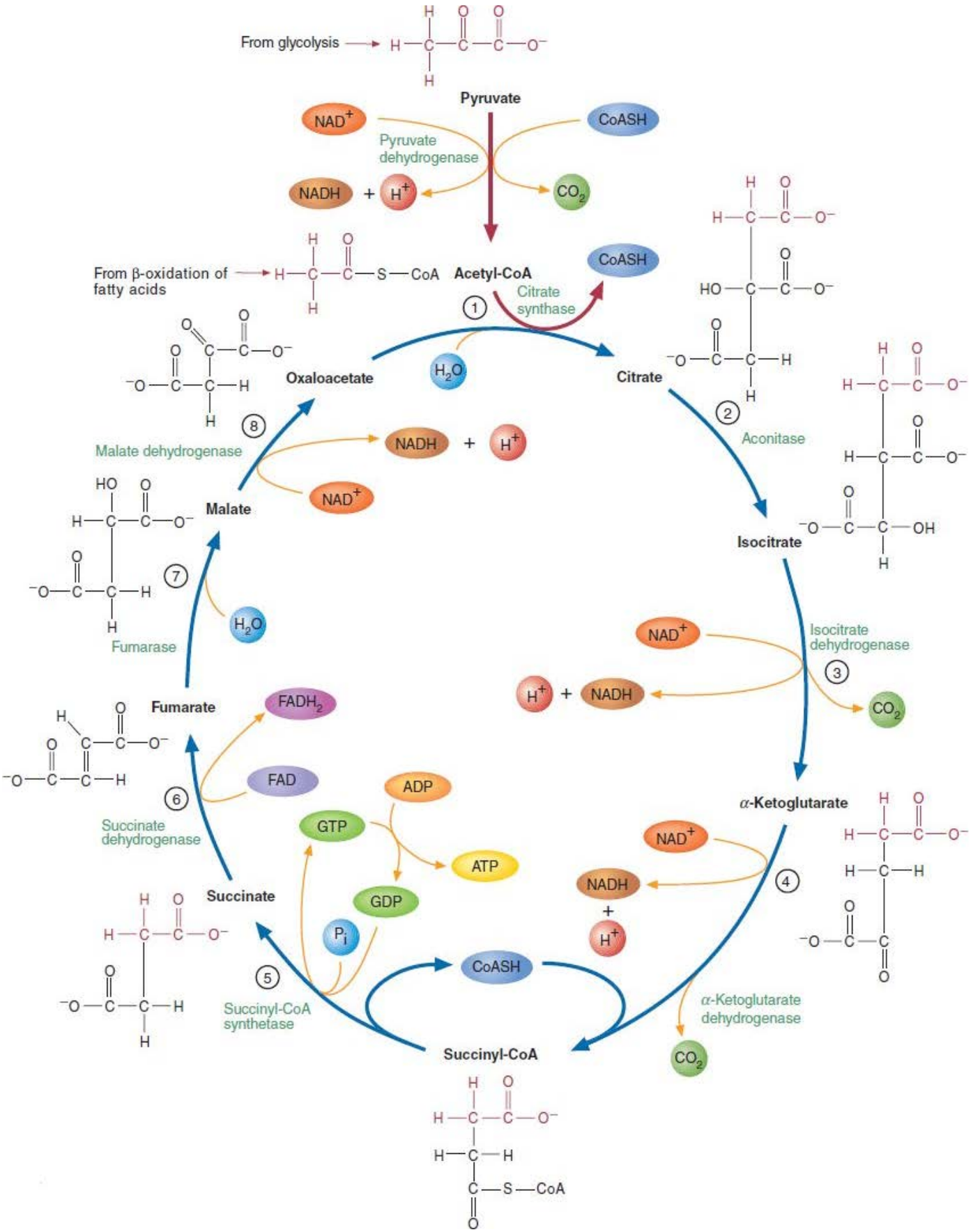


FIGURE 9.8

The Citric Acid Cycle

In each turn of the cycle, acetyl-CoA from the glycolytic pathway or from fatty acid catabolism enters, and two fully oxidized carbon molecules leave as CO_2. Three molecules of NAD^+ and the molecule of FAD are reduced. One molecule of GTP (interconvertible with ATP) is generated in a substrate-level phosphorylation reaction.

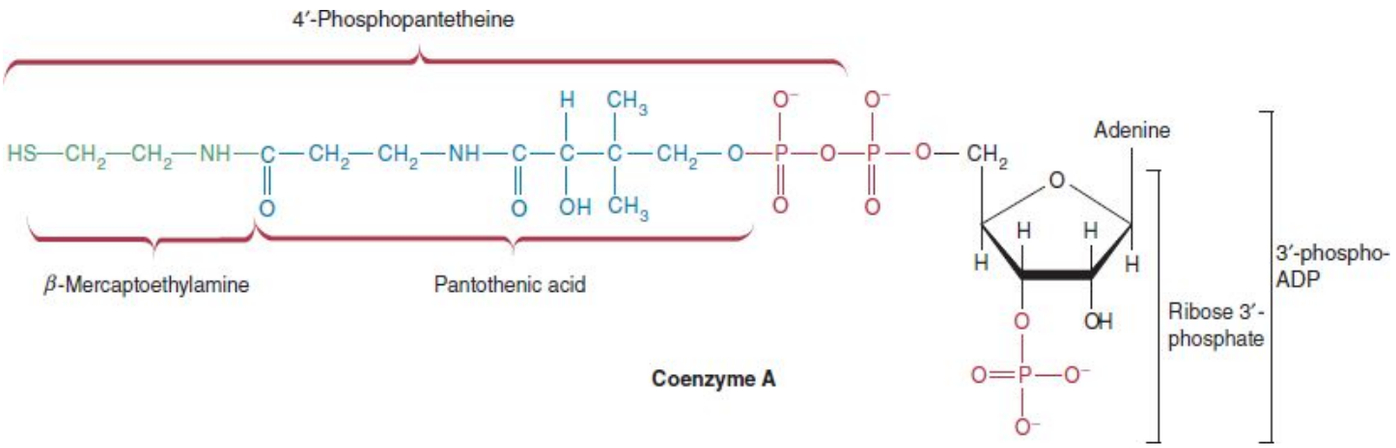


FIGURE 9.9

Coenzyme A

In coenzyme A, a 3′-phosphate derivative of ADP is linked to pantothenic acid via a phosphate ester bond. The β-mercaptoethylamine group of coenzyme A is attached to pantothenic acid by an amide bond. Coenzyme A is a carrier of acyl groups that range in size from the acetyl group to long-chain fatty acids. Because the reactive SH group forms a thioester bond with acyl groups, coenzyme A is often abbreviated as CoASH. Note that sulfur is a better leaving group than oxygen. Consequently, the carbon–sulfur linkage of a thioester is a high-energy bond that is more easily cleaved than the carbon–oxygen bond of an ester.

In the first reaction of the citric acid cycle, a two-carbon acetyl group condenses with a four-carbon molecule (oxaloacetate) to form a six-carbon molecule (citrate). During the subsequent seven reactions, in which two CO_2 molecules are produced and four pairs of electrons are removed from carbon compounds, citrate is reconverted to oxaloacetate. During one step in the cycle, the high-energy molecule GTP is produced during a substrate-level phosphorylation. The net reaction for the citric acid cycle is as follows:

$$\text{Acetyl-CoA} + 3\,NAD^+ + FAD + GDP + P_i + 2\,H_2O \rightarrow 2\,CO_2 + 3\,NADH + FADH_2 + CoASH + GTP + 2H^+$$

In addition to its role in energy production, the citric acid cycle plays another important role in metabolism. Cycle intermediates are substrates in a variety of biosynthetic reactions. **Table 9.2** provides a summary of the roles of coenzymes in the citric acid cycle.

Conversion of Pyruvate to Acetyl-CoA

After its transport into the mitochondrial matrix, pyruvate is converted to acetyl-CoA in a series of reactions catalyzed by the enzymes in the pyruvate dehydrogenase complex (PDHC). The net reaction, an oxidative decarboxylation, is as follows:

$$\text{Pyruvate} + NAD^+ + \text{CoASH} \rightarrow \text{Acetyl-CoA} + NADH + CO_2 + H^+$$

TABLE 9.2 Summary of the Coenzymes in the Citric Acid Cycle

Coenzyme	Functions
Thiamine pyrophosphate (TPP)	Decarboxylation and aldehyde group transfer
Lipoic acid	Carrier of hydrogens or acetyl groups
NADH	Electron carrier
$FADH_2$	Electron carrier
Coenzyme A (CoASH)	Acetyl group carrier

Despite the apparent simplicity of this highly exergonic reaction ($\Delta G°' = -33.5$ kJ/mol), its mechanism is one of the most complex known. The PDHC (**Figure 9.10**) is a large multienzyme structure that contains multiple copies of three enzyme activities: pyruvate dehydrogenase (E_1), dihydrolipoyl transacetylase (E_2), and dihydrolipoyl dehydrogenase (E_3). **Table 9.3** summarizes the number of copies of each enzyme and the required coenzymes of the *E. coli* pyruvate dehydrogenase complex.

In the first step, pyruvate dehydrogenase catalyzes the decarboxylation of pyruvate (**Figure 9.11**). A nucleophile is formed when a basic residue of the enzyme extracts a proton from the thiazole ring of **thiamine pyrophosphate** (TPP). (TPP is the coenzyme form of thiamine, also called vitamin B_1.) The intermediate, hydroxyethyl-TPP (HETPP), forms after the nucleophilic thiazole ring has attacked the carbonyl group of pyruvate, with the resulting loss of CO_2.

In the next several steps, the hydroxyethyl group of HETPP is converted to acetyl-CoA by dihydrolipoyl transacetylase. **Lipoic acid** (**Figure 9.12**), an acyl transfer coenzyme that contains two thiol groups that can be reversibly oxidized, plays a crucial role in this transformation. Lipoic acid is bound to the enzyme through an amide linkage with the ε-amino group of a lysine side chain, forming a flexible arm. Lipoamide reacts with HETPP to form an acetylated lipoamide and free TPP. The lipoamide then transfers the acetyl group to the sulfhydryl group of coenzyme A. Subsequently, the reduced lipoamide is transferred to the dihydrolipoyl dehydrogenase active site where it is reoxidized by a bound FAD to yield $FADH_2$. The reduction potential of $FADH_2$ is sufficiently altered by its association with the enzyme that it can donate a hydride ion to an incoming NAD^+. (Ordinarily, $FADH_2$ cannot reduce NAD^+ because NAD^+ has a more negative reduction potential.) The mobile NADH can deliver its electrons to the ETC and is replaced in the enzyme by another NAD^+ molecule so that the cycle can begin again.

KEY CONCEPTS

Pyruvate is converted to acetyl-CoA by the enzymes in the pyruvate dehydrogenase complex. TPP, FAD, NAD^+, Coenzyme A, and lipoic acid are required coenzymes.

PDHC is stringently regulated because of its central role in energy metabolism, linking glycolysis to the citric acid cycle. Its activity is controlled largely through allosteric effectors and

+

covalent modification. The enzyme complex is allosterically activated by NAD , CoASH, and AMP. It is inhibited by high concentrations of ATP and the reaction products acetyl-CoA and NADH. In mammals, acetyl-CoA and NADH also activate pyruvate dehydrogenase kinase 1 (PDK1), which converts the active PDHC to an inactive phosphorylated form. High concentrations of the substrates pyruvate, CoASH, and NAD^+ inhibit the activity of the kinase. The PDHC is reactivated by a dephosphorylation reaction catalyzed by pyruvate dehydrogenase phosphatase (PDP). PDP is activated when the mitochondrial ATP concentration is low. PDP is also activated by Ca^{2+} and insulin.

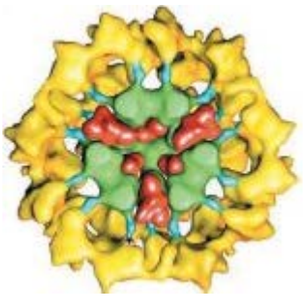


FIGURE 9.10

Pyruvate Dehydrogenase Complex

This cross section of the PDHC (*Saccharomyces cerevisiae*) shows the relative positions of the E_1 subunits (pyruvate dehydrogenase, yellow), the E_2 subunits (dihydrolipoyl transacetylase, green), the lipoyl domains of E_2 (blue, reaching outward to the active sites of E_1), and the E_3 subunits (dihydrolipoyl dehydrogenase) in the core (red), which can also be reached by the swinging arms of the lipoamide domains of E_2.

TABLE 9.3 *E. coli* Pyruvate Dehydrogenase Complex

Enzyme Activity	Function	Copies per Complex*	Coenzymes
Pyruvate dehydrogenase (E_1)	Decarboxylates pyruvate	24 (20–30)	TPP
Dihydrolipoyl transacetylase (E_2)	Catalyzes transfer of acetyl group to CoASH	24 (60)	Lipoic acid, CoASH
Dihydrolipoyl dehydrogenase (E_3)	Reoxidizes dihydrolipoamide	12 (20–30)	NAD^+, FAD

* The number of molecules of each enzyme activity found in mammalian pyruvate dehydrogenase is shown in parentheses.

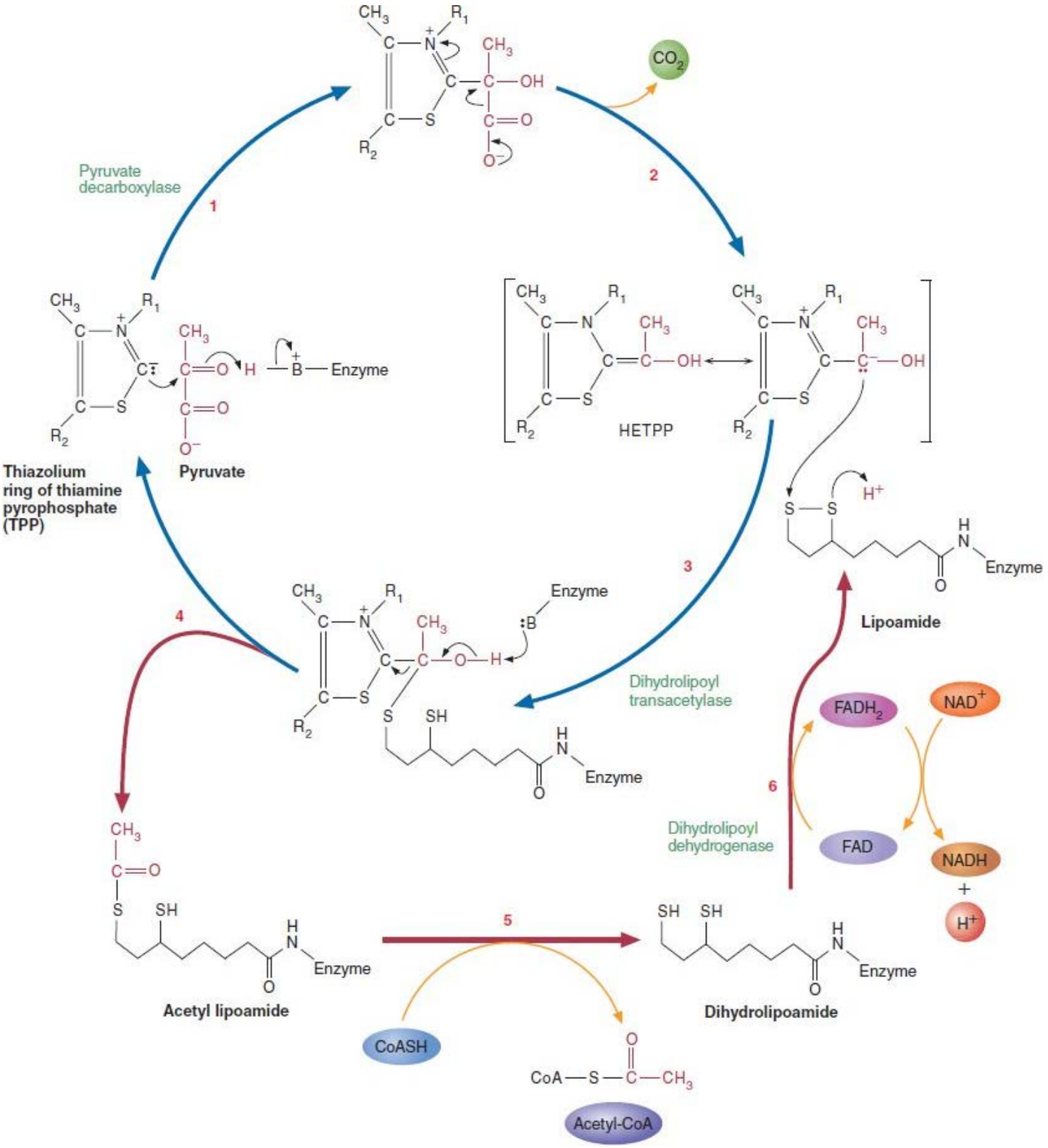


FIGURE 9.11

The Reactions Catalyzed by the Pyruvate Dehydrogenase Complex

Initially, pyruvate dehydrogenase (decarboxylase) catalyzes the formation of HETTP (hydroxyethyl-TPP) using TPP as a coenzyme. In step 1, TPP in a zwitterionic form performs a nucleophilic attack on the C2 carbonyl of pyruvate to produce the reactive intermediate that decarboxylates (step 2) to yield CO_2 and HETPP. In step 3, dihydrolipoyl transacetylase uses lipoamide (lipoic acid covalently linked to a lysine side chain) to convert the hydroxyethyl group of HETPP to a thioester of lipoamide. A resonance stabilized HETPP attacks S1 of oxidized lipoamide (step 4) to yield acetyl lipoamide, thereby converting S2 into a sulfhydryl group. A subsequent transacetylation reaction (step 5) transfers the acetyl group to the thiol group of coenzyme A. In step 6, the FAD in dihydrolipoyl dehydrogenase reoxidizes the reduced lipoamide, yielding $FADH_2$ and

dihydrolipoamide. FAD is regenerated when $FADH_2$ donates a hydride ion to NAD^+. (Refer to Figure 9.12 for the structure of lipoamide.)

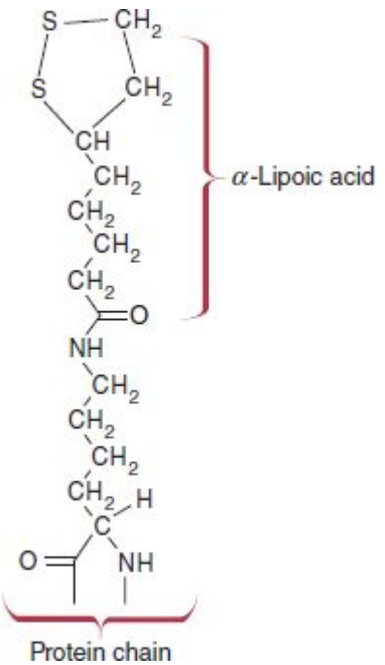


FIGURE 9.12

Lipoamide

Lipoic acid is covalently bonded to the enzyme through an amide linkage with the ε-amino group of a lysine residue.

3D animation of lipoamide

Reactions of the Citric Acid Cycle

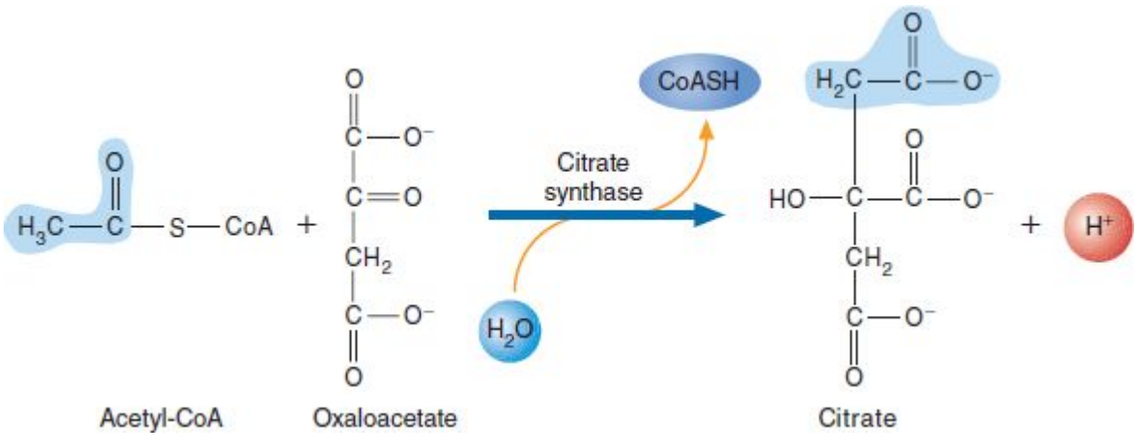
The citric acid cycle is composed of eight reactions that occur in two stages:

1. The two-carbon acetyl group of acetyl-CoA enters the cycle by reacting with the four-carbon compound oxaloacetate (OAA); two molecules of CO_2 are subsequently liberated (reactions 1–4).
2. OAA is regenerated, so it can react with another acetyl-CoA (reactions 5–8).

The enzymes that catalyze these reactions associate through noncovalent interactions into metabolons (p. 232), multienzyme complexes that ensure efficient channeling of the product of each reaction to the next enzyme in the pathway.

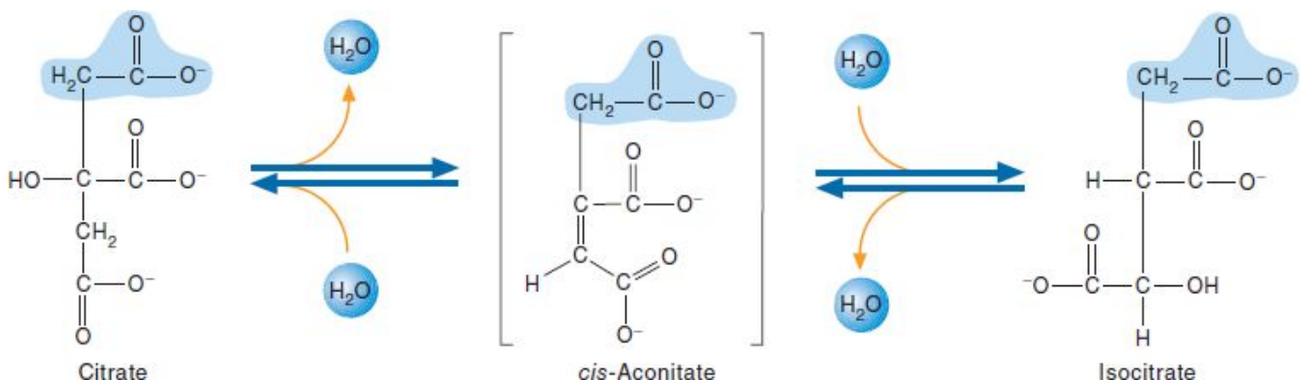
The reactions of the citric acid cycle are as follows.

1. **Introduction of two carbons as acetyl-CoA.** The citric acid cycle begins with the condensation of acetyl-CoA with oxaloacetate to form citrate:



The enzyme citrate synthase is a homodimer in which the active site is a cleft between the two subunits. The binding of oxaloacetate in the open form of the enzyme causes a structural change that induces a transition to the closed form while simultaneously forming the acetyl-CoA binding site. In this reaction (**Figure 9.13**). the enzyme removes a proton from the methyl group of acetyl-CoA, thereby converting it to an enol (p. 290). The enol subsequently attacks the C2 carbonyl carbon of oxaloacetate. The product, citroyl-CoA, rapidly hydrolyzes to form citrate and CoASH. Because of the hydrolysis of the high-energy thioester bond, the overall standard free energy change is −33.5 kJ/mol, and citrate formation is highly exergonic.

2. **Citrate is isomerized to form a secondary alcohol that can be easily oxidized.** In the next reaction of the cycle, citrate, which contains a tertiary alcohol, is reversibly converted to isocitrate by aconitase. The reaction begins with the protonation of the C3 hydroxyl group by a histidine side chain, which promotes the release of a water molecule. A serine residue then abstracts a proton from C2 to form the double bond of the intermediate *cis*-aconitate. After flipping of the intermediate within the active site, the enzyme reverses the steps. It catalyzes the hydration of the double bond when a now basic histidine abstracts a proton from a water molecule, thus triggering a nucleophilic attack on C2, and the protonated serine is deprotonated by the double bond to yield the product isocitrate, a more reactive molecule. Although the standard free energy change of citrate isomerization is positive ($\Delta G^{\circ\prime}$ = 13.3 kJ), the reaction is pulled forward by the rapid removal of isocitrate by the next reaction.



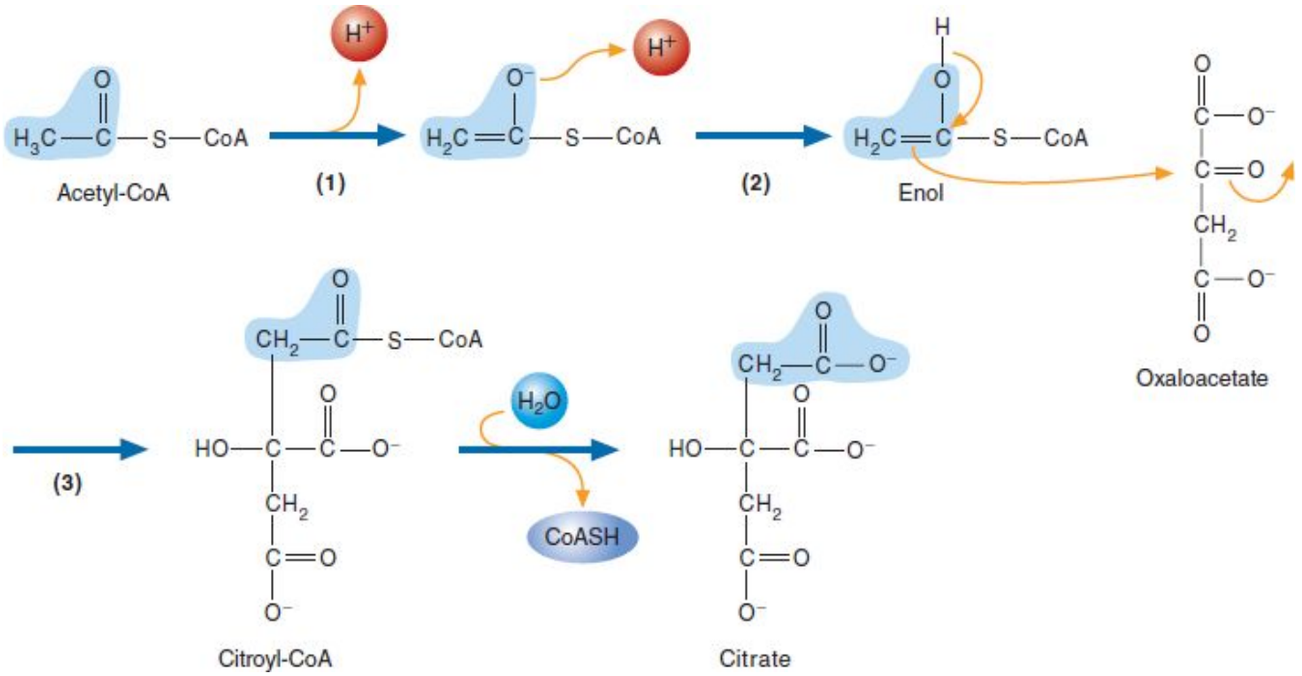
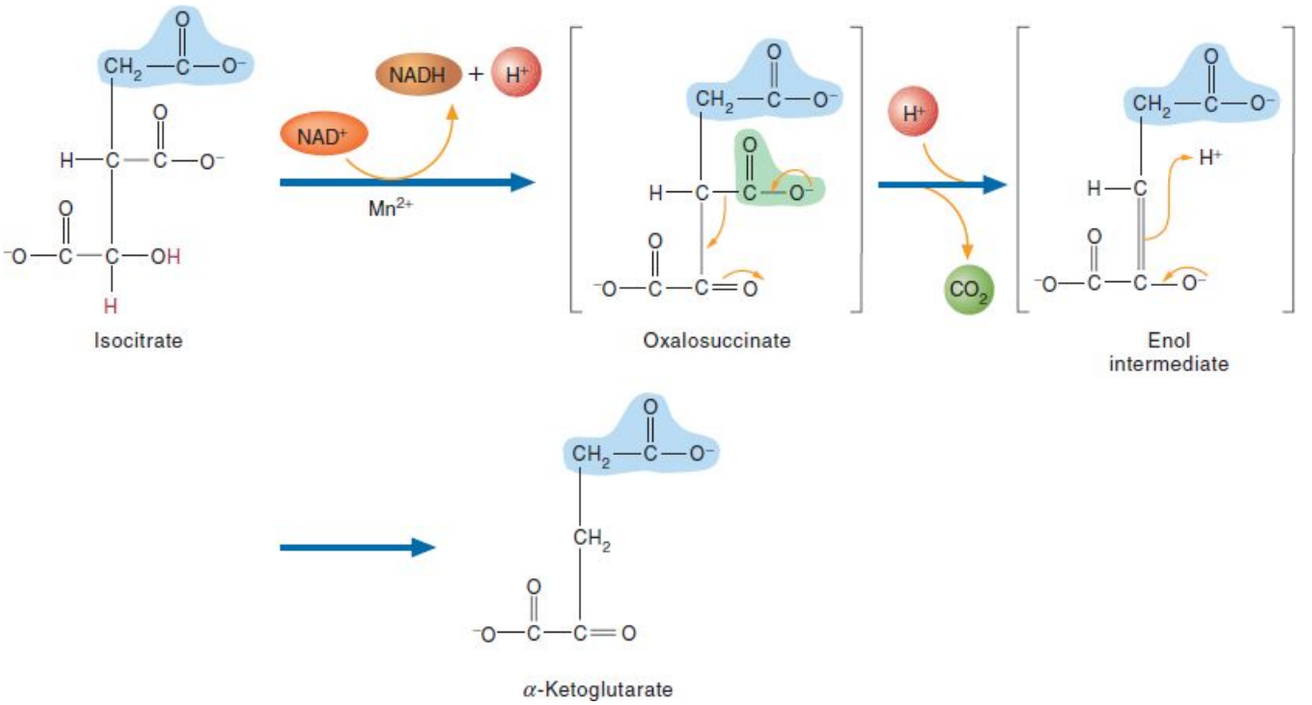


FIGURE 9.13

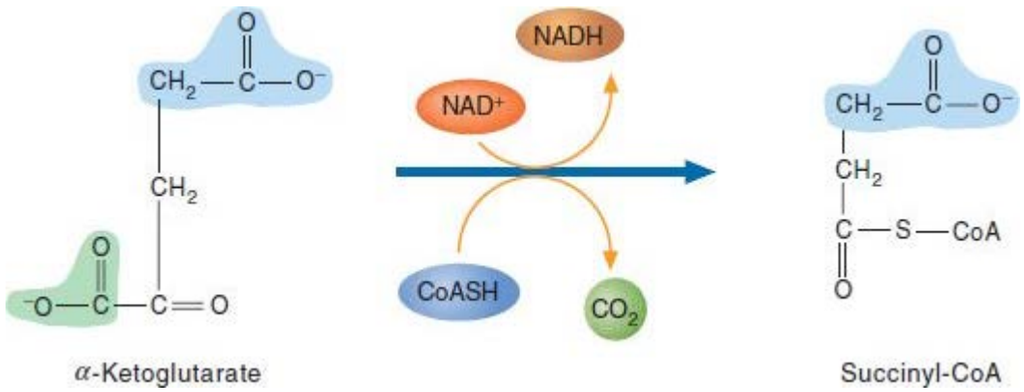
Citrate Synthesis

(1) A side chain carboxylate group (Asp) of the enzyme citrate synthase, acting as a base, removes a proton from the methyl group of acetyl-CoA to form an enol. (2) Simultaneously, a side chain NH group of a histidine residue donates a proton to the carbonyl oxygen, thus generating an enol intermediate. (3) The same histidine side chain then deprotonates the enol to produce an enolate anion that launches a nucleophilic attack on the carbonyl carbon of oxaloacetate. (4) The product, citroyl-CoA, is then hydrolyzed in a nucleophilic acyl substitution reaction when the oxygen of a nearby water molecule (deprotonated by a second histidine side chain) attacks the thioester bond, to yield citrate and CoASH.

3. **Isocitrate is oxidized to form α-ketoglutarate and CO_2.** The oxidative decarboxylation of isocitrate, catalyzed by isocitrate dehydrogenase, proceeds with a $\Delta G^{\circ\prime}$ of –8.4kJ/mol. There are three isoenzyme forms of isocitrate dehydrogenase in mammals. The NAD^+-requiring isozyme (IDH3) is found only within mitochondria. The other isozymes, IDH1 (cytoplasm) and IDH2 (mitochondria), use $NADP^+$ as a cofactor. NADPH is required in biosynthetic processes and quenching of reactive oxygen species (ROS). Note that the NADH produced in the conversion of isocitrate to α-ketoglutarate is the first link between the citric acid cycle and the ETC and oxidative phosphorylation. The reaction begins when the α-carbon (C2) hydroxy group of isocitrate is deprotonated to form the carbonyl group of the intermediate oxalosuccinate. As the carbonyl group forms, a hydride ion is extracted from the α-carbon and then donated to $NAD(P)^+$ to yield NAD(P)H. A manganese (Mn^{2+}) cofactor facilitates the polarization of the newly formed ketone group in oxalosuccinate. The decarboxylation of oxalosuccinate is initiated when a tyrosine side chain group polarizes the C-3 carboxyl group, causing the formation of the enol intermediate (i.e., electrons flow toward the α-carbon to form the double bond between carbons 2 and 3). The enol intermediate rapidly rearranges to form the more stable molecule α-ketoglutarate, an α-keto acid.



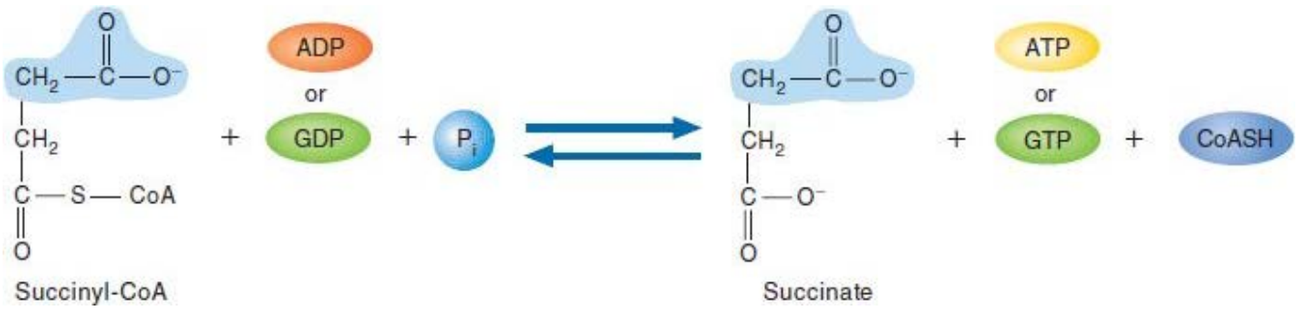
4. ***α*-Ketoglutarate is oxidized to form a second molecule each of NADH and CO_2.** The conversion of *α*-ketoglutarate to succinyl-CoA is catalyzed by the enzyme activities in the *α*-ketoglutarate dehydrogenase complex: *α*-ketoglutarate dehydrogenase, dihydrolipoyl transsuccinylase, and dihydrolipoyl dehydrogenase.



This highly exergonic reaction ($\Delta G^{\circ\prime}$ = −33.5 kJ/mol), an oxidative decarboxylation, is analogous to the conversion of pyruvate to acetyl-CoA catalyzed by the pyruvate dehydrogenase complex. Both reactions have energy-rich thioester molecules as products, that is, acetyl-CoA and succinyl-CoA. Other similarities between the two multienzyme complexes are that the same cofactors (TPP, CoASH, lipoic acid, NAD^+, and FAD) are required, and the same or similar allosteric effectors are inhibitors. *α*-Ketoglutarate dehydrogenase is inhibited by succinyl-CoA, NADH, ATP, and GTP. An important difference between the two complexes is that the control mechanism of the *α*-ketoglutarate dehydrogenase complex does not involve covalent modification.

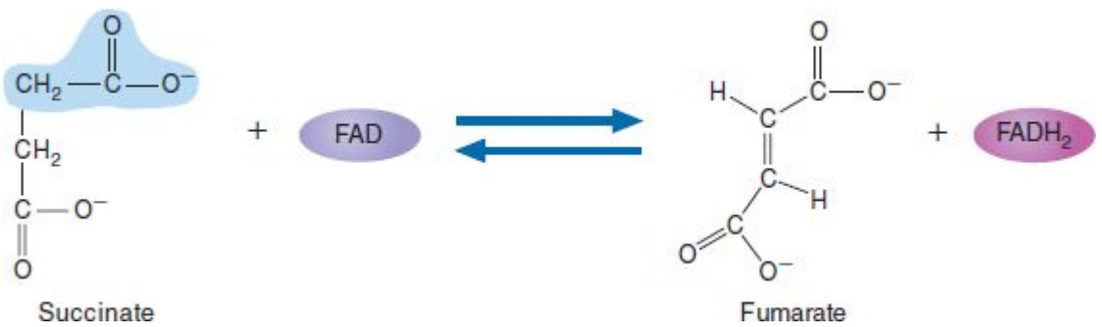
5. **The cleavage of succinyl-CoA is coupled to a substrate-level phosphorylation.** Succinyl-CoA contains a thioester bond with a free energy of hydrolysis value ($\Delta G^{\circ\prime}$) of –36 kJ/mol. The cleavage of the high-energy bond of succinyl-CoA to form succinate is a reversible reaction catalyzed by succinyl-CoA synthetase (also called succinate thiokinase). The reaction begins with the nucleophilic displacement of CoASH to yield succinyl phosphate.

The phosphate is subsequently removed by the enzyme with the formation of a covalent bond between a histidine side chain and the phosphate group. The phosphate group is then transferred to a nucleoside diphosphate. In mammals, the reaction is coupled to the substrate-level phosphorylation of ADP or GDP. Since the substrate level phosphorylation has a $\Delta G°'$ value of +32.2 kJ/mol, the overall reaction is reversible, with a net $\Delta G°'$ value of approximately −3.8 kJ/mol. There are two forms of succinyl-CoA synthetase: one is specific for ATP and the other for GTP. In many tissues, both enzymes are produced, although their relative amounts vary.



The direction of the reaction depends on the relative concentrations of nucleoside diphosphates (ADP and/or GDP) and nucleotide triphosphates (ATP and/or GTP). GTP is used for RNA, DNA, and protein synthesis within mitochondria. The phosphoryl group of GTP can also be donated to ADP in a reversible reaction catalyzed by nucleoside diphosphate kinase.

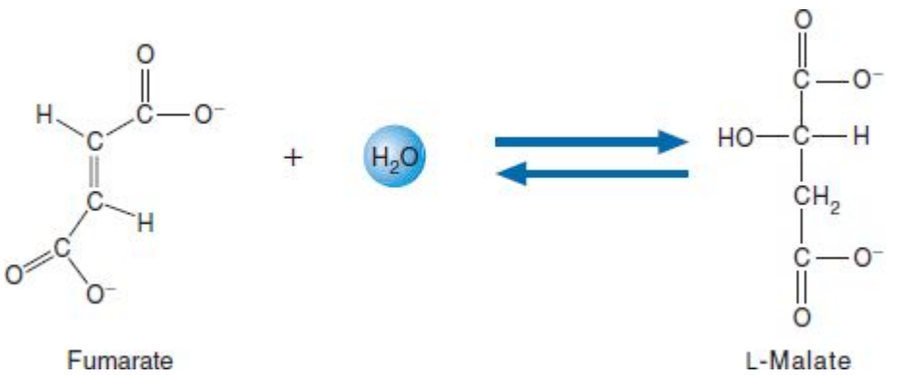
6. **The four-carbon molecule succinate is oxidized to form fumarate and $FADH_2$.** Succinate dehydrogenase catalyzes the reversible oxidation of succinate to form fumarate:



A general base within the enzyme's active site removes a proton from C-2 to form a C-2–C-3 double bond followed by the release of a C-3 hydride to FAD to yield $FADH_2$ and fumarate. Unlike the other citric acid cycle enzymes, succinate dehydrogenase is not found within the mitochondrial matrix. Instead, it is tightly bound to the inner mitochondrial membrane. Succinate dehydrogenase is a flavoprotein that uses FAD to drive the oxidation of succinate to fumarate. (Note that FAD is used instead of NAD^+ because it is a stronger oxidizing agent; in other words, FAD's more positive reduction potential [see Table 9.1] allows the oxidation of a carbon–carbon single bond to form a double bond.) Succinate dehydrogenase is composed of four subunits: (1) ShdA, which contains the succinate binding site and a covalently bound FAD; (2) ShdB, which possesses three iron-sulfur

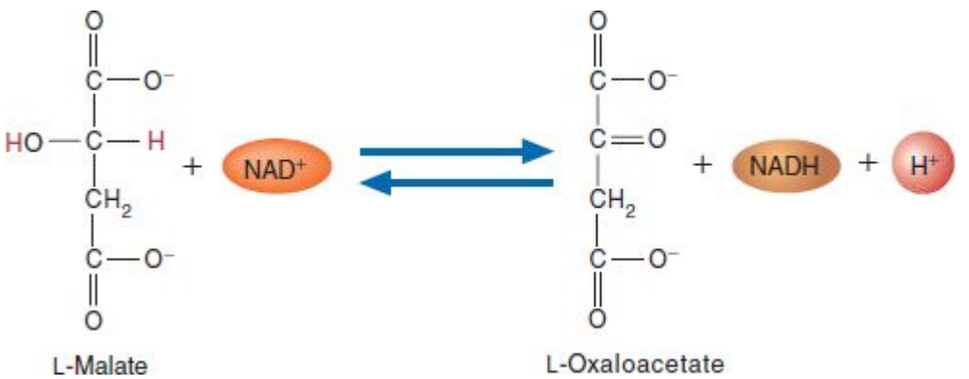
clusters (p. 364) that function as electron carriers between $FADH_2$ and coenzyme Q, a component of the ETC; and (3, 4) subunits ShdC and ShdD, which are hydrophobic molecules that anchor the enzyme complex into the inner membrane. The $\Delta G^{\circ\prime}$ for succinate oxidation is −5.6 kJ/mol. Succinate dehydrogenase is activated by high concentrations of succinate, ADP, and P_i and inhibited by OAA. The enzyme is also inhibited by malonate (see Figure 6.18 on p. 227), a structural analogue of succinate, whose accumulation redirects energy to fatty acid synthesis. Hans Krebs (1900–1981) used this inhibitor in his pioneering work on the citric acid cycle.

7. **Fumarate is hydrated.** Fumarate is converted to L-malate in a reversible stereospecific hydration reaction ($\Delta G^{\circ\prime}$ = −3.8 kJ/mol) catalyzed by fumarase (also referred to as fumarate hydratase):



Fumarase is a lyase that catalyzes a stereospecific hydration. A general base of the enzyme deprotonates a water molecule that then attacks the double bond to form a hydroxyl group at C-2 and a carbanion at C-3. Next, a general acid protonates the carbanion to yield L-malate.

8. **Malate is oxidized to form OAA and a third NADH.** Finally, OAA is regenerated with the oxidation of L-malate by malate dehydrogenase:



A histidine side chain in the active site of the enzyme removes a hydrogen from the hydroxyl group at C-2 of L-malate. Simultaneously, with the formation of the carbonyl group, a hydride ion is transferred to NAD^+ to yield NADH. Malate dehydrogenase uses NAD^+ as the oxidizing agent in a highly endergonic reaction ($\Delta G^{\circ\prime}$ = +29 kJ/mol). The reaction is pulled to completion because of the removal of oxaloacetate in the next round of the cycle. There are two isoenzyme forms of malate dehydrogenase. One is located in the mitochondrial matrix, whereas the other, found in the cytoplasm, is involved in the malate–aspartate shuttle (see Figure 10.21b on p. 382).

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on Hans Krebs and the citric acid cycle.

Fate of Carbon Atoms in the Citric Acid Cycle

In each turn of the citric acid cycle, two carbon atoms enter as the acetyl group of acetyl-CoA, and two molecules of CO_2 are released. A careful review of Figure 9.8 reveals that the two carbon atoms released as CO_2 molecules are not the same two carbons that just entered the cycle. Instead, the released carbon atoms are derived from OAA, which reacted with the incoming acetyl-CoA. The incoming carbon atoms subsequently form one-half of succinate. Because of succinate's symmetric structure, the carbon atoms derived from the incoming acetyl group become evenly distributed in all of the molecules derived from succinate. Consequently, incoming carbon atoms are released as CO_2 only after two or more turns of the cycle.

KEY CONCEPTS

- The citric acid cycle begins with the condensation of a molecule of acetyl-CoA with oxaloacetate to form citrate, which is eventually reconverted to oxaloacetate.
- During this process, two molecules of CO_2, three molecules of NADH, one molecule of $FADH_2$, and one molecule of GTP are produced.

QUESTION 9.5

Trace the labeled carbon in $CH_3{}^{14}C(=O)—SCoA$ through one round of the citric acid cycle. After examining **Figure 9.8**, show why more than two turns of the cycle are required before all the labeled carbon atoms are released as $^{14}CO_2$.

QUESTION 9.6

A mutated IDH1 isoenzyme is found in a high percentage of a type of brain cancer called glioblastoma. Instead of converting isocitrate to α-ketoglutarate, mutated IDH1 converts its substrate to 2-hydroxyglutarate, a circumstance that disrupts the citric acid cycle, among other effects. Review the structures of isocitrate and α-ketoglutarate and determine the structure of 2-hydroxyglutarate.

The Amphibolic Citric Acid Cycle

Amphibolic pathways can function in both anabolic and catabolic processes. The citric acid cycle is obviously catabolic: acetyl groups are oxidized to form CO_2, and energy is conserved in reduced coenzyme molecules. The citric acid cycle is also anabolic, since several citric acid cycle intermediates are precursors in biosynthetic pathways (**Figure 9.14**). For example, OAA is a

gluconeogenesis substrate (Chapter 8) and a precursor in the synthesis of the amino acids aspartate, lysine, threonine, isoleucine, and methionine (Chapter 14). α-Ketoglutarate also plays an important role in amino acid synthesis as a precursor of glutamate, glutamine, proline, and arginine. The synthesis of porphyrins such as heme (**Figure 5.35**) requires succinyl-CoA (Chapter 14). Finally, excess citrate molecules are transported into the cytoplasm, where they are cleaved to form OAA and acetyl-CoA. The latter molecule is used to synthesize fatty acids and steroid molecules such as cholesterol (Chapter 12).

Anabolic processes drain the citric acid cycle of the molecules required to sustain its role in energy generation. Several reactions, referred to as **anaplerotic reactions**, replenish them. One of the most important anaplerotic reactions is catalyzed by pyruvate carboxylase (p. 302). A high concentration of acetyl-CoA, an indicator of an insufficient OAA concentration, activates pyruvate carboxylase. As a result, OAA concentration increases. Other anaplerotic reactions include the synthesis of succinyl-CoA from certain fatty acids (Chapter 12) and the α-keto acids α-ketoglutarate and OAA from the amino acids glutamate and aspartate, respectively, via transamination reactions (Chapter 14).

KEY CONCEPTS

- The citric acid cycle is an amphibolic pathway; that is, it plays a role in both anabolism and catabolism.
- The citric acid cycle intermediates used in anabolic processes are replenished by several anaplerotic reactions.

QUESTION 9.7

Pyruvate carboxylase deficiency, a disease that is usually fatal, is caused when the enzyme that converts pyruvate to OAA is missing or defective. It is characterized by varying degrees of mental retardation and disturbances in several metabolic pathways, especially those involving amino acids and their degradation products. A prominent symptom of this malady is *lactic aciduria* (lactic acid in the urine). After reviewing the function of pyruvate carboxylase, explain why this symptom occurs.

Pyruvate Carboxylase Deficiency

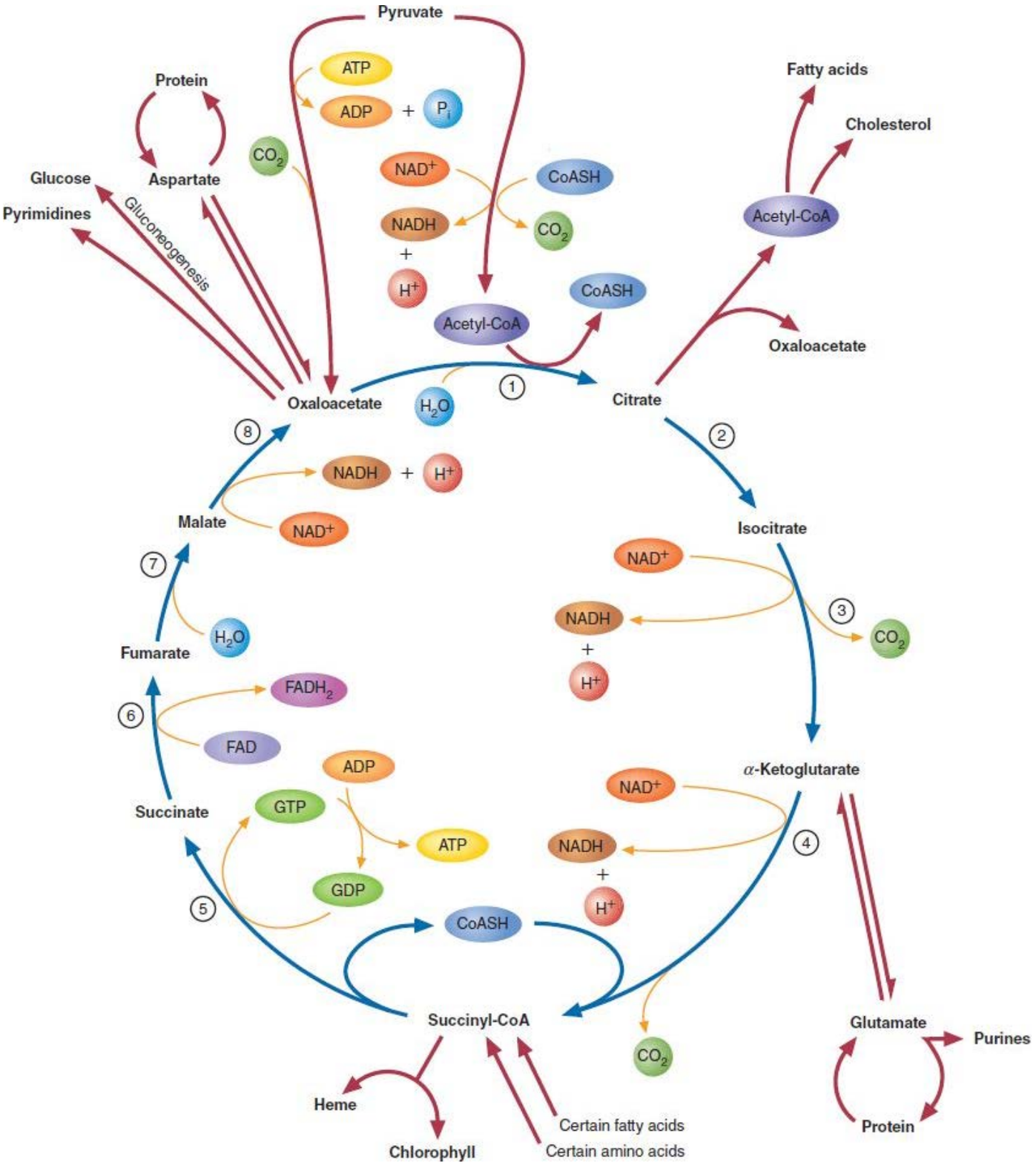


FIGURE 9.14

The Amphibolic Citric Acid Cycle

The citric acid cycle operates in both anabolic processes (e.g., the synthesis of fatty acids, cholesterol, heme, and glucose) and catabolic processes (e.g., amino acid degradation and energy production).

Citric Acid Cycle Regulation

The citric acid cycle is precisely regulated to meet the cell's energy and biosynthetic requirements (**Figure 9.15**). Regulation is achieved via control of three irreversible enzymes within the cycle:

citrate synthase, isocitrate dehydrogenase, and α-ketoglutarate dehydrogenase. These three enzymes operate far from equilibrium (i.e., with highly negative $\Delta G^{\circ\prime}$ values), and they also catalyze reactions that represent important metabolic branch points. Strategies of control include substrate availability, product inhibition, and competitive feedback inhibition. Increased matrix levels of Ca^{2+} also activate all three enzymes.

CITRATE SYNTHASE Citrate synthase, the first enzyme in the cycle, catalyzes the formation of citrate from acetyl-CoA and OAA. The concentrations of acetyl-CoA and OAA are low in mitochondria in relation to the amount of the enzyme. Any increase in substrate availability stimulates citrate synthesis. Citrate synthase is inhibited by its product, citrate. OAA is the product of an endergonic reaction. Therefore, its concentration in mitochondria is quite low relative to malate unless the $NADH/NAD^+$ ratio is low. In many Gram-negative bacteria (*E. coli*, for instance), ATP, NADH, and succinyl-CoA allosterically inhibit citrate synthase.

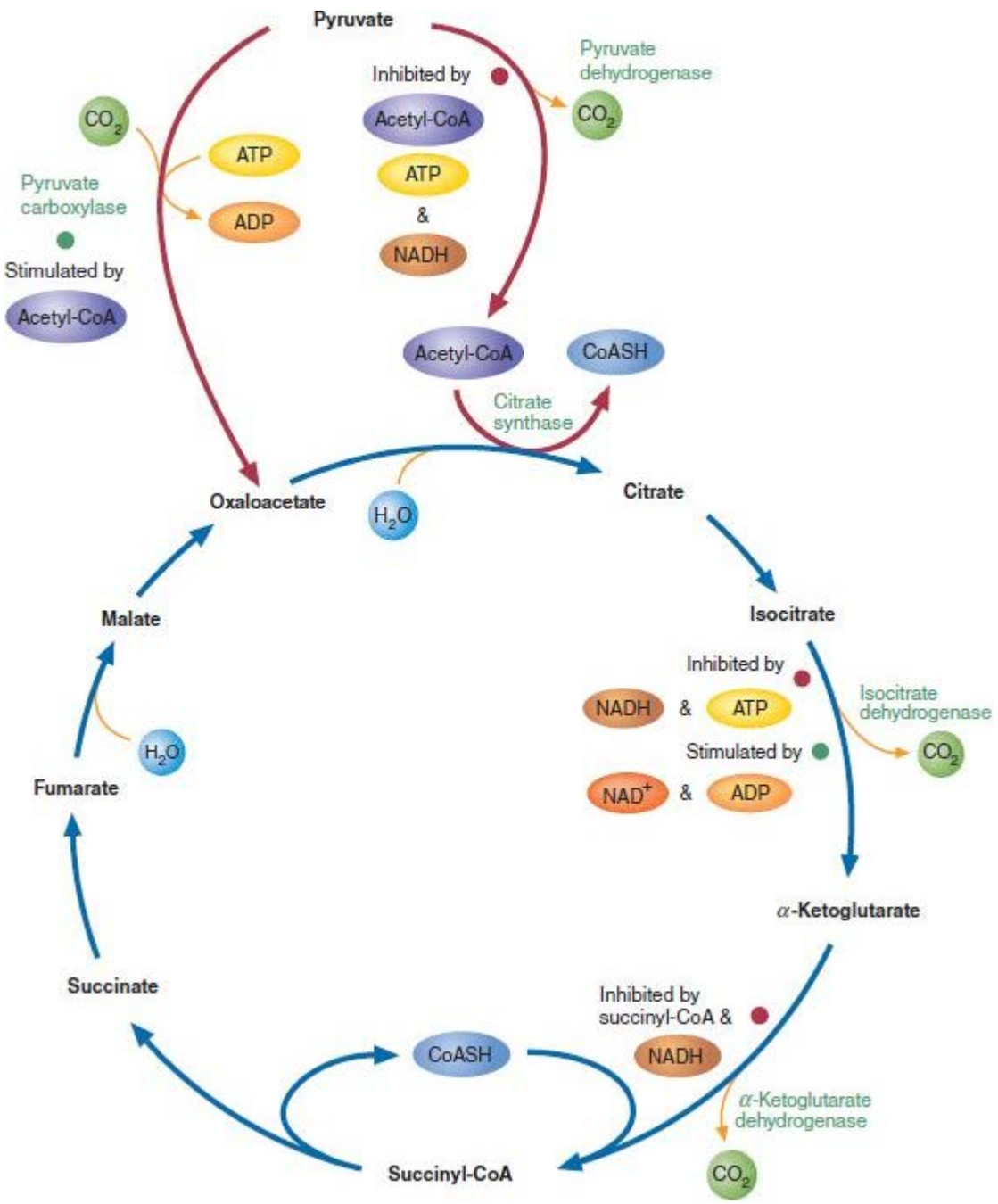


FIGURE 9.15

Control of the Citric Acid Cycle

The major regulatory sites of the cycle are indicated. Activators and inhibitors of regulated enzymes are shown in color: red sphere (inhibition) and green sphere (activation).

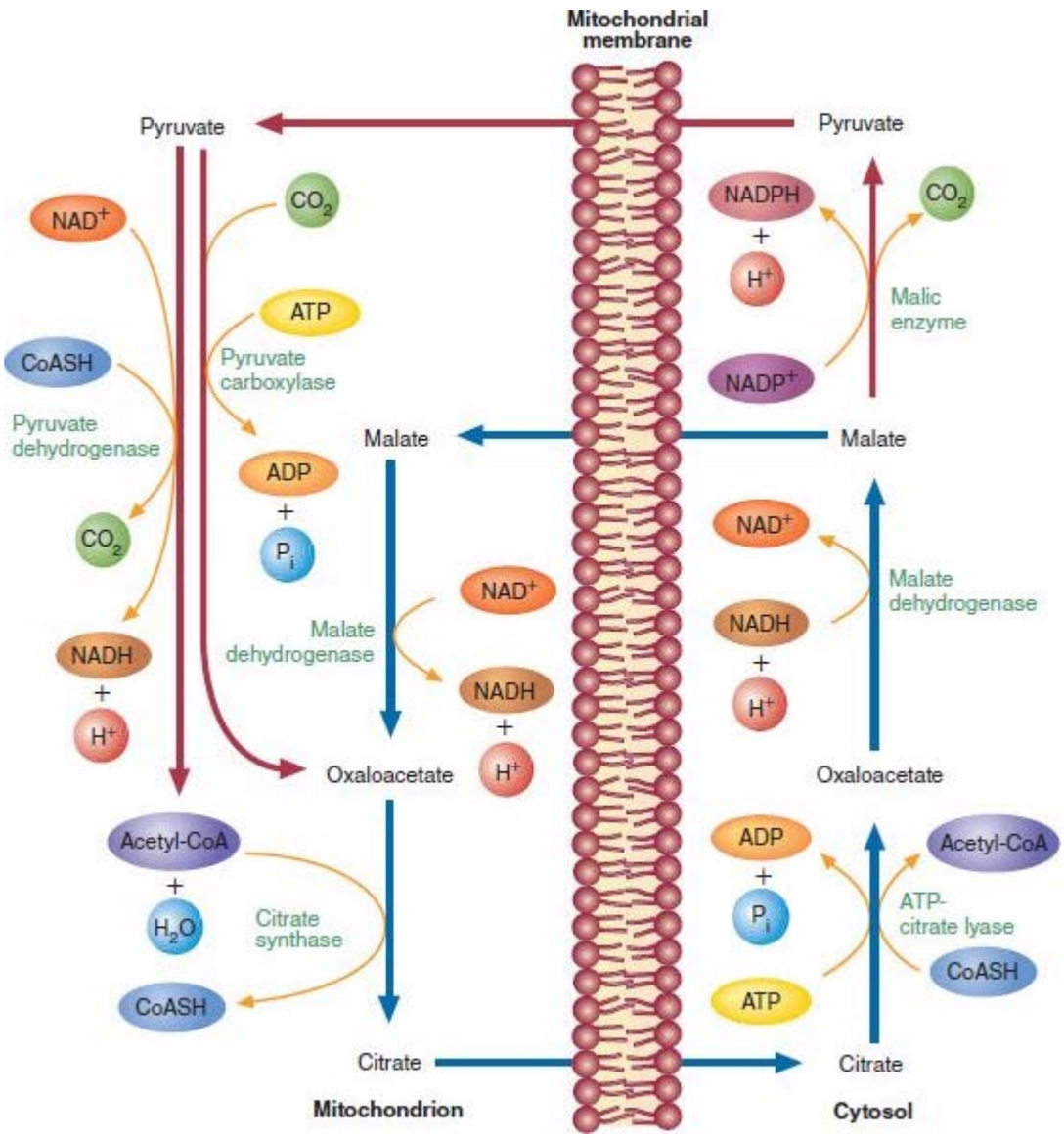


FIGURE 9.16

Citrate Metabolism

When citrate, a citric acid cycle intermediate, moves from the mitochondrial matrix into the cytoplasm, it is cleaved to form acetyl-CoA and oxaloacetate by citrate lyase. The citrate lyase reaction is driven by ATP hydrolysis. Most of the oxaloacetate is reduced to malate by malate dehydrogenase. Acetyl-CoA molecules can be used in biosynthetic pathways such as fatty acid synthesis. Malate may then be oxidized to pyruvate and CO_2 by malic enzyme. The NADPH produced in this reaction is used in cytoplasmic biosynthetic processes, such as fatty acid synthesis. Pyruvate enters the mitochondria, where it may be converted to oxaloacetate or acetyl-CoA. Malate may also reenter the mitochondria, where it is reoxidized to form oxaloacetate.

ISOCITRATE DEHYDROGENASE Isocitrate dehydrogenase's activity is stimulated by relatively

high concentrations of ADP and NAD and inhibited by ATP and NAD(P)H. Isocitrate dehydrogenase is closely regulated because of its important role in citrate metabolism (**Figure 9.16**). As described earlier, the conversion of citrate to isocitrate is reversible. An equilibrium mixture of the two molecules consists largely of citrate. (The reaction is driven forward because isocitrate is rapidly transformed to α-ketoglutarate.) Of the two molecules, only citrate can penetrate the mitochondrial inner membrane. When cellular energy demands are met, excess citrate molecules are transported out of mitochondria and into the cytoplasm. Citrate is then cleaved by ATP-citrate lyase to yield acetyl-CoA and oxaloacetate. The acetyl-CoA formed is used in the synthesis of fatty acids and other lipids. OAA is used in biosynthetic reactions, or it can be converted to malate. Malate either reenters the mitochondrion, where it is reconverted to OAA, or is converted in the cytoplasm to pyruvate by malic enzyme. Pyruvate then reenters the mitochondrion. In addition to being a precursor of acetyl-CoA and OAA in the cytoplasm, citrate also acts directly to regulate several cytoplasmic processes. Citrate is an allosteric activator of the first reaction of fatty acid synthesis. In addition, citrate metabolism provides some of the NADPH used in fatty acid synthesis. Finally, because citrate is an inhibitor of PFK-1, it inhibits glycolysis.

α-KETOGLUTARATE DEHYDROGENASE The activity of α-ketoglutarate dehydrogenase is strictly regulated because of the important role of α-ketoglutarate in several metabolic processes (e.g., amino acid metabolism). When a cell's energy stores are low, α-ketoglutarate dehydrogenase is activated, and α-ketoglutarate is retained within the cycle at the expense of biosynthetic processes. As the cell's supply of NADH rises, the enzyme is inhibited, and α-ketoglutarate molecules become available for biosynthetic reactions. The enzyme is also inhibited by its product succinyl-CoA and activated by AMP, a critical indicator of low energy charge.

Two enzymes outside the citric acid cycle profoundly affect its regulation. The relative activities of PDHC and pyruvate carboxylase determine the degree to which pyruvate is used to generate energy and biosynthetic precursors. For example, if a cell is using a cycle intermediate such as α-ketoglutarate in biosynthesis, the concentration of OAA falls and acetyl-CoA accumulates. Because acetyl-CoA is an activator of pyruvate carboxylase (and an inhibitor of PDHC), more OAA is produced from pyruvate, thus replenishing the cycle.

KEY CONCEPTS

- The citric acid cycle is closely regulated, thus ensuring that the cell's energy and biosynthetic needs are met.
- Allosteric effectors and substrate availability primarily regulate the enzymes citrate synthase, isocitrate dehydrogenase, α-ketoglutarate dehydrogenase, pyruvate dehydrogenase, and pyruvate carboxylase.

CALCIUM REGULATION The signal transduction mechanisms by which cells respond to a variety of stimuli (e.g., hormones, growth factors, and neurotransmitters) often involve transient increases in cytoplasm [Ca^{2+}], followed rapidly by increases in [Ca^{2+}] in the mitochondrial matrix (p. 60). A principal role of Ca^{2+} in the matrix is the stimulation of ATP synthesis by activating the enzymes that regulate the pace of the citric acid cycle. Calcium ions stimulate PDHC activity by activating the dephosphorylating enzyme PDP (p. 342). Both isocitrate dehydrogenase and α-ketoglutarate dehydrogenase are activated directly by Ca^{2+} when the ion binds to a regulatory site on each enzyme. The linkage of the cell's response to a stimulus-driven signal transduction pathway with the uptake of Ca^{2+} into the mitochondrial matrix thus serves to match energy demand with energy production.

The Citric Acid Cycle and Human Disease

Though rarely, several human diseases have been attributed to deficits in citric acid cycle enzymes. Because of the brain's high-energy requirements, the most commonly observed illnesses are severe forms of *encephalopathy* (brain dysfunction characterized by cognitive deficits, tremor, and seizures). For example, encephalopathies have been linked to mutations in the genes that code for α-ketoglutarate dehydrogenase, the A subunit of succinate dehydrogenase, fumarase, and succinyl-CoA synthetase. Several rare cancers are also caused by citric acid cycle enzyme deficits (p. 356). SHB and SHD mutations can cause a pheochromocytoma, an adrenal tumor that secretes excessive amounts of the hormone/neurotransmitter molecules epinephrine and norepinephrine. Symptoms include excessive heart rate and sweating, high blood pressure, and anxiety. A form of renal cell cancer is caused by mutations in fumarase. Several mutated versions of citric acid cycle enzymes contribute to the metabolic disorders in numerous cancers. (Refer to the Biochemistry in Perspective reading on carcinogenesis on p. 356.)

Encephalopathy and Cancers

The Glyoxylate Cycle

Plants and some fungi, algae, protozoans, and bacteria can grow using two-carbon compounds. (Molecules such as ethanol, acetate, and acetyl-CoA, derived from fatty acids, are the most common substrates.) The series of reactions responsible for this capability, referred to as the **glyoxylate cycle**, is a modified version of the citric acid cycle. In plants, the glyoxylate cycle occurs in organelles called glyoxysomes, a type of peroxisome (p. 61) found in germinating seeds. In the absence of photosynthesis, growth in germinating seed is supported by the conversion of oil reserves (triacylglycerol) to carbohydrate. In other eukaryotic organisms and in bacteria, glyoxylate enzymes occur in cytoplasm.

The glyoxylate cycle (**Figure 9.17**) consists of five reactions. The first two reactions (the synthesis of citrate and isocitrate) are familiar because they also occur in the citric acid cycle. However, the formation of citrate from OAA and acetyl-CoA and the isomerization of citrate to form isocitrate are catalyzed by glyoxysome-specific isozymes. The next two reactions are unique to the glyoxylate cycle. Isocitrate is split into two molecules (succinate and glyoxylate) by isocitrate lyase. (This reaction is an aldol cleavage.) Succinate, a four-carbon molecule, is eventually converted to malate by mitochondrial enzymes (**Figure 9.18**). The two-carbon molecule glyoxylate reacts with a second molecule of acetyl-CoA to form malate in a reaction catalyzed by malate synthase. The cycle is completed as malate is converted to OAA by malate dehydrogenase.

The glyoxylate cycle allows for the net synthesis of larger molecules from two-carbon molecules. The decarboxylation reactions of the citric acid cycle, in which two molecules of CO_2 are lost, are bypassed. Using two molecules of acetyl-CoA, the glyoxylate cycle produces one molecule each of succinate and OAA. The succinate product is used in the synthesis of metabolically important molecules such as glucose. The oxaloacetate product is used to sustain the glyoxylate cycle. In organisms, such as animals, that do not possess isocitrate lyase and malate synthase, gluconeogenesis substrates are always molecules with at least three carbon atoms. In these organisms, there is no net synthesis of glucose from fatty acids.

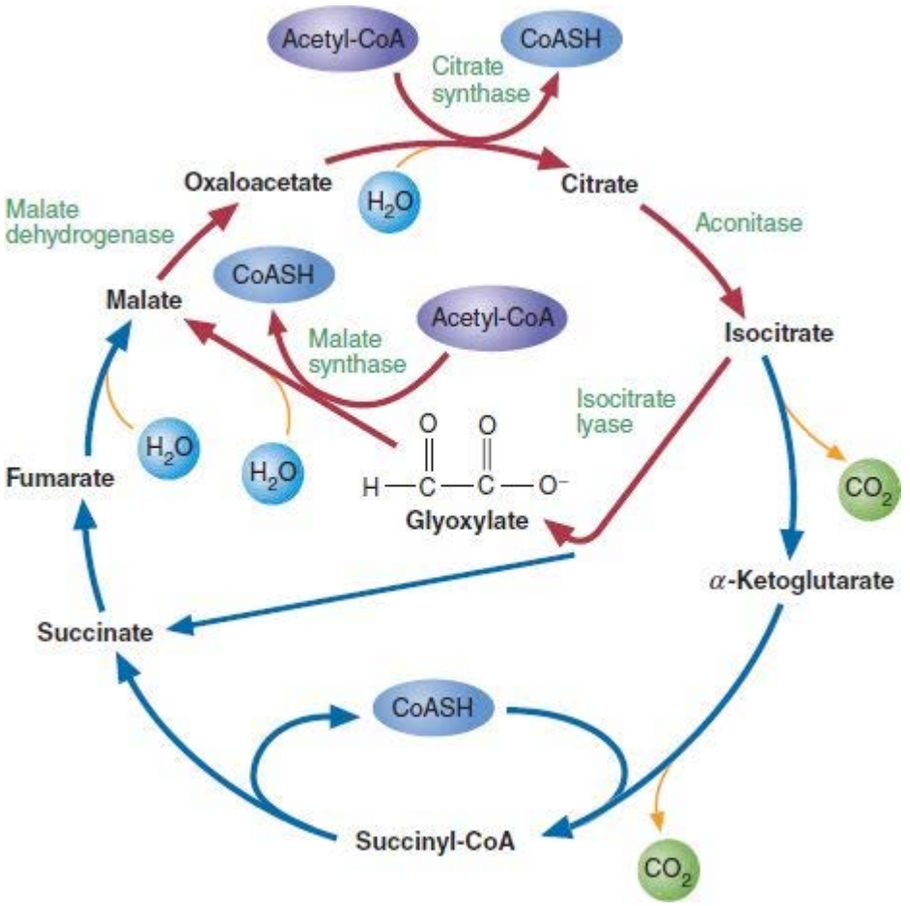


FIGURE 9.17

The Glyoxylate Cycle

Using some of the enzymes of the citric acid cycle, the glyoxylate cycle converts two molecules of acetyl-CoA to one molecule of oxaloacetate. Both decarboxylation reactions of the citric acid cycle are bypassed.

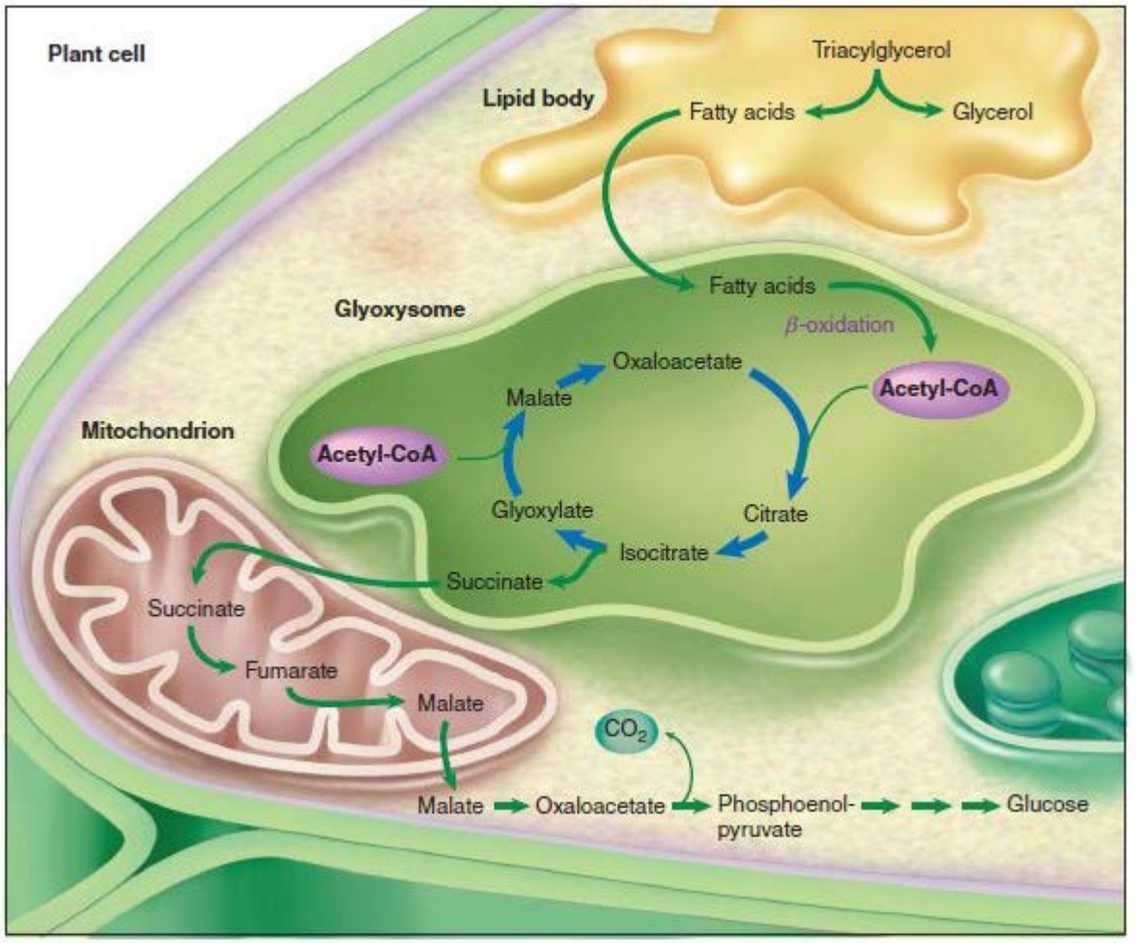


FIGURE 9.18

Role of the Glyoxylate Cycle in Gluconeogenesis

The acetyl-CoA used in the glyoxylate cycle is derived from the breakdown of fatty acids (*β*-oxidation, see Chapter 12). In organisms with the appropriate enzymes, such as some bacteria, glucose can be produced from two-carbon compounds such as ethanol and acetate. In plants, the reactions that convert fatty acids into glucose are localized in lipid bodies, glyoxysomes, mitochondria, and the cytoplasm.

KEY CONCEPTS

- Organisms in which the glyoxylate cycle occurs can use two-carbon molecules to sustain growth.
- In plants, the glyoxylate cycle is found in organelles called glyoxysomes.

Biochemistry IN PERSPECTIVE

Carcinogenesis: The Warburg Effect and Metabolic Reprogramming

What is aerobic glycolysis, and how does it impact cancerous cells? Cancer is a

group of diseases in which there is uncontrolled proliferation of abnormal cells. In the past, cancer research focused on investigations of mutated genes. More recently, cancer has been increasingly recognized as an intricate and heterogeneous set of metabolic disorders. The first abnormal metabolic cancer discovery is credited to Otto Warburg (1883–1970), who reported in 1924 that many tumor cells generate energy principally by glycolysis instead of aerobic respiration, even if O_2 is plentiful. He believed that this exception to the Pasteur effect (p. 301) in cancerous cells, now referred to as the *Warburg effect*, is caused by mitochondrial damage. Although there is mitochondrial damage in some forms of cancer, emerging research has revealed that "aerobic glycolysis" (the conversion of glucose to lactate) is more closely associated with altered cell signaling pathways that cause a reprogramming of metabolic processes. Such alterations give rise to selective advantages to cancerous cells in energy generation and biomolecule synthesis.

Cancerous cells require a large amount of energy and a large supply of metabolic precursor molecules to sustain rapid growth. Although glycolytic ATP synthesis is inefficient, the process of carcinogenesis involves metabolic changes that include markedly increased glucose uptake and glycolytic activity in "transformed" cells. Increased glycolytic flux (10 to 100 times faster than complete glucose oxidation) provides precursors for synthesis of the molecules required for cell division: amino acids, lipids, and nucleic acids. For example, glucose-6-phosphate is a precursor of ribose-5-phosphate; dihydroxyacetone phosphate is converted to glycerol, used in triacylglycerol synthesis; and glycerate-3-phosphate is a precursor of serine, glycine, and cysteine. Cancer cells also use the pentose phosphate pathway, among other processes, to synthesize large quantities of reducing power in the form of NADPH. The evasions of normal regulatory mechanisms that allow cells to rewire their metabolism involve mutations in a diverse group of genes. Prominent examples include c-myc, HIF-1, PKB, and p53. Although many cancers have normal oxidative phosphorylation, increased reliance on aerobic glycolysis in some cells can also be caused by tumor cell mitochondria with mutations in certain citric acid cycle enzymes.

MYC The *myc* gene codes for a transcription factor that regulates about 15% of all genes, including many that are involved in cell division and apoptosis (p. 59). Among myc's target genes are those that code for most of the glycolytic enzymes and several glucose transporters (GLUTs, p. 619). Tumors with mutated myc, called c-myc (representing approximately 70% of all human cancers), have increased lactic acid production and diversion of glycolytic intermediates into pathways that synthesize nucleotides, amino acids, and lipids. C-myc also promotes glutaminolysis, an alternative energy-generating pathway in which glutamine is transported into mitochondria and converted into the citric acid cycle intermediate α-ketoglutarate, which is subsequently converted to citrate, the source of acetyl CoA for fatty acid synthesis (p. 461). Cells also use IDH1 to synthesize NADPH. The glutamine transporter, glutaminase (the enzyme that converts glutamine to glutamate), LDHA (the gene that codes for the M subunit of lactate dehydrogenase), and the lactate plasma membrane transporter are a few of the molecules upregulated by c-myc.

HIF-1 HIF-1 (hypoxia inducible factor 1) is a transcription factor that responds to decreases in cellular oxygen levels by inducing the transcription of genes that promote survival under low-oxygen conditions. The most prominent examples are genes for glycolytic enzymes, glucose transporters, and PDK1, the enzyme that inactivates pyruvate dehydrogenase (p. 342), thereby diverting pyruvate away from mitochondrial oxidation. When oxygen is available, HIF-1 is inactivated by oxygen- and α-ketoglutarate-dependent prolyl hydroxylation. HIF-1 is stabilized in some tumors by mutations in succinate dehydrogenase and fumarase. Succinate or fumarate accumulations result in leakage of these molecules into cytoplasm, where they act as competitive

inhibitors of prolyl hydroxylase. In addition, mutated versions of IDH1 or IDH2 can result in the synthesis of 2-hydroxyglutarate, also an inhibitor of prolyl hydroxylase.

PKB Protein kinase B (PKB) (p. 612), also referred to as Akt, is a serine/threonine protein kinase that regulates glucose metabolism, among several other processes. Activated PKB enhances glucose uptake and glycolysis by inducing increased glucose transporter insertion into the plasma membrane, hexokinase activation, and PFK-2-dependent activation of PFK-1 (p. 298). PKB also promotes the diversion of glucose carbon into the biosynthetic pathways that produce fatty acids and cholesterol and other lipids required for rapid growth. PKB facilitates transport of pyruvate into mitochondria and its rapid conversion into acetyl-CoA. The subsequent increase in citrate results in its export into the cytoplasm where acetyl-CoA, the substrate for lipid synthesis, is regenerated by the PKB-activated enzyme ATP-citrate lyase. Mutated versions of PKB that cause it to be hyperactivated cause some malignancies, while those that decrease its activity cause others.

P53 With more than 100 target genes, p53 is one of the most important tumor suppressors. In addition to its well-known roles in DNA repair, cell cycle regulation, and apoptosis induction, p53 is now known as an important regulator of energy metabolism. P53 coordinates the regulation of glycolysis and oxidative phosphorylation. It does so by maintaining oxidative phosphorylation and down-regulating glycolysis (via decreases in glucose transporters and glycolytic enzymes). P53-induced regulation of energy metabolism is exerted through inducing the synthesis and/or activation of enzymes such as PKB, mTOR (p. 612), and TIGAR (TP53-induced glycolysis and apoptosis regulator), which decreases glycolytic flux by lowering the intracellular concentrations of fructose 2,6-bisphosphate (the PFK-1 allosteric activator). Aerobic glycolysis is one of many consequences of p53 inactivation.

SUMMARY Aerobic glycolysis is a process in tumor cells in which rapid glycolysis-generated ATP synthesis occurs even when O_2 is present. Loss of glycolysis regulation is one facet of carcinogenesis, the set of mechanisms whereby normal cells are gradually transformed into cancerous cells that are no longer responsive to the body's regulatory signals.

Chapter Summary

1. Aerobic organisms have an enormous advantage over anaerobic organisms, that is, a greater capacity to obtain energy from organic food molecules. To use oxygen to generate energy requires the following biochemical pathways: the citric acid cycle, the electron transport pathway, and oxidative phosphorylation.
2. Most reactions that capture or release energy are redox reactions. In these reactions, electrons are transferred between an electron donor (reducing agent) and an electron acceptor (oxidizing agent). In most biochemical reactions, hydride ions are transferred to NAD^+ or $NADP^+$ or hydrogen atoms are transferred to FAD or FMN. A substance's tendency to gain electron(s) is called its reduction potential. Electrons flow spontaneously from a substance with a less positive (more negative) reduction potential to a substance with a more positive (less negative) reduction potential. In favorable redox reactions $\Delta E^{\circ\prime}$ is positive and $\Delta G^{\circ\prime}$ is negative.
3. The citric acid cycle is a series of biochemical reactions that eventually completely oxidize organic substrates, such as glucose and fatty acids, to form CO_2, H_2O, and the reduced coenzymes NADH and

$FADH_2$. Pyruvate, the product of the glycolytic pathway, is converted to acetyl-CoA, the citric acid cycle substrate.

4. In addition to its role in energy generation, the citric acid cycle plays an important role in several biosynthetic processes, such as gluconeogenesis, amino acid synthesis, and porphyrin synthesis.
5. The glyoxylate cycle, found in plants and some fungi, algae, protozoans, and bacteria, is a modified version of the citric acid cycle in which two-carbon molecules, such as acetate, are converted to precursors of glucose.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on the citric acid cycle to help you prepare for exams.

Chapter 9 Review Quiz

Suggested Readings

Anderson NM, et al. 2017. The emerging role and targetability of the TCA cycle in cancer metabolism. Protein Cell doi:10.1007/s13238-017-0451-1.

Bilgen T. 2004. Metabolic evolution and the origin of life. In: Storey KB, editor. Functional metabolism: regulation and adaptation. Hoboken (NJ): Wiley-Liss. p. 557–82.

Colombini M. 2012. VDAC structure, selectivity and dynamics. Biochim Biophys Acta 1818(6):1457–65.

Kornberg H. 2000. Krebs and his trinity of cycles. Nat Rev Mol Cell Biol 1(3):225–7.

Krebs HA. 1970. The history of the tricarboxylic cycle. Perspect Biol Med 14:154–70.

Lane N. 2002. Oxygen: the molecule that made the world. Oxford (UK): Oxford University Press.

Lane N. 2015. Power, sex and suicide and the meaning of life. Oxford (UK): Oxford University Press.

Mailloux RJ. 2015. Still at the center of it all: novel functions of the oxidative Krebs cycle. Bioenergetics 4:122. doi:10.4172/2167-7662.1000122.

Potter M., et al. 2016. The Warburg effect: 80 years on. Biochem Soc Trans 44:1499–505.

Schmitt DL, An S. 2017. Spatial organization of metabolic enzyme complexes in cells. Biochem 56:3184–96.

Key Words

aerotolerant anaerobe, *331*

amphibolic pathway, *349*

anaplerotic, *349*

coenzyme A, *339*

conjugate redox pair, *333*

facultative anaerobe, *331*

flavin adenine dinucleotide (FAD), *335*

flavin mononucleotide (FMN), *335*
flavoprotein, *335*
glyoxylate cycle, *354*
lipoic acid, *342*
nicotinamide adenine dinucleotide (NAD), *335*
nicotinamide adenine dinucleotide phosphate (NADP), *335*
obligate aerobe, *331*
obligate anaerobe, *331*
reduction potential, *333*
standard reduction potential, *333*
thiamine pyrophosphate, *342*

Review Questions

SECTION 9.1

Comprehension Questions

1. Define the following terms:
    a. obligate anaerobe
    b. aerotolerant anaerobe
    c. facultative anaerobe
    d. obligate aerobe
    e. reactive oxygen species
2. Define the following terms:
    a. electron transport chain
    b. oxidation-reduction reactions
    c. conjugate redox pair
    d. reduction potential
    e. standard reduction potential
3. Define the following terms:
    a. aerobic metabolism
    b. aerobic respiration
    c. $\Delta G^{\circ\prime}$
    d. $\Delta E^{\circ\prime}$
    e. F
4. Define the following terms:
    a. FMN
    b. FAD
    c. NAD^+
    d. ETC
    e. salt bridge

Fill in the Blanks

5. The tendency for a substance to gain electrons is called the _____ potential.
6. _____ in the outer mitochondrial membrane allows the diffusion of small hydrophilic molecules.
7. Together an electron donor and an electron acceptor in a reaction is called a _____ .
8. A significant portion of the energy captured as electrons pass down the electron transport chain is used to synthesize _____.

Short-Answer Questions

9. Describe in general terms how the appearance of molecular oxygen in Earth's atmosphere about 3 billion years ago affected the history of living organisms.
10. List the biochemical processes required to obtain energy from glucose using O_2 as an electron acceptor.
11. Calculate the free energy changes that occur in the following reactions:

    a. $\frac{1}{2}\ O_2 + NADH + H^+ \rightarrow H_20 + NAD^+$

    b. $S + NADH + H^+ \rightarrow H_2S + NAD^+$

12. Referring to Question 11, calculate the difference in energy produced by the oxidation of NADH by O_2 and sulfur.

Critical-Thinking Questions

13. Determine the standard free energy ($\Delta G°'$) for the following reactions:

    a. $NADH + H^+ + \frac{1}{2}\ O_2 \rightarrow NAD^+ + H_2O$

    b. Cytochrome c (Fe^{2+}) + $\frac{1}{2}\ O_2$ → cytochrome c (Fe^{3+}) + H_2O

14. Using the data in Table 9.1 and the equation $\Delta G°' = -nF\Delta E°'$, calculate the free energy ($\Delta G°'$) produced by the reduction of sulfur to hydrogen sulfide and oxygen to water by NADH. How much more free energy is produced by the reduction of oxygen compared to that of sulfur?
15. Calculate the standard free energy change ($\Delta G°'$) for the following reaction:

    $\frac{1}{2}\ O_2 + FADH_2 \rightarrow H_2O + FAD$

16. The value of the standard reduction potential ($E°'$) for the redox half-reaction

    $Acetaldehyde + 2H^+ + 2e^- \rightarrow Ethanol$

    is −0.20 V. Calculate the free energy ($\Delta G°'$) for the following reaction:

    $Ethanol + NAD^+ \rightarrow Acetaldehyde + NADH + H^+$

SECTION 9.2

Comprehension Questions

17. Define the following terms:

    a. coenzyme A

b. TPP
c. lipoic acid
d. PDHC
e. PDK1

18. Define the following terms:

a. HETPP
b. nucleoside diphosphate kinase
c. amphibolic pathway
d. anaplerotic reaction
e. citric acid cycle

19. Define the following terms:

a. vitamin B_1
b. PDP
c. OAA
d. glyoxylate cycle
e. electron acceptor

Fill in the Blanks

20. _____ pathways can function in both anabolic and catabolic processes.
21. In eukaryotic cells, the citric acid cycle occurs in the _____.
22. The most pivotal event in the history of life on Earth, allowing for the emergence of eukaryotic cells, was the _____.
23. _____ produced in the citric acid cycle is an important carbon source for gluconeogenesis.
24. _____ reactions replenish the intermediate molecules of the citric acid cycle.
25. The acetyl group of acetyl-CoA is derived from the breakdown of lipids, carbohydrates, and certain _____.
26. Pyruvate is converted to acetyl-CoA by the enzymes of the _____ enzyme complex.
27. _____, also referred to as Akt, is a serine/threonine kinase that regulates glucose metabolism.

Short-Answer Questions

28. How do calcium ions regulate the citric acid cycle?
29. How do germinating seeds convert their triacylglycerol reserves to the glucose molecules required in the synthesis of complex carbohydrate such as cellulose?
30. How does pyruvate carboxylase deficiency result in lactic aciduria, an illness in which lactate appears in the urine?
31. Give examples of biochemical pathways that utilize citric acid cycle intermediates as precursor molecules.
32. The citric acid cycle operates only when O_2 is present, yet O_2 is not a substrate for the cycle. Explain.
33. Describe in detail the structure of the pyruvate dehydrogenase complex and the role of each enzyme, cofactor, and coenzyme.
34. If a small amount of [1–^{14}C] glucose is added to an aerobic yeast culture, where will the ^{14}C

label initially appear in citrate molecules?

35. Provide examples of molecules derived from citric acid intermediates.
36. Write the net equation for the citric acid cycle.
37. Discuss the mechanisms of control of the irreversible steps in the citric acid cycle.
38. Write balanced equations for each reaction in the citric acid cycle.
39. What coenzymes are required in each of the reactions in Question 38?
40. Each of the following molecules has a role in carcinogenic aerobic glycolysis: c-myc, HIF-1, PKB, and p53. Describe the function of each molecule.
41. Explain why animals cannot produce glucose from two carbon molecules, such as acetate or ethanol.
42. What steps of the citric acid cycle are regulated? How and why are they regulated?
43. The redox half reaction

    $$\text{Oxaloacetate} + 2H^+ + 2e^- \rightarrow \text{Malate}$$

    has a standard reduction potential value of 0.166 V. Calculate the free energy ($\Delta G°'$) for the following reaction:

    $$\text{Malate} + NAD^+ \rightarrow \text{Oxaloacetate} + NADH + H^+$$

44. Using the data in Table 9.1, calculate the free energy ($\Delta G°'$) for the synthesis of NADH in the conversion of isocitrate to α-ketoglutarate in the citric acid cycle.
45. Malonate (p. 227) poisons the citric acid cycle because it inhibits succinate dehydrogenase. After reviewing its structure, describe how the inhibitory effect of malonate can be overcome.

Critical-Thinking Questions

46. One consequence of ethanol addiction is fatty liver disease, an illness in which liver cells accumulate large amounts of triacylglycerols, the esters derived from glycerol and fatty acids. Ethanol is oxidized in the cytoplasm of liver cells by alcohol dehydrogenase and aldehyde dehydrogenase to yield acetate and 2 NADH. Acetate is then transported into the mitochondrion, where it is converted to acetyl-CoA and metabolized in the citric acid cycle. When alcohol is consumed in excessive quantities, the resulting high levels of NADH cause metabolic abnormalities, one of which is high levels of fatty acid synthesis. Fatty acid synthesis, also a cytoplasmic process, uses acetyl-CoA as a substrate and NADPH as a reducing agent. Speculate about how a high level of cytoplasmic NADH provides a source of NADPH for fatty acid synthesis.
47. Despite the absence of the glyoxylate cycle in mammals, when ^{14}C-labeled acetate is fed to lab animals, small amounts of the radioactive label later appear in glycogen stores. Explain.
48. One effect of chronic alcohol abuse is thiamine deficiency, caused by impaired absorption of the vitamin through the intestinal wall and diminished storage in a damaged liver. When thiamine levels are inadequate, cellular energy generation is diminished. List three enzymes involved in cellular metabolism that require thiamine. Describe the possible metabolic consequences of inadequate thiamine levels.
49. Pyruvate dehydrogenase deficiency is a fatal disease usually diagnosed in children. Symptoms include severe neurological damage. Elevated blood levels of lactate, pyruvate, and alanine are also seen. Explain how the deficiency of pyruvate dehydrogenase causes these elevated values.
50. The plant toxin fluoroacetate ($F\text{-}CH_2COO^-$) is easily converted to fluorocitrate when an animal ingests the plant. The enzyme aconitase has a high affinity for the 2-fluorocitrate molecule. Consider the purpose of this enzyme and speculate as to why the reaction to produce isocitrate

does not occur as planned. Why is fluoroacetate considered a poison? [*Hint*: Fluorine is more electronegative than oxygen.]

51. Systemic fungal infections (*Candida* and *Cryptococcus*) and tuberculosis (*Mycobacterium tuberculosis*) are on the rise. These organisms have high virulence in part because they fare better following macrophage phagocytosis than other microorganisms. Phagocytosis activates the glyoxylate cycle in fungi and mycobacteria, allowing them to use two-carbon substrates to sustain growth in an otherwise nutrient-poor environment of the phagolysosome. Suggest some likely molecules that could be processed through the glyoxylate cycle in this circumstance
52. What is the significance of substrate-level phosphorylation reactions? Which of the reactions in the citric acid cycle involve a substrate-level phosphorylation? Name another example from a biochemical pathway with which you are familiar.
53. You have just consumed a piece of fruit. Trace the carbon atoms in the glucose in the fruit through the biochemical pathways between their uptake into tissue cells and their conversion to CO_2.
54. Fatty acid degradation stimulates the citric acid cycle through the activation of pyruvate carboxylase by acetyl-CoA. Why would the activation of pyruvate carboxylase increase energy generation from fatty acids?
55. Dichloroacetate, which inhibits the enzyme pyruvate dehydrogenase kinase, has been used with limited results to treat lactic acidosis. The phosphorylation of the α-subunit of the pyruvate dehydrogenase component of the PDHC by pyruvate dehydrogenase kinase causes complete loss of enzymatic activity. Describe the hypothesis behind the clinical use of dichloroacetate.
56. Shock (failure of the circulatory system) is an abnormal condition in which blood flow is inadequate. The most common causes of shock are massive blood loss and obstruction of blood flow. As a result of shock, there is a failure to provide cells with oxygen and nutrients. In this circumstance, cells swell and lysosomal membranes rupture, among other effects. Describe how energy is generated during shock and why cell structure becomes destabilized.
57. Lactic acidosis occurs as a result of shock. Explain why low oxygen levels promote lactate production.
58. The large amount of energy used during aerobic exercise (e.g., running) requires large amounts of oxaloacetate. Explain why acetyl-CoA cannot be used to produce oxaloacetate in this circumstance. What is the likely source of oxaloacetate molecules during aerobic activity?

MCAT Study Questions

59. In which of the following reactions is the reactant oxidized?

    a. ATP → AMP
    b. NAD^+ → NADH
    c. NADH → NAD^+
    d. acetaldehyde → ethanol

60. Which of the following statements is true when glucose is oxidized in the presence of O_2?

    a. All ATP synthesis requires the presence of O_2.
    b. O_2 is the final electron acceptor.
    c. No water molecules are formed.
    d. The electron transport chain is responsible for all ATP synthesis.

61. Which of the following is true concerning the function of lipoic acid in pyruvate metabolism?

a. transfer of acetyl groups
b. pyruvate decarboxylation
c. conversion of pyruvate to lactate
d. oxidation of $FADH_2$ to yield NADH

62. Which of the following is not a product of the citric acid cycle?

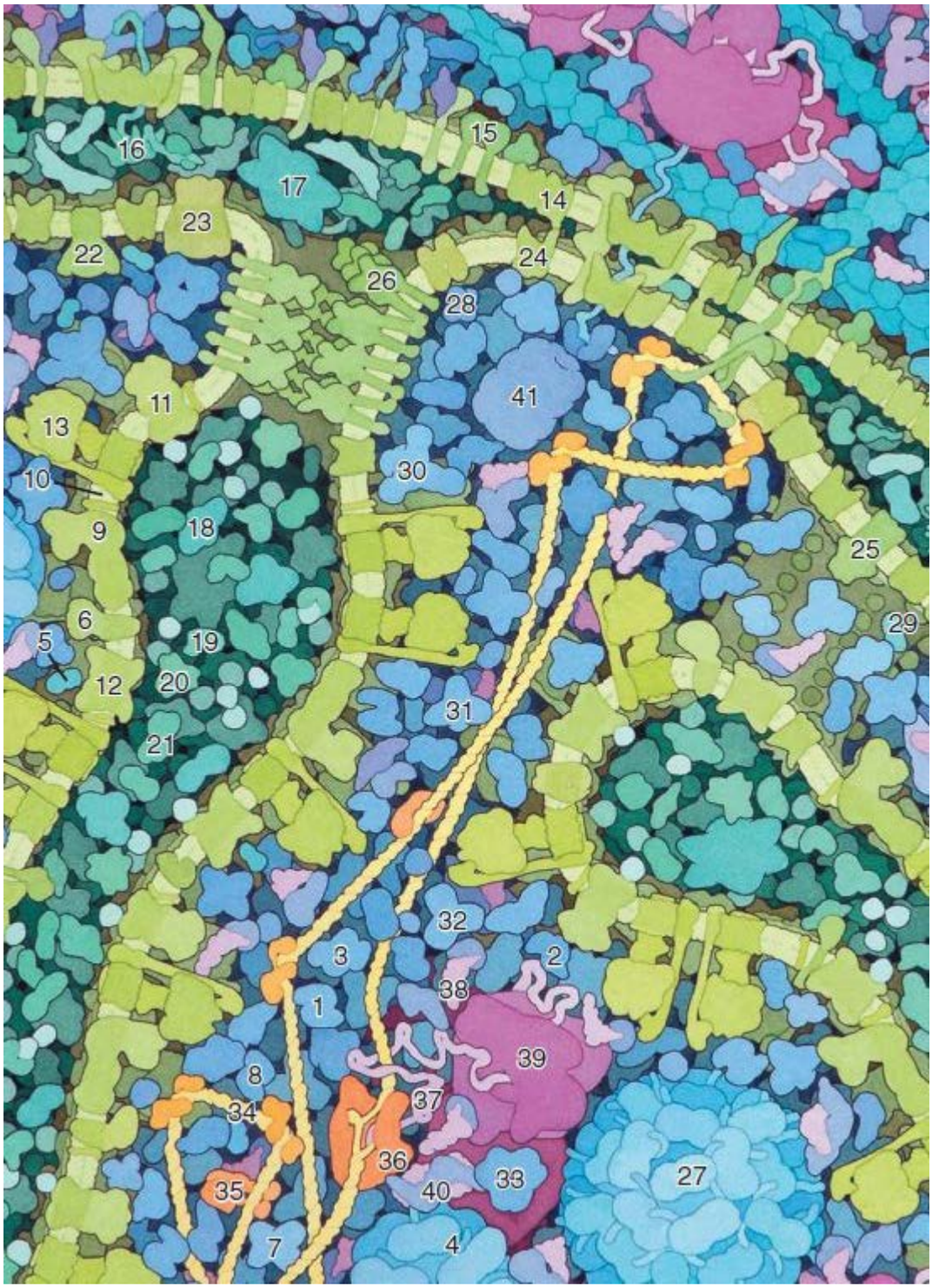
a. $FADH_2$
b. NADH
c. AMP
d. CO_2

63. The coenzymes required for the activity of the pyruvate dehydrogenase complex are

a. biotin, thiamine pyrophosphate, CoASH, and lipoic acid
b. lipoic acid, NAD^+, thiamine pyrophosphate and FAD
c. NAD^+, lipoic acid, tetrahydrofolate, CoASH
d. biotin, tetrahydrofolate, vitamin B6

CHAPTER 9

Aerobic Metabolism I: The Citric Acid Cycle



Cross Section of a Mitochondrion This illustration, based on electron microscopy and X-ray crystallography, shows a cross section of a mitochondrion in magnified molecular detail (1×10^6). Each component is numbered. The citric acid cycle enzymes are (1) citrate synthase, (2) aconitase, (3) isocitrate dehydrogenase, (4) α-ketoglutarate dehydrogenase, (5) succinyl CoA synthetase, (6) succinate dehydrogenase, (7) fumarase, and (8) malate dehydrogenase. The inner membrane electron transport chain components are (9) NADH dehydrogenase, (6) succinate dehydrogenase, (10) coenzyme Q, (11) cytochrome bc1 reductase, (12) cytochrome oxidase, and (13) ATP synthase. Note that for clarity the citric acid cycle enzymes and the electron transport system molecules (Chapter 10) are shown as separate molecules. They are believed to function as metabolons. Other mitochondrion components mentioned elsewhere are (14) VDAC (p. 380), (15) monoamine oxidase, (16) TIM (p. 783), (17) creatine kinase (p. 378), (18) glycerol-3-phosphate dehydrogenase (p. 380), (19) adenylate kinase (p. 154), (20) nucleoside diphosphate kinase (p. 557), (21) sulfate oxidase (p. 587), (22) Mg^{2+} transporter, (23) calcium channel, (24) ADP/ATP translocator (p. 378), (25) potassium channel, (26) dynamin-like protein (regulates mitochondrial fusion and cristae structure), (27) pyruvate dehydrogenase (p.

342), (28) Mn-superoxide dismutase (p. 389), (29) Cu-Zn superoxide dismutase (p. 389), (30) pyruvate carboxylase (p. 302), (31) acyl CoA dehydrogenase (p.453), (32) ornithine transcarbamoylase (p. 578), (33) glutamate dehydrogenase (p. 536), (34) DNA, (35) steroid receptor (p. 614), (36) RNA polymerase (p. 725), (37) mRNA, (38) tRNA, (39) ribosome, (40) amino acyl-tRNA transferase (p. 757), (41) Hsp60 (p. 172).

OUTLINE

OXYGEN AND EVOLUTION: CHANCE AND NECESSITY

9.1 OXIDATION-REDUCTION REACTIONS
Redox Coenzymes
Aerobic Metabolism

9.2 CITRIC ACID CYCLE
Conversion of Pyruvate to Acetyl-CoA
Reactions of the Citric Acid Cycle
Fate of Carbon Atoms in the Citric Acid Cycle
The Amphibolic Citric Acid Cycle
Citric Acid Cycle Regulation
The Citric Acid Cycle and Human Disease
The Glyoxylate Cycle

Biochemistry in Perspective
Carcinogenesis: The Warburg Effect and Metabolic Reprogramming

AVAILABLE ONLINE
Biochemistry in Perspective
Hans Krebs and the Citric Acid Cycle
Biochemistry in Perspective
The Evolutionary History of the Citric Acid Cycle

Oxygen and Evolution: Chance and Necessity

Aerobic organisms use oxygen (O_2) to extract energy from nutrient molecules. *Aerobic respiration*, the oxygen-requiring energy-generating mechanism, generates significantly more energy than does fermentation, an anaerobic (without oxygen) process. When it first appeared about 2 billion years ago (bya), aerobic respiration marked a critical turning point in the evolution of life. The large increases in energy it supplied to organisms with the molecular equipment to exploit oxygen provided the resources for evolutionary innovation. It is not coincidental that eukaryotic cells (about 1.5 bya) and multicellular organisms (about 1 bya) originated after aerobic respiration became a common means of energy generation. Oxygen accumulated in the primordial atmosphere because it is a waste product of *oxygenic photosynthesis* (Chapter 13), a 3-billion-year-old process initiated by cyanobacteria that uses sunlight to drive the synthesis of biomolecules from CO_2 and the hydrogen atoms of water. Oxygenic photosynthesis is so complex and thermodynamically challenging, however, that it might never have come about at all.

Why Oxygen?

Why is oxygen so useful in energy generation? The answer lies in the nature of energy generation and the chemistry of oxygen and water. Energy is captured when electrons are transferred from an electron donor to an electron acceptor. In aerobic respiration, oxygen acts as the terminal acceptor of electrons removed from organic nutrients. As O_2 combines with these electrons as well as protons, water molecules are formed. Oxygen is an excellent electron acceptor for two major reasons. First, the element oxygen is abundant, which makes it far more useful as an oxidizing agent than less commonly available elements such as sulfur. Second, oxygen is a powerful oxidizing agent because it is highly electronegative (i.e., it has a considerable affinity for electrons). Consequently, there is a correspondingly large energy release with every electron that is transferred from carbon to oxygen.

Why Is Oxygenic Photosynthesis Such a Challenge?

The origin of oxygenic photosynthesis is improbable for two major reasons. First, because oxygen atoms are powerful electron acceptors, water is a very poor reducing agent. In contrast, nonoxygenic photosynthesis, the precursor of the oxygen-generating process, used more powerful reducing agents such as H_2 and H_2S. Significantly more light energy is required to oxidize water molecules than for H_2 and H_2S. As a result, the photosynthetic pigment molecules that absorb light energy had to be adapted to absorbing higher-energy photons (electromagnetic particles). Second, photosynthetic electron transport (one electron at a time) and water oxidation (a sequential and concerted four-electron process) are seemingly incompatible. The cyanobacteria solved this extraordinarily complex set of problems with a multisubunit pigment–enzyme complex now referred to as *photosystem II* (PSII).

The One and Only Oxygen-Evolving Complex

Within PSII, water molecules are split into electrons, protons, and O_2 by the *oxygen-evolving complex* (OEC), a protein complex that contains a unique inorganic cube-like cofactor: the Mn_3CaO_4 cubane. Of all of the problems solved by the cyanobacteria, it is the OEC that provides insight into the tenuous origin of oxygenic photosynthesis. Unlike all other biochemical devices, the OEC was invented only once. Consequently, all modern photosynthesizing organisms contain an identical water-splitting mechanism. The emergence of the water-splitting OEC of PSII was the most pivotal event in the history of life on Earth. Without the OEC, all modern organisms would be anaerobic prokaryotes.

Overview

MODERN AEROBIC ORGANISMS TRANSDUCE THE CHEMICAL BOND ENERGY OF FOOD MOLECULES INTO THE BOND ENERGY OF ATP. THEY PERFORM THIS feat because oxygen is used as the terminal acceptor of the electrons extracted from food molecules. The capacity to use oxygen to oxidize nutrients such as glucose and fatty acids yields a substantially greater amount of energy than does fermentation.

As atmospheric O_2 began accumulating on Earth, about 2 bya, existing organisms were confronted with a serious problem: molecular oxygen forms toxic oxygen ions and peroxides called *reactive oxygen species* (ROS). ROS react with and damage or destroy biomolecules. Consequently, exposure to O_2 acted as a severe selection pressure. Species in existence during this period either evolved a means of adapting to O_2 or became extinct. Modern

organisms are classified based on the strategies they use to cope with ROS or use O_2 in energy generation:

1. **Obligate anaerobes** are organisms that grow only in the absence of O_2 (i.e., they live in highly reduced environments such as soil) and use fermentation to generate energy.
2. **Aerotolerant anaerobes** depend on fermentation but possess enzymes and antioxidant molecules that detoxify ROS.
3. **Facultative anaerobes** not only possess the biochemical mechanisms required for detoxifying ROS but also can use O_2 as an electron acceptor when the gas is available.
4. **Obligate aerobes** are highly dependent on O_2 for energy production. They protect themselves from ROS with elaborate detoxifying mechanisms that are composed of enzymes and numerous endogenous and exogenous antioxidant molecules.

Aerobic metabolism consists of three biochemical processes: the citric acid cycle, the electron transport pathway, and oxidative phosphorylation (**Figure 9.1**). In eukaryotes, these processes occur within the mitochondrion (**Figure 9.2**). The citric acid cycle is a metabolic pathway in which two-carbon fragments derived from organic fuel molecules are oxidized to form CO_2 and the coenzymes NAD^+ and FAD are reduced to form NADH and $FADH_2$, respectively. The electron transport pathway, also called the electron transport chain (ETC), is a mechanism by which electrons are transferred from NADH and $FADH_2$ to a series of electron carriers that are sequentially reduced and then oxidized. The terminal electron acceptor is O_2. In oxidative phosphorylation, the energy released by electron transport is captured in the form of a proton gradient that drives the synthesis of ATP, the energy currency of living organisms.

Chapter 9 begins with a review of oxidation-reduction reactions and the relationship between electron flow and energy transduction. This review is followed by a detailed discussion of the citric acid cycle, the central pathway in aerobic metabolism, and its roles in energy generation and biosynthesis. In Chapter 10, the discussion of aerobic metabolism continues with an examination of electron transport and oxidative phosphorylation, the means by which aerobic organisms use oxygen to generate significant amounts of ATP. Chapter 10 ends with a review of *oxidative stress*, a series of reactions in which toxic oxygen species are created and subsequently damage cells. Chapter 10 also describes the principal mechanisms used by living organisms to protect themselves from oxidative stress.

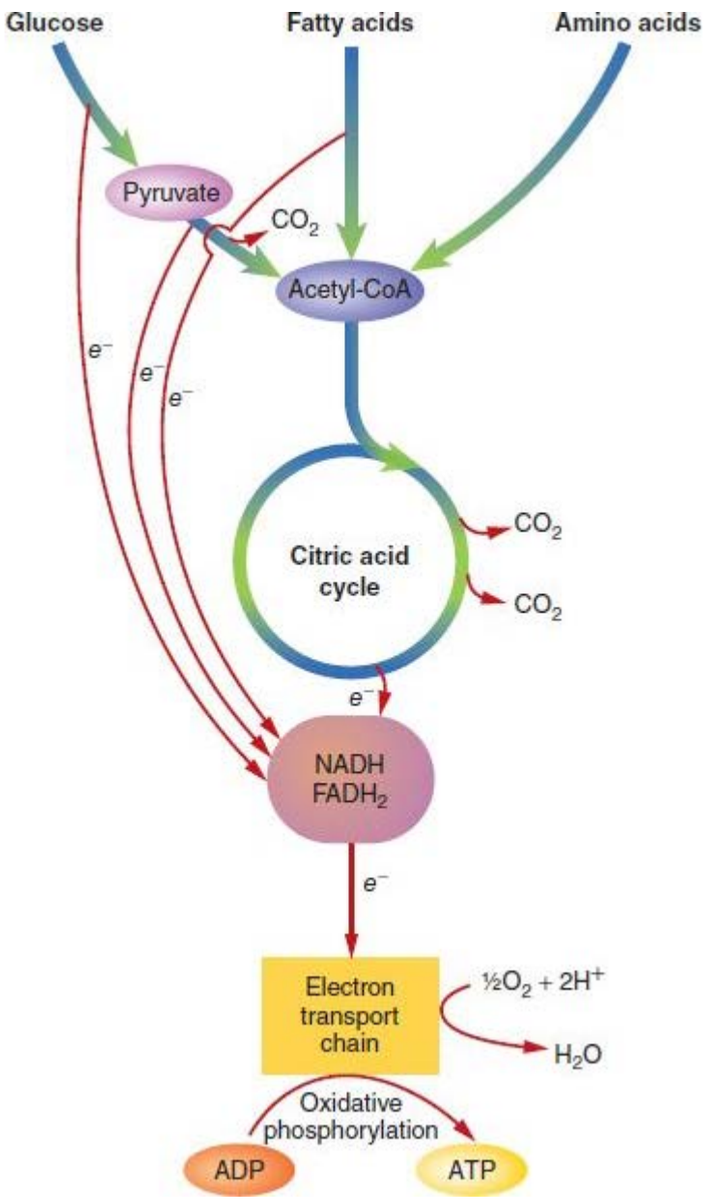


FIGURE 9.1

Overview of Aerobic Metabolism

In aerobic metabolism, the nutrient molecules glucose, fatty acids, and some amino acids are degraded to form acetyl-CoA. Acetyl-CoA then enters the citric acid cycle. Electron carriers (NADH and $FADH_2$) produced by glucose and fatty acid degradation and several citric acid cycle reactions donate electrons (e^-) to the electron transport chain. Energy captured by the electron transport chain is then used to synthesize ATP in a process referred to as oxidative phosphorylation. Note that O_2, the terminal electron acceptor in aerobic metabolism, combines with protons to form water molecules.

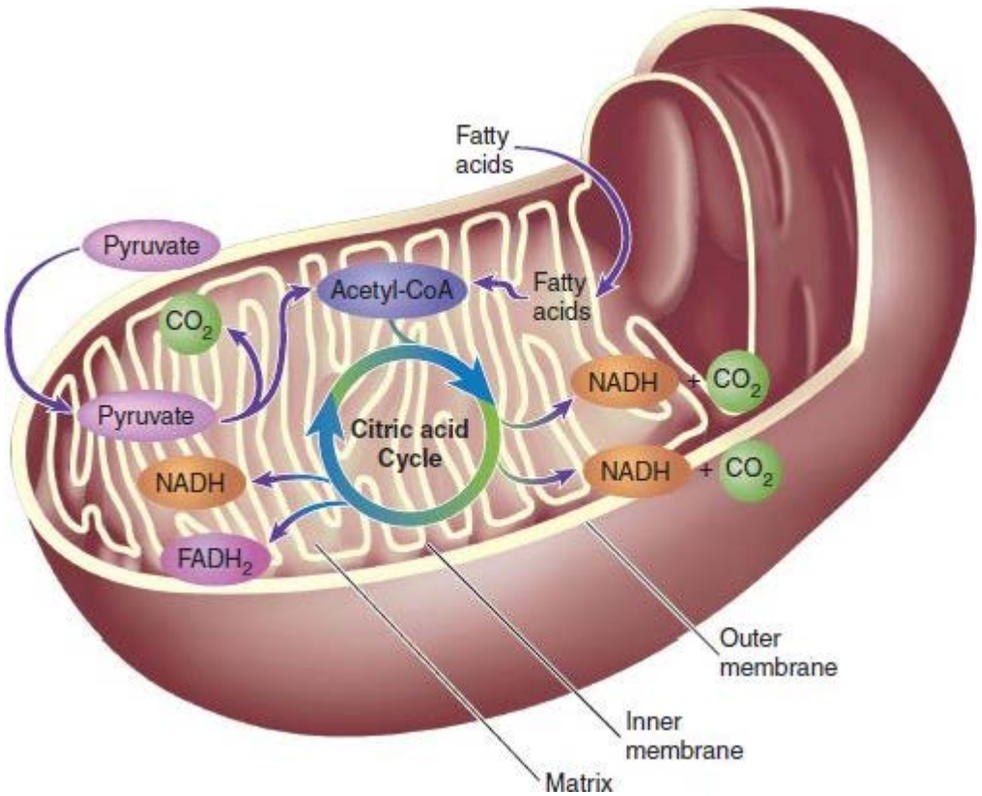


FIGURE 9.2

Aerobic Metabolism in the Mitochondrion

In eukaryotic cells, aerobic metabolism occurs within the mitochondrion. Acetyl-CoA, the oxidation product of pyruvate, fatty acids, and certain amino acids (not shown), is oxidized by the reactions of the citric acid cycle within the mitochondrial matrix. The principal products of the cycle are the reduced coenzymes, NADH and $FADH_2$, and CO_2. The high-energy electrons of NADH and $FADH_2$ are subsequently donated to the electron transport chain (ETC), a series of electron carriers in the inner membrane. The terminal electron acceptor for the ETC is O_2. The energy derived from the electron transport mechanism drives ATP synthesis by creating a proton gradient across the inner membrane. The large folded surface of the inner membrane is studded with ETC complexes, transport proteins of numerous types, and ATP synthase, the enzyme complex responsible for ATP synthesis. Small organic anions such as pyruvate, malate, and succinate, as well as ATP, ADP, NADH, and NAD^+, are exchanged across the outer mitochondrial membrane by VDAC (not shown). VDAC (voltage-dependent anion channels) allow the diffusion of small hydrophilic molecules as well as Ca^{2+} into and out of the intermembrane space. Proteins such as hexokinase, glucokinase (p. 297), glycerol kinase (p. 306), and creatine kinase (p. 380) can bind to VDAC.

9.1 OXIDATION-REDUCTION REACTIONS

In living organisms, both energy-capturing and energy-releasing processes consist largely of redox reactions. Recall that redox reactions occur when electrons are transferred between an electron donor (a reducing agent) and an electron acceptor (an oxidizing agent). In some redox reactions, only electrons are transferred. For example, in the reaction

$$Cu^+ + Fe^{3+} \rightleftharpoons Cu^{2+} + Fe^{2+}$$

an electron is transferred from Cu^+ to Fe^{3+}. Cu^+, the reducing agent, is oxidized to form Cu^{2+}. Meanwhile, Fe^{3+} is reduced to Fe^{2+}. In many reactions, however, both electrons and protons are transferred. For example, the reaction catalyzed by lactate dehydrogenase begins with the transfer

of a hydride ion ($H:^-$), that is, a hydrogen nucleus and two electrons, from NADH to pyruvate. A proton (H^+) is gained from the environment (**Figure 9.3**) to form the final products lactate and NAD^+.

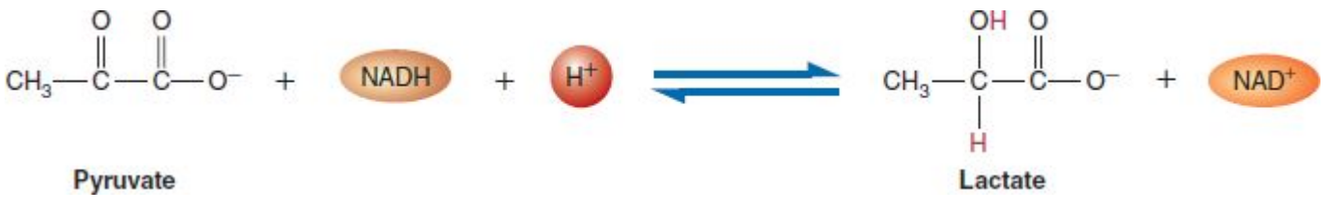


FIGURE 9.3

Reduction of Pyruvate by NADH

In this redox reaction, a hydride ion ($H:^-$) is transferred from NADH to pyruvate, and the product is protonated from the surrounding medium to form lactate.

Redox reactions are more easily understood if they are separated into half-reactions. In the reaction between copper and iron, the Cu^+ ion loses an electron to become Cu^{2+}:

$$Cu^+ \rightleftharpoons Cu^{2+} + e^-$$

This equation indicates that Cu^+ is the electron donor. (Together Cu^+ and Cu^{2+} constitute a **conjugate redox pair**.) As Cu^+ loses an electron to form Cu^{2+}, Fe^{3+} gains an electron to form Fe^{2+}:

$$Fe^{3+} + e^- \rightleftharpoons Fe^{2+}$$

In this half-reaction, Fe^{3+} is an electron acceptor. The separation of redox reactions emphasizes that electrons are always the common intermediates between half-reactions.

The constituents of half-reactions may be observed in an electrochemical cell (**Figure 9.4**). Each half-reaction takes place in a separate container, or *half-cell*. The movement of electrons generated in the half-cell undergoing oxidation (e.g., $Cu^+ \rightarrow Cu^{2+} + e^-$) creates a voltage (or potential difference) between the two half-cells. The sign of the voltage (measured by a voltmeter) is positive or negative according to the direction of the electron flow. The magnitude of the potential difference is a measure of the energy that drives the reaction.

The tendency for a specific substance to gain electrons is called its **reduction potential**. The **standard reduction potential** ($E°$) of a substance is measured in a galvanic cell relative to a standard hydrogen electrode. A standard cell has all solutes at 1.0 M concentration, all gases at 1 atm pressure, and the temperature at 25°C. The reduction potential for the half-reaction, $2H^+ + 2e^- \rightarrow H_2$ (g), against the standard hydrogen electrode is set at 0.00 V.

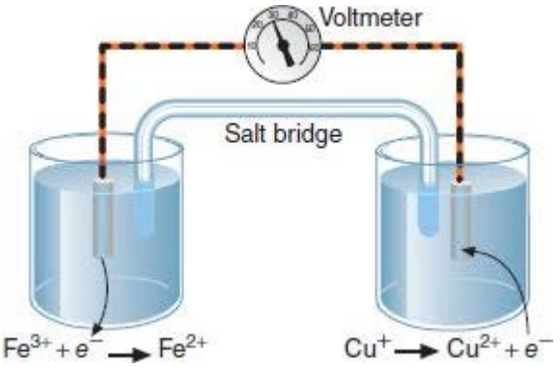


FIGURE 9.4

An Electrochemical Cell

Electrons flow from the Cu^{2+}/Cu^+ half-cell (cathode) through the voltmeter to the Fe^{3+}/Fe^{2+} half-cell (anode). The salt bridge containing KCl completes the electrical circuit. The voltmeter measures the electrical potential, which drives electrons from one half-cell to the other.

In biochemistry, the reference half-reaction is

$$2H^+ + 2e^- \rightleftharpoons H_2$$

when pH = 7
temperature = 25°
pressure = 1 atm

Under these conditions, the reduction potential ($E°'$) of the hydrogen electrode is –0.42 V when measured against the standard hydrogen electrode in which the hydrogen ion concentration is 1 M. Substances with reduction potentials lower than –0.42 V (i.e., those with more negative values) have a lower affinity for electrons than does H^+. Substances with higher reduction potentials (i.e., those with more positive values) have a greater affinity for electrons (**Table 9.1**). The pH in the test electrode is 7.0 for each of the redox half-reactions, and the pH of the reference standard electrode is 0 or the $[H^+]$ is 1.0 M.

A substance with a more negative (less positive) reduction potential will receive electrons from a substance with a more positive reduction potential, and the overall cell potential ($\Delta E°'$) will be positive. The relationship between $\Delta E°'$ and $\Delta G°'$ is

$$\Delta G°' = -nF\,\Delta E°'$$

TABLE 9.1 Standard Reduction Potentials*

Redox Half-Reaction	Standard Reduction Potentials ($E°'$) (V)
$2H^+ + 2e^- \rightarrow H_2$	−0.42
α-Ketoglutarate + CO_2 + $2H^+$ + $2e^-$ → isocitrate	−0.38
$NADP^+ + H^+ + 2e^- \rightarrow$ NADPH	−0.324
+ + −	

$NAD + H + 2e \rightarrow NADH$	−0.32
$S + 2H^+ + 2e^- \rightarrow H_2S$	−0.23
$FAD + 2H^+ + 2e^- \rightarrow FADH_2$	−0.22
Acetaldehyde + $2H^+ + 2e^- \rightarrow$ ethanol	−0.20
Pyruvate + $2H^+ + 2e^- \rightarrow$ lactate	−0.19
Oxaloacetate + $2H^+ + 2e^- \rightarrow$ malate	−0.166
$Cu^{2+} + e^- \rightarrow Cu^+$	+0.16
Fumarate + $2H^+ + 2e^- \rightarrow$ succinate	+0.031
Cytochrome b (Fe^{3+}) + $e^- \rightarrow$ cytochrome b (Fe^{2+})	+0.075
Cytochrome c_1 (Fe^{3+}) + $e^- \rightarrow$ cytochrome c_1 (Fe^{2+})	+0.22
Cytochrome c (Fe^{3+}) + $e^- \rightarrow$ cytochrome c (Fe^{2+})	+0.235
Cytochrome a (Fe^{3+}) + $e^- \rightarrow$ cytochrome a (Fe^{2+})	+0.29
$NO_3^- + 2H^+ + 2e^- \rightarrow NO_2^- + H_2O$	+0.42
$NO_2^- + 8H^+ + 6e^- \rightarrow NH_4^+ + 2H_2O$	+0.44
$Fe^{3+} + e^- \rightarrow Fe^{2+}$	+0.77
$^1/_2O_2 + 2H^+ + 2e^- \rightarrow H_2O$	+0.82

* By convention, redox reactions are written with the reducing agent to the right of the oxidizing agent and the number of electrons transferred. In this table, the redox pairs are listed in order of increasing $E°'$ values. The more negative the $E°'$ value is for a redox pair, the lower the affinity of the oxidized component for electrons. The more positive the $E°'$ value is, the greater the affinity of the oxidized component of the redox pair for electrons. Under appropriate conditions, a redox half-reaction reduces any of the half-reactions below it in the table.

where $\Delta G°'$ = the standard free energy at pH 7
n = the number of electrons transferred
F = the Faraday constant (96,485 J/V·mol)
$\Delta E°'$ = the difference in reduction potential between the electron donor and the electron acceptor under standard conditions

Living organisms utilize redox coenzymes as high-energy electron carriers. The most prominent examples are described next.

Redox Coenzymes

The coenzyme forms of the vitamin molecules nicotinic acid and riboflavin are universal electron carriers. Their structural and functional properties are as follows.

NICOTINIC ACID There are two coenzyme forms of nicotinic acid: **nicotinamide adenine dinucleotide (NAD)** and **nicotinamide adenine dinucleotide phosphate (NADP)**. These coenzymes occur in oxidized forms (NAD^+ and $NADP^+$) and reduced forms (NADH and NADPH). The structures of NAD^+ and $NADP^+$ both contain adenosine and the *N*-ribosyl derivative of nicotinamide (derived from the vitamin niacin), which are linked together through a pyrophosphate group (**Figure 9.5a**). $NADP^+$ has an additional phosphate attached to the 2′ OH group of adenosine. (The ring atoms of the sugar in a nucleotide are designated with a prime to distinguish them from atoms in the base.) Both NAD^+ and $NADP^+$ carry electrons for several enzymes in a group known as the dehydrogenases. (Dehydrogenases catalyze hydride transfer reactions. Many dehydrogenases that catalyze reactions involved in energy generation use the coenzyme NADH. The enzymes that require NADPH usually catalyze biosynthetic reactions (e.g., fatty acid synthesis, pp. 460–66, and quench [detoxify] ROS, pp. 389–90). A small number of dehydrogenases can use either NADH or NADPH.)

Recall that alcohol dehydrogenase catalyzes the reversible oxidation of ethanol to form acetaldehyde (p. 216).

During this reaction NAD^+ accepts a hydride ion from ethanol, the substrate molecule undergoing oxidation. The product deprotonates to form the acetaldehyde molecule. The reversible reduction of NAD^+ is illustrated in **Figure 9.5b**.

In most reactions catalyzed by dehydrogenases, the NAD^+ (or $NADP^+$) is bound only transiently to the enzyme. After the reduced version of the coenzyme is released from the enzyme, it donates the hydride ion to another molecule, called an *electron acceptor*, with a more positive reduction potential than NADH.

RIBOFLAVIN Riboflavin (vitamin B_2) is a component of two coenzymes: **flavin mononucleotide (FMN)** and **flavin adenine dinucleotide (FAD)** (**Figure 9.6**). FMN and FAD function as tightly bound prosthetic groups in a class of enzymes known as the **flavoproteins**. Flavoproteins are a diverse group of redox enzymes; they function as dehydrogenases, oxidases, and hydroxylases. These enzymes use the isoalloxazine group of FAD or FMN as a donor or acceptor of two hydrogen atoms. FMN plays a key role in the link between two-electron transfer reactions in the mitochondrial matrix and the one-electron transfer reactions of the electron transport chain because it can transfer one hydrogen atom at a time. Succinate dehydrogenase is a prominent example of a flavoprotein. It catalyzes the oxidation of succinate by FAD to form fumarate and $FADH_2$, an important reaction in the citric acid cycle.

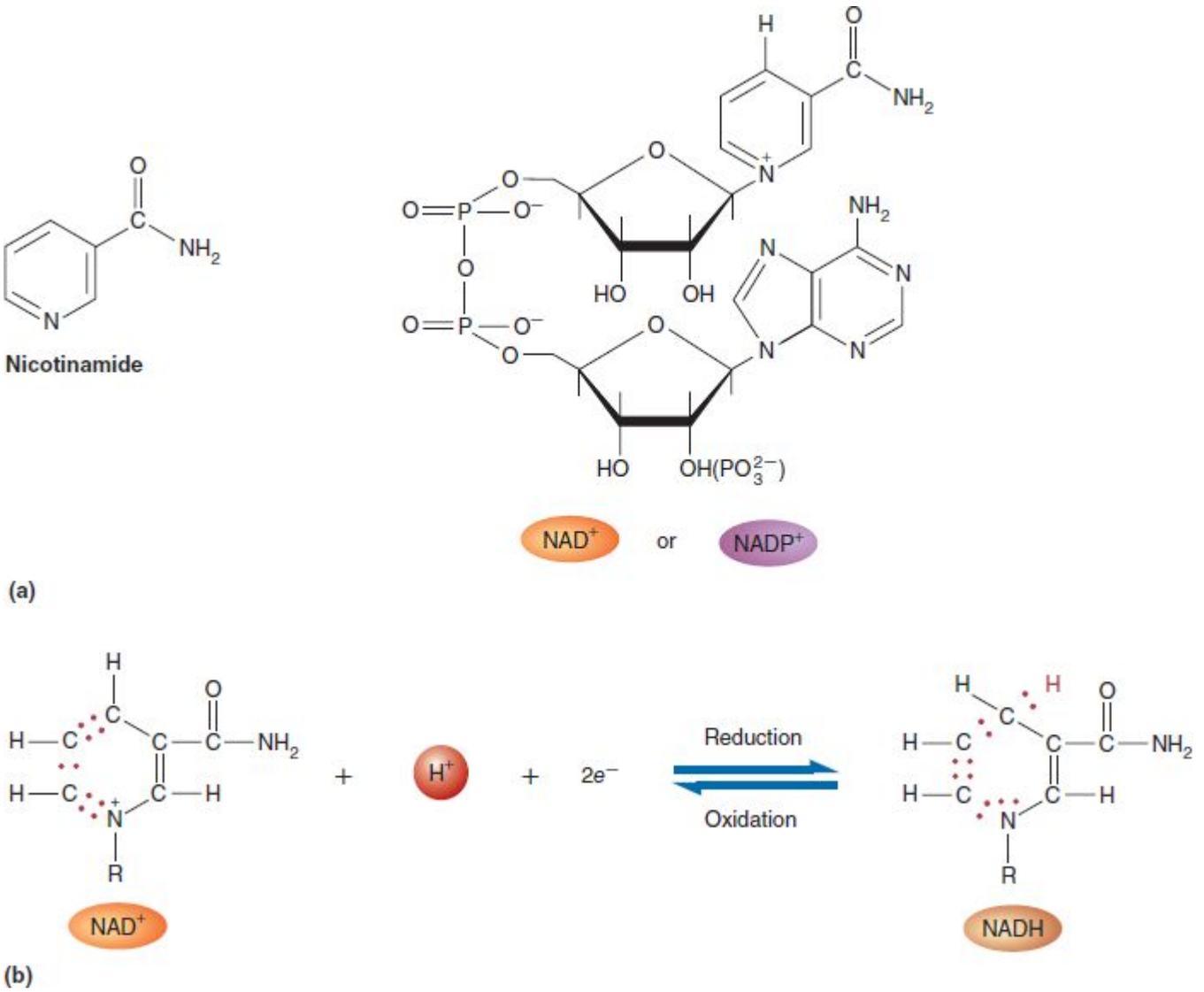


FIGURE 9.5

Nicotinamide Adenine Dinucleotide (NAD)

(a) Nicotinamide and $NAD(P)^+$. (b) Reversible reduction of NAD^+ to NADH. To simplify the equation, only the nicotinamide ring is shown. The rest of the molecule is designated R.

QUESTION 9.1

Use **Table 9.1** to determine which of the following reactions will proceed as written:

$$CH_3CH_2OH + 2\ \text{cyt b}\ (Fe^{3+}) \rightarrow CH_3CHO + 2\ \text{cyt b}\ (Fe^{2+}) + 2H^+$$

$$NO_2^- + H_2O + 2\ \text{cyt b}\ (Fe^{3+}) \rightarrow 2\ \text{cyt b}\ (Fe^{2+}) + NO_3^- + 2H^+$$

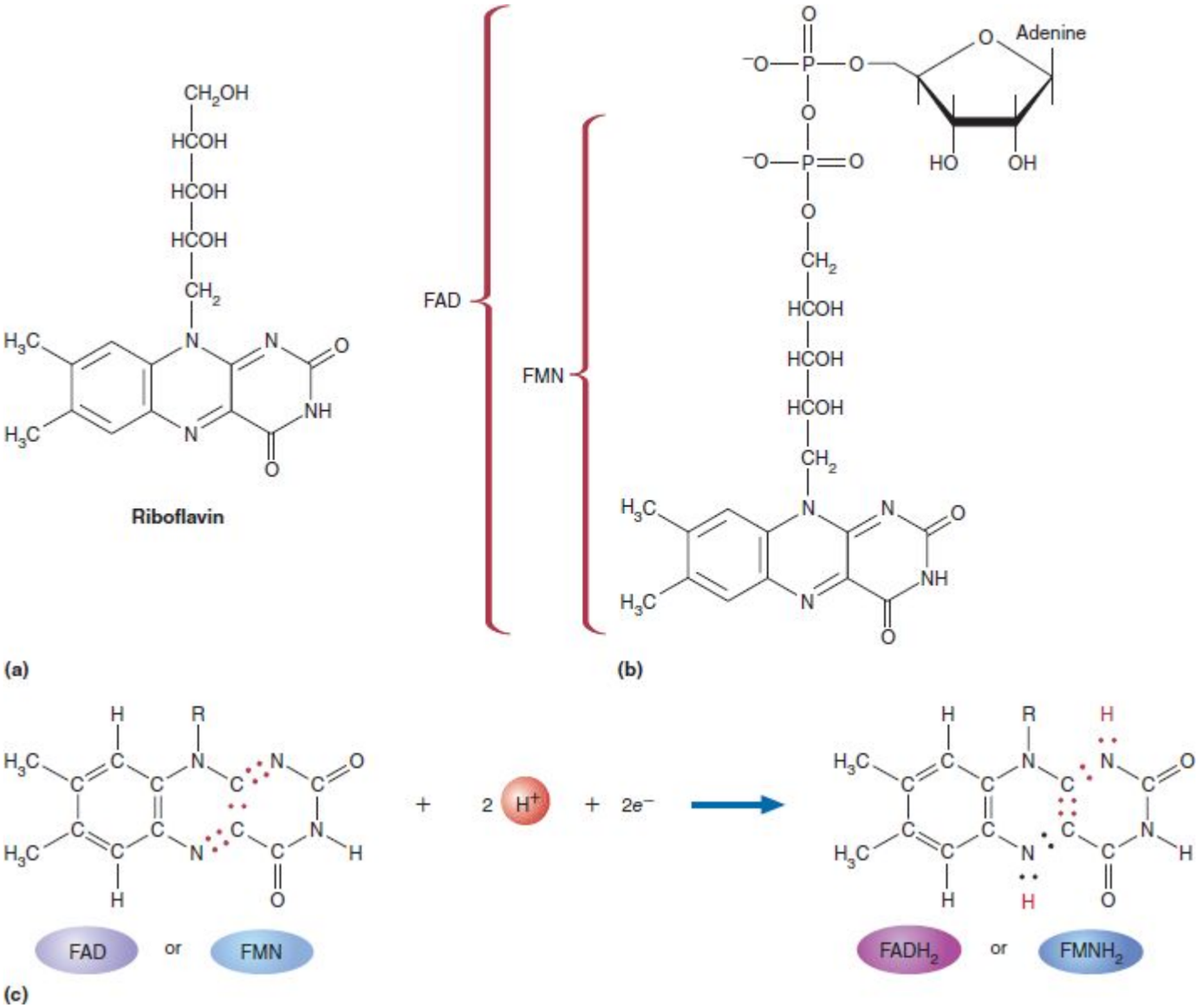


FIGURE 9.6

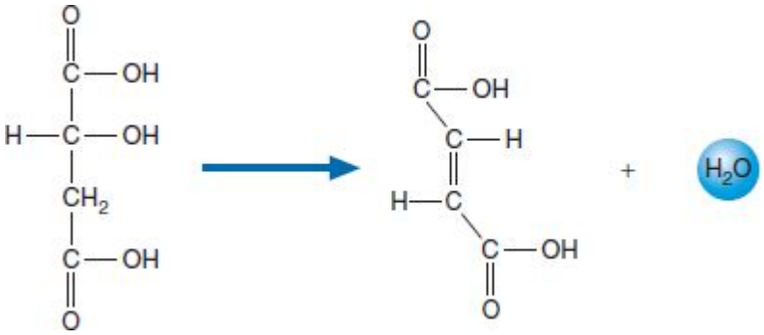
Flavin Coenzymes

(a) The vitamin riboflavin consists of an isoalloxazine ring system linked to ribitol (an alcohol formed by the reduction of ribose). (b) Structure of FAD and FMN. (c) Reversible reduction of flavin coenzymes: to simplify the equation, only the isoalloxazine ring system is shown. The rest of the coenzyme is designated R.

QUESTION 9.2

Which of the following reactions are redox reactions? For each redox reaction, identify the oxidizing and reducing agents.

1. Glucose + ATP → glucose-1-phosphate + ADP

2. 

3. Lactate + $NAD^+ \rightarrow$ pyruvate + NADH + H^+

4. $NO_2^- + 8H^+ + 6$ cyt b $(Fe^{2+}) \rightarrow NH_4^+ + 2H_2O + 6$ cyt b (Fe^{3+})

5. $CH_3CHO + NADH + H^+ \rightarrow CH_3CH_2OH + NAD^+$

WORKED PROBLEM 9.1

Use the following half-cell potentials to calculate (a) the overall cell potential and (b) $\Delta G^{\circ\prime}$.

$$\text{Succinate} + {}^1/_2O_2 \rightarrow \text{fumarate} + H_2O$$

The half-reactions are

$$\text{Succinate} \rightarrow \text{fumarate} + 2H^+ + 2e^- \qquad (E^{\circ\prime} = -0.031\ \text{V})$$
$$\tfrac{1}{2}\ O_2 + 2H^+ + 2e^- \rightarrow H_2O \qquad (E^{\circ\prime} = +0.82\ \text{V})$$

SOLUTION

Write the fumarate reaction as an oxidation (lower reduction potential), balance if the number of electrons transferred differs between the two half-reactions (not necessary here), and add the two reactions to get the net reaction.

$$\text{Succinate} + {}^1/_2O_2 \rightarrow \text{fumarate} + H_2O$$

a. The overall potential is defined by the following:

$$\Delta E^{\circ\prime} = E^{\circ\prime}\ (\text{electron acceptor}) - E^{\circ\prime}\ (\text{electron donor})$$
$$\Delta E^{\circ\prime} = (+0.82\ \text{V}) - (-0.031\ \text{V})$$
$$\Delta E^{\circ\prime} = +0.85\ \text{V}$$

b. Use the formula to find $\Delta G^{\circ\prime}$.

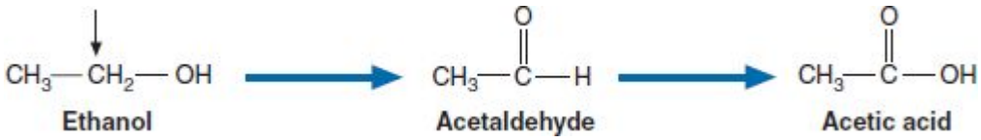
$$\Delta G^{\circ\prime} = -nF\Delta E^{\circ\prime}$$
$$= -(2)(96.5\ \text{kJ/V.mol})(0.85\ \text{V})$$
$$= -164.05\ \text{kJ/mol}$$
$$= -164\ \text{kJ/mol}$$

■

QUESTION 9.3

Because redox reactions play an important role in living processes, biochemists need to determine the oxidation state of the atoms in a molecule. In one method, the oxidation state of an atom is determined by assigning numbers to carbon atoms based on the type of groups attached to them. For example, a bond to a hydrogen is assigned the value −1. A bond to another carbon atom is valued at 0, and a bond to an electronegative atom such as oxygen or nitrogen is valued at +1. The values of a single carbon atom in a molecule may range from −4 (e.g., CH_4) to +4 (CO_2). Note that methane is a high-energy molecule and carbon dioxide is a low-energy molecule. As carbon changes its oxidation state from −4 to +4, a large amount of energy is released. This process is therefore highly exothermic.

Ethanol is degraded in the liver by a series of redox reactions. Identify the oxidation state of the indicated carbon atom in each molecule in the following reaction sequence:



QUESTION 9.4

As CO_2 is incorporated into a sugar molecule during photosynthesis, is it being oxidized or reduced?

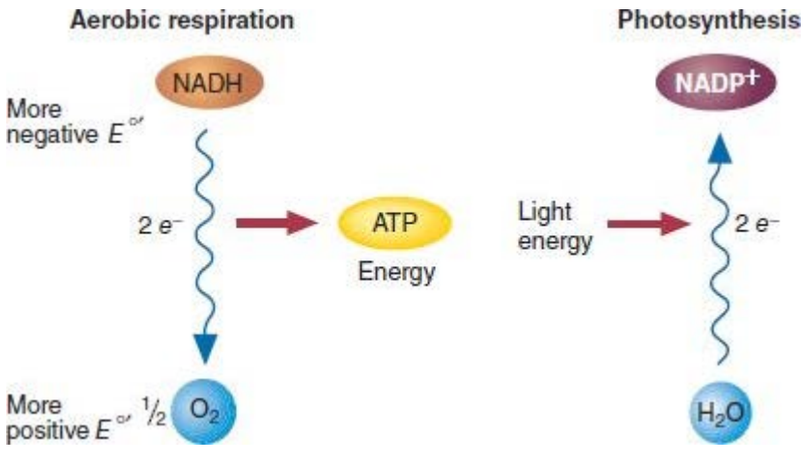


FIGURE 9.7

Electron Flow and Energy

Electron flow is used to generate and capture energy in aerobic respiration. Radiant (light) energy is used to drive electron flow in photosynthesis. Note that the energy captured by photosynthesis in the chemical bonds of sugars and other biomolecules is released by aerobic respiration and used to synthesize ATP.

Aerobic Metabolism

Most of the aerobic cell's free energy is captured by the mitochondrial ETC. During this process,

electrons are transferred from a redox pair with a more negative reduction potential ($NADH/NAD^+$) to those with more positive reduction potentials. The last component in the system is the $H_2O/½\ O_2$ pair:

$$½\ O_2 + NADH + H^+ \rightarrow H_2O + NAD^+$$

The free energy released as a pair of electrons passes from NADH to O_2 under standard conditions is calculated as follows:

$$\begin{aligned}\Delta G^{\circ\prime} &= -nF\Delta E^{\circ\prime} \\ &= -2(96.5\ kJ/V{\cdot}mol)[0.815 - (-0.32)] \\ &= -220\ kJ/mol\end{aligned}$$

A significant portion of the free energy generated as electrons move from NADH to O_2 in the ETC is used to drive ATP synthesis.

KEY CONCEPTS

- In living organisms, both energy-capturing and energy-releasing processes consist primarily of redox reactions.
- In redox reactions, electrons move between an electron donor and an electron acceptor.
- In many reactions, both electrons and protons are transferred.
- In biological systems, most redox reactions involve hydride ion transfer ($NAD(P)H/NAD(P)^+$) or hydrogen atom transfer ($FADH_2/FAD$).

In several metabolic processes, electrons move from redox pairs with more positive reduction potentials to those with more negative reduction potentials. Of course, energy is required. The most prominent example of this phenomenon is photosynthesis (Chapter 13). Photosynthetic organisms use captured light energy to drive electrons from electron donors, such as water, to electron acceptors with more negative reduction potentials (**Figure 9.7**). The energized electrons eventually flow back to acceptors with more positive reduction potentials, thereby providing energy for ATP synthesis and CO_2 reduction to form carbohydrate.

In Section 9.2, the citric acid cycle is examined. In this pathway, which is the first phase of aerobic metabolism, the energy released by the oxidation of two-carbon fragments derived from glucose, fatty acids, and some amino acids is captured by and carried in the reduced coenzymes NAD(P)H and $FADH_2$.

9.2 CITRIC ACID CYCLE

The citric acid cycle (**Figure 9.8**) is a series of biochemical reactions that aerobic organisms use to release chemical energy stored in the two-carbon acetyl group in acetyl-CoA. Acetyl-CoA is composed of an acetyl group derived from the breakdown of carbohydrates, lipids, and some amino acids that is linked to the acyl carrier molecule **coenzyme A** (**Figure 9.9**). Acetyl-CoA is synthesized from pyruvate and is also the product of fatty acid catabolism (described in Chapter 11) and certain reactions in amino acid metabolism (Chapter 15). In the citric acid cycle, the acetyl group's carbon atoms are eventually oxidized to CO_2, and the electrons are transferred to NAD^+ and FAD.

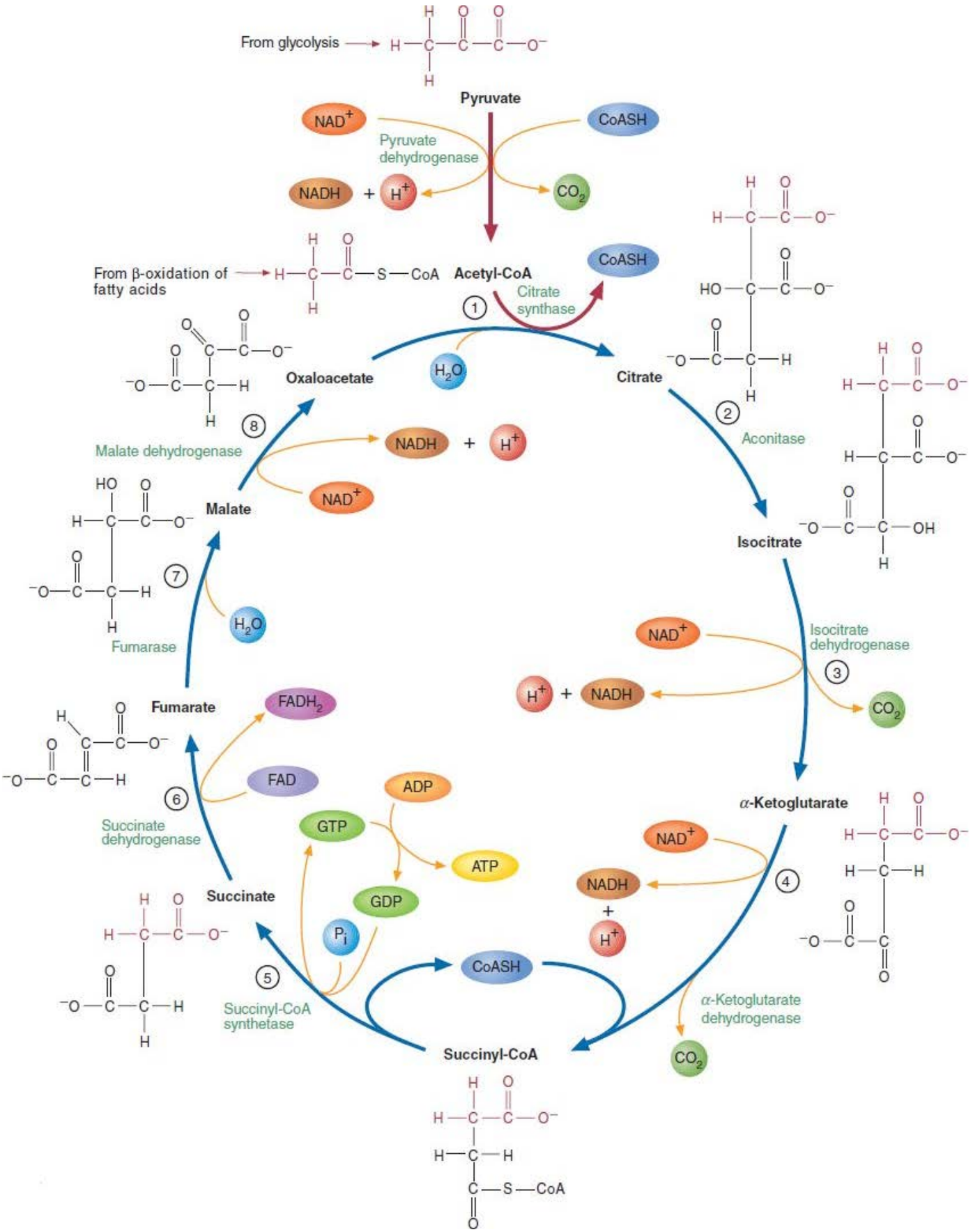


FIGURE 9.8

The Citric Acid Cycle

In each turn of the cycle, acetyl-CoA from the glycolytic pathway or from fatty acid catabolism enters, and two fully oxidized carbon molecules leave as CO_2. Three molecules of NAD^+ and the molecule of FAD are reduced. One molecule of GTP (interconvertible with ATP) is generated in a substrate-level phosphorylation reaction.

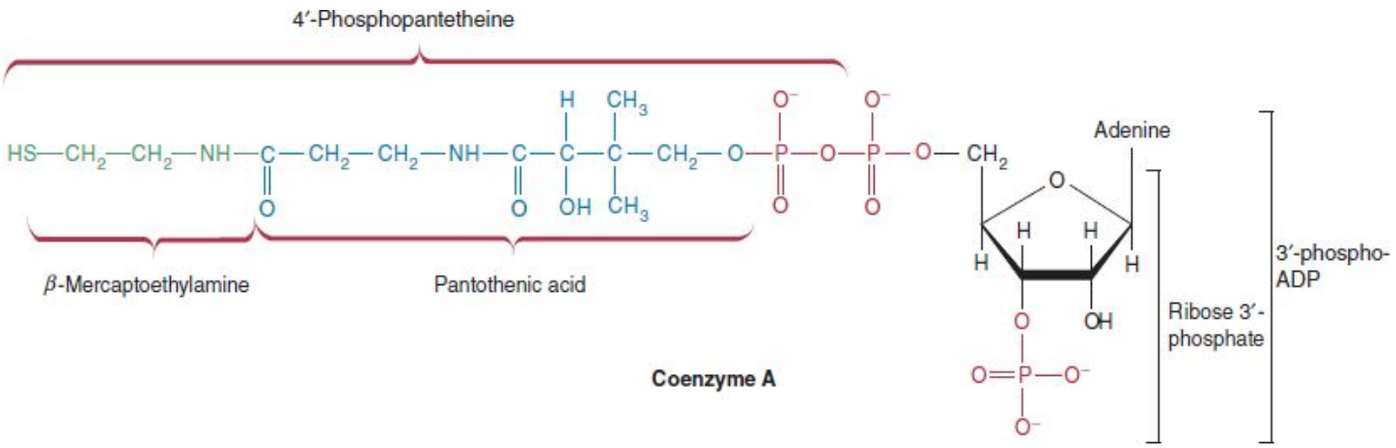


FIGURE 9.9

Coenzyme A

In coenzyme A, a 3′-phosphate derivative of ADP is linked to pantothenic acid via a phosphate ester bond. The β-mercaptoethylamine group of coenzyme A is attached to pantothenic acid by an amide bond. Coenzyme A is a carrier of acyl groups that range in size from the acetyl group to long-chain fatty acids. Because the reactive SH group forms a thioester bond with acyl groups, coenzyme A is often abbreviated as CoASH. Note that sulfur is a better leaving group than oxygen. Consequently, the carbon–sulfur linkage of a thioester is a high-energy bond that is more easily cleaved than the carbon–oxygen bond of an ester.

In the first reaction of the citric acid cycle, a two-carbon acetyl group condenses with a four-carbon molecule (oxaloacetate) to form a six-carbon molecule (citrate). During the subsequent seven reactions, in which two CO_2 molecules are produced and four pairs of electrons are removed from carbon compounds, citrate is reconverted to oxaloacetate. During one step in the cycle, the high-energy molecule GTP is produced during a substrate-level phosphorylation. The net reaction for the citric acid cycle is as follows:

$$\text{Acetyl-CoA} + 3\,NAD^+ + FAD + GDP + P_i + 2\,H_2O \rightarrow 2\,CO_2 + 3\,NADH + FADH_2 + CoASH + GTP + 2H^+$$

In addition to its role in energy production, the citric acid cycle plays another important role in metabolism. Cycle intermediates are substrates in a variety of biosynthetic reactions. **Table 9.2** provides a summary of the roles of coenzymes in the citric acid cycle.

Conversion of Pyruvate to Acetyl-CoA

After its transport into the mitochondrial matrix, pyruvate is converted to acetyl-CoA in a series of reactions catalyzed by the enzymes in the pyruvate dehydrogenase complex (PDHC). The net reaction, an oxidative decarboxylation, is as follows:

$$\text{Pyruvate} + NAD^+ + \text{CoASH} \rightarrow \text{Acetyl-CoA} + NADH + CO_2 + H^+$$

TABLE 9.2 Summary of the Coenzymes in the Citric Acid Cycle

Coenzyme	Functions
Thiamine pyrophosphate (TPP)	Decarboxylation and aldehyde group transfer
Lipoic acid	Carrier of hydrogens or acetyl groups
NADH	Electron carrier
$FADH_2$	Electron carrier
Coenzyme A (CoASH)	Acetyl group carrier

Despite the apparent simplicity of this highly exergonic reaction ($\Delta G°' = -33.5$ kJ/mol), its mechanism is one of the most complex known. The PDHC (**Figure 9.10**) is a large multienzyme structure that contains multiple copies of three enzyme activities: pyruvate dehydrogenase (E_1), dihydrolipoyl transacetylase (E_2), and dihydrolipoyl dehydrogenase (E_3). **Table 9.3** summarizes the number of copies of each enzyme and the required coenzymes of the *E. coli* pyruvate dehydrogenase complex.

In the first step, pyruvate dehydrogenase catalyzes the decarboxylation of pyruvate (**Figure 9.11**). A nucleophile is formed when a basic residue of the enzyme extracts a proton from the thiazole ring of **thiamine pyrophosphate** (TPP). (TPP is the coenzyme form of thiamine, also called vitamin B_1.) The intermediate, hydroxyethyl-TPP (HETPP), forms after the nucleophilic thiazole ring has attacked the carbonyl group of pyruvate, with the resulting loss of CO_2.

In the next several steps, the hydroxyethyl group of HETPP is converted to acetyl-CoA by dihydrolipoyl transacetylase. **Lipoic acid** (**Figure 9.12**), an acyl transfer coenzyme that contains two thiol groups that can be reversibly oxidized, plays a crucial role in this transformation. Lipoic acid is bound to the enzyme through an amide linkage with the ε-amino group of a lysine side chain, forming a flexible arm. Lipoamide reacts with HETPP to form an acetylated lipoamide and free TPP. The lipoamide then transfers the acetyl group to the sulfhydryl group of coenzyme A. Subsequently, the reduced lipoamide is transferred to the dihydrolipoyl dehydrogenase active site where it is reoxidized by a bound FAD to yield $FADH_2$. The reduction potential of $FADH_2$ is sufficiently altered by its association with the enzyme that it can donate a hydride ion to an incoming NAD^+. (Ordinarily, $FADH_2$ cannot reduce NAD^+ because NAD^+ has a more negative reduction potential.) The mobile NADH can deliver its electrons to the ETC and is replaced in the enzyme by another NAD^+ molecule so that the cycle can begin again.

KEY CONCEPTS

Pyruvate is converted to acetyl-CoA by the enzymes in the pyruvate dehydrogenase complex. TPP, FAD, NAD^+, Coenzyme A, and lipoic acid are required coenzymes.

PDHC is stringently regulated because of its central role in energy metabolism, linking glycolysis to the citric acid cycle. Its activity is controlled largely through allosteric effectors and

+

covalent modification. The enzyme complex is allosterically activated by NAD , CoASH, and AMP. It is inhibited by high concentrations of ATP and the reaction products acetyl-CoA and NADH. In mammals, acetyl-CoA and NADH also activate pyruvate dehydrogenase kinase 1 (PDK1), which converts the active PDHC to an inactive phosphorylated form. High concentrations of the substrates pyruvate, CoASH, and NAD^+ inhibit the activity of the kinase. The PDHC is reactivated by a dephosphorylation reaction catalyzed by pyruvate dehydrogenase phosphatase (PDP). PDP is activated when the mitochondrial ATP concentration is low. PDP is also activated by Ca^{2+} and insulin.

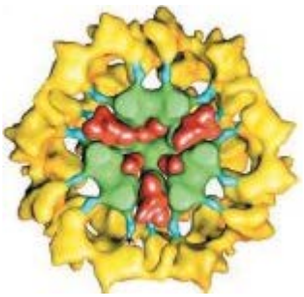


FIGURE 9.10

Pyruvate Dehydrogenase Complex

This cross section of the PDHC (*Saccharomyces cerevisiae*) shows the relative positions of the E_1 subunits (pyruvate dehydrogenase, yellow), the E_2 subunits (dihydrolipoyl transacetylase, green), the lipoyl domains of E_2 (blue, reaching outward to the active sites of E_1), and the E_3 subunits (dihydrolipoyl dehydrogenase) in the core (red), which can also be reached by the swinging arms of the lipoamide domains of E_2.

TABLE 9.3 *E. coli* Pyruvate Dehydrogenase Complex

Enzyme Activity	Function	Copies per Complex*	Coenzymes
Pyruvate dehydrogenase (E_1)	Decarboxylates pyruvate	24 (20–30)	TPP
Dihydrolipoyl transacetylase (E_2)	Catalyzes transfer of acetyl group to CoASH	24 (60)	Lipoic acid, CoASH
Dihydrolipoyl dehydrogenase (E_3)	Reoxidizes dihydrolipoamide	12 (20–30)	NAD^+, FAD

* The number of molecules of each enzyme activity found in mammalian pyruvate dehydrogenase is shown in parentheses.

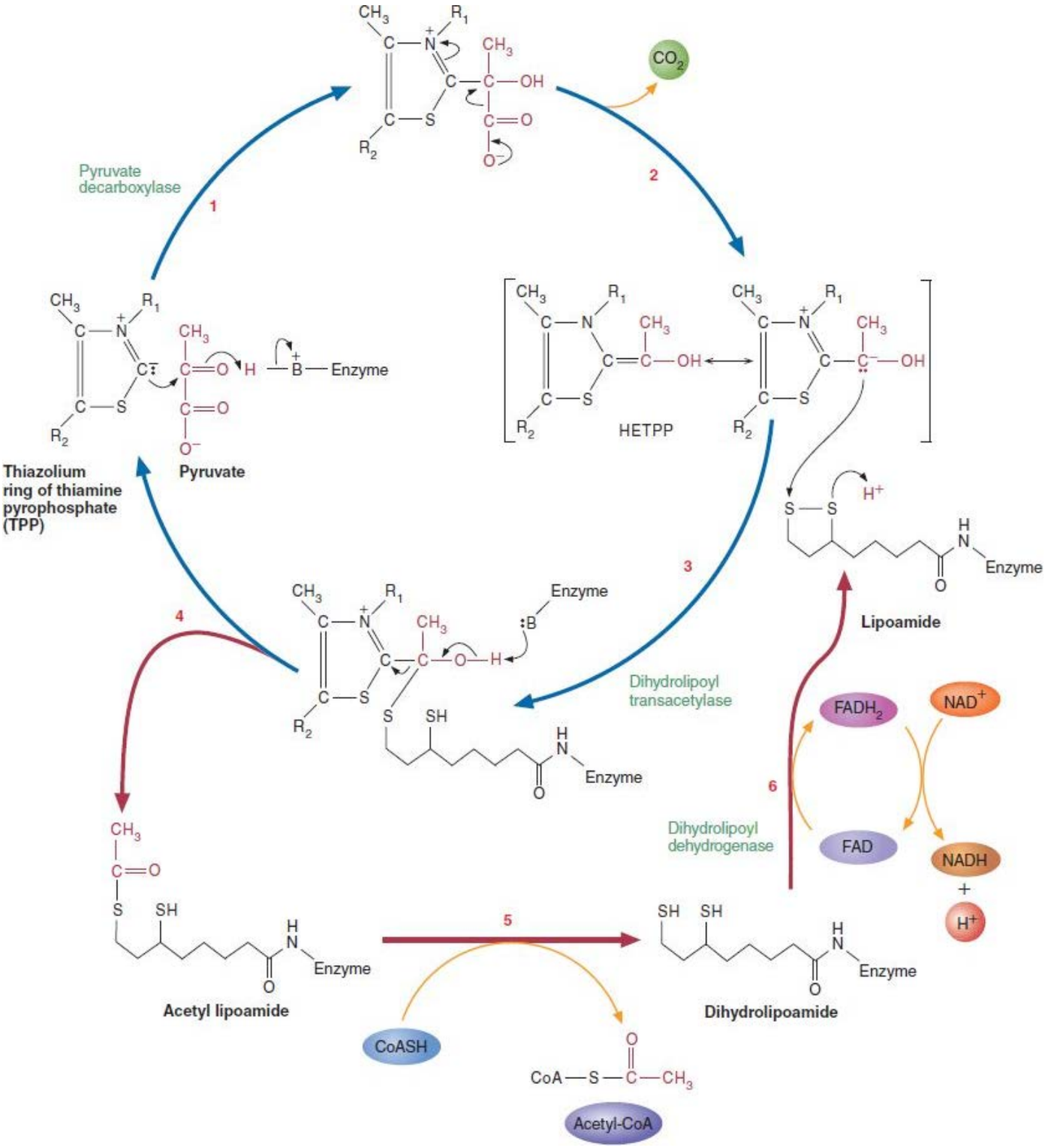


FIGURE 9.11

The Reactions Catalyzed by the Pyruvate Dehydrogenase Complex

Initially, pyruvate dehydrogenase (decarboxylase) catalyzes the formation of HETTP (hydroxyethyl-TPP) using TPP as a coenzyme. In step 1, TPP in a zwitterionic form performs a nucleophilic attack on the C2 carbonyl of pyruvate to produce the reactive intermediate that decarboxylates (step 2) to yield CO_2 and HETPP. In step 3, dihydrolipoyl transacetylase uses lipoamide (lipoic acid covalently linked to a lysine side chain) to convert the hydroxyethyl group of HETPP to a thioester of lipoamide. A resonance stabilized HETPP attacks S1 of oxidized lipoamide (step 4) to yield acetyl lipoamide, thereby converting S2 into a sulfhydryl group. A subsequent transacetylation reaction (step 5) transfers the acetyl group to the thiol group of coenzyme A. In step 6, the FAD in dihydrolipoyl dehydrogenase reoxidizes the reduced lipoamide, yielding $FADH_2$ and

dihydrolipoamide. FAD is regenerated when $FADH_2$ donates a hydride ion to NAD^+. (Refer to Figure 9.12 for the structure of lipoamide.)

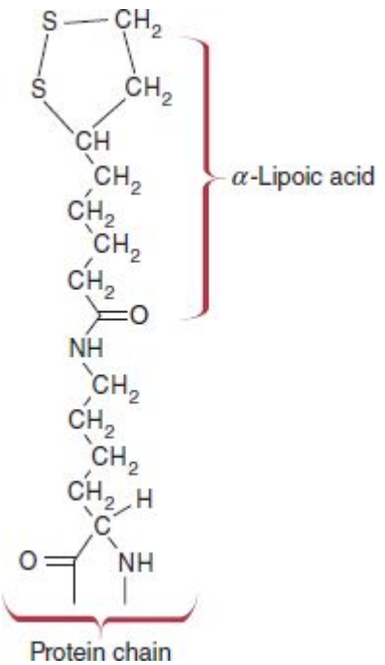


FIGURE 9.12

Lipoamide

Lipoic acid is covalently bonded to the enzyme through an amide linkage with the ε-amino group of a lysine residue.

3D animation of lipoamide

Reactions of the Citric Acid Cycle

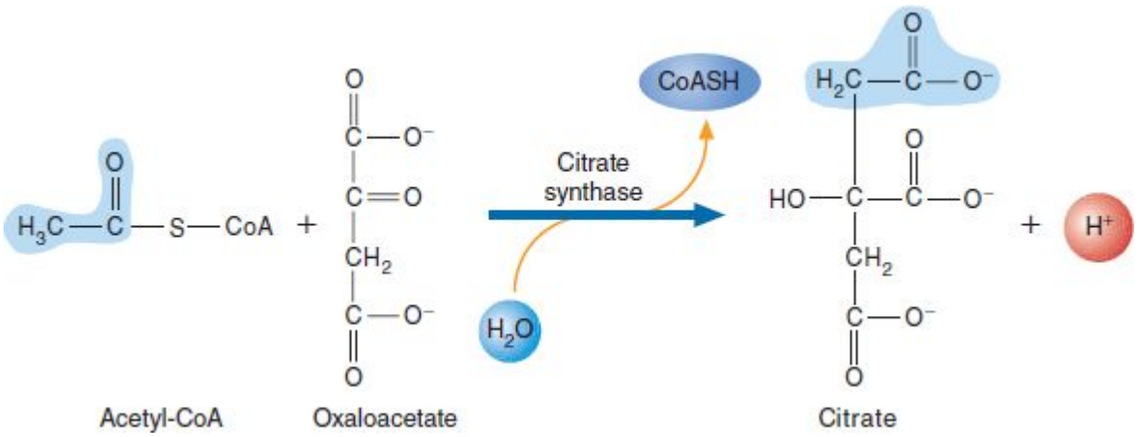
The citric acid cycle is composed of eight reactions that occur in two stages:

1. The two-carbon acetyl group of acetyl-CoA enters the cycle by reacting with the four-carbon compound oxaloacetate (OAA); two molecules of CO_2 are subsequently liberated (reactions 1–4).
2. OAA is regenerated, so it can react with another acetyl-CoA (reactions 5–8).

The enzymes that catalyze these reactions associate through noncovalent interactions into metabolons (p. 232), multienzyme complexes that ensure efficient channeling of the product of each reaction to the next enzyme in the pathway.

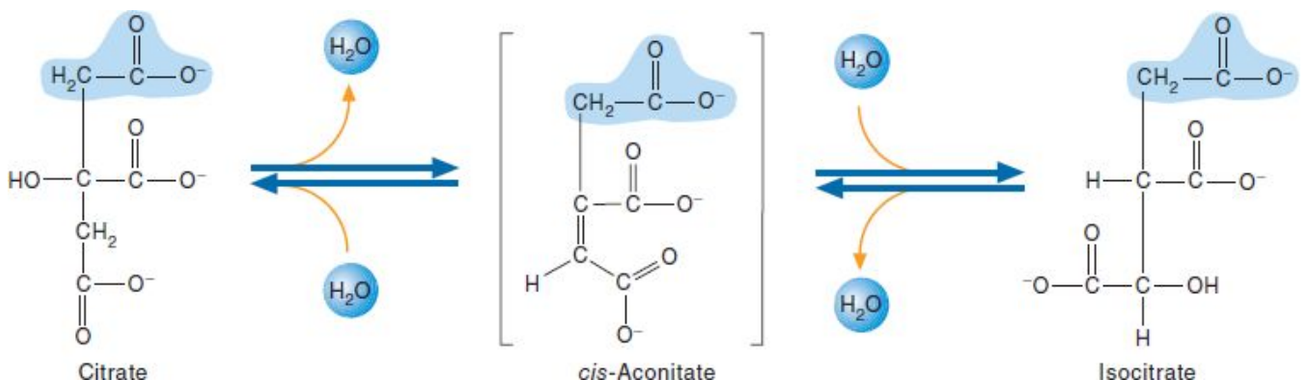
The reactions of the citric acid cycle are as follows.

1. **Introduction of two carbons as acetyl-CoA.** The citric acid cycle begins with the condensation of acetyl-CoA with oxaloacetate to form citrate:



The enzyme citrate synthase is a homodimer in which the active site is a cleft between the two subunits. The binding of oxaloacetate in the open form of the enzyme causes a structural change that induces a transition to the closed form while simultaneously forming the acetyl-CoA binding site. In this reaction (**Figure 9.13**). the enzyme removes a proton from the methyl group of acetyl-CoA, thereby converting it to an enol (p. 290). The enol subsequently attacks the C2 carbonyl carbon of oxaloacetate. The product, citroyl-CoA, rapidly hydrolyzes to form citrate and CoASH. Because of the hydrolysis of the high-energy thioester bond, the overall standard free energy change is −33.5 kJ/mol, and citrate formation is highly exergonic.

2. **Citrate is isomerized to form a secondary alcohol that can be easily oxidized.** In the next reaction of the cycle, citrate, which contains a tertiary alcohol, is reversibly converted to isocitrate by aconitase. The reaction begins with the protonation of the C3 hydroxyl group by a histidine side chain, which promotes the release of a water molecule. A serine residue then abstracts a proton from C2 to form the double bond of the intermediate *cis*-aconitate. After flipping of the intermediate within the active site, the enzyme reverses the steps. It catalyzes the hydration of the double bond when a now basic histidine abstracts a proton from a water molecule, thus triggering a nucleophilic attack on C2, and the protonated serine is deprotonated by the double bond to yield the product isocitrate, a more reactive molecule. Although the standard free energy change of citrate isomerization is positive ($\Delta G^{\circ\prime}$ = 13.3 kJ), the reaction is pulled forward by the rapid removal of isocitrate by the next reaction.



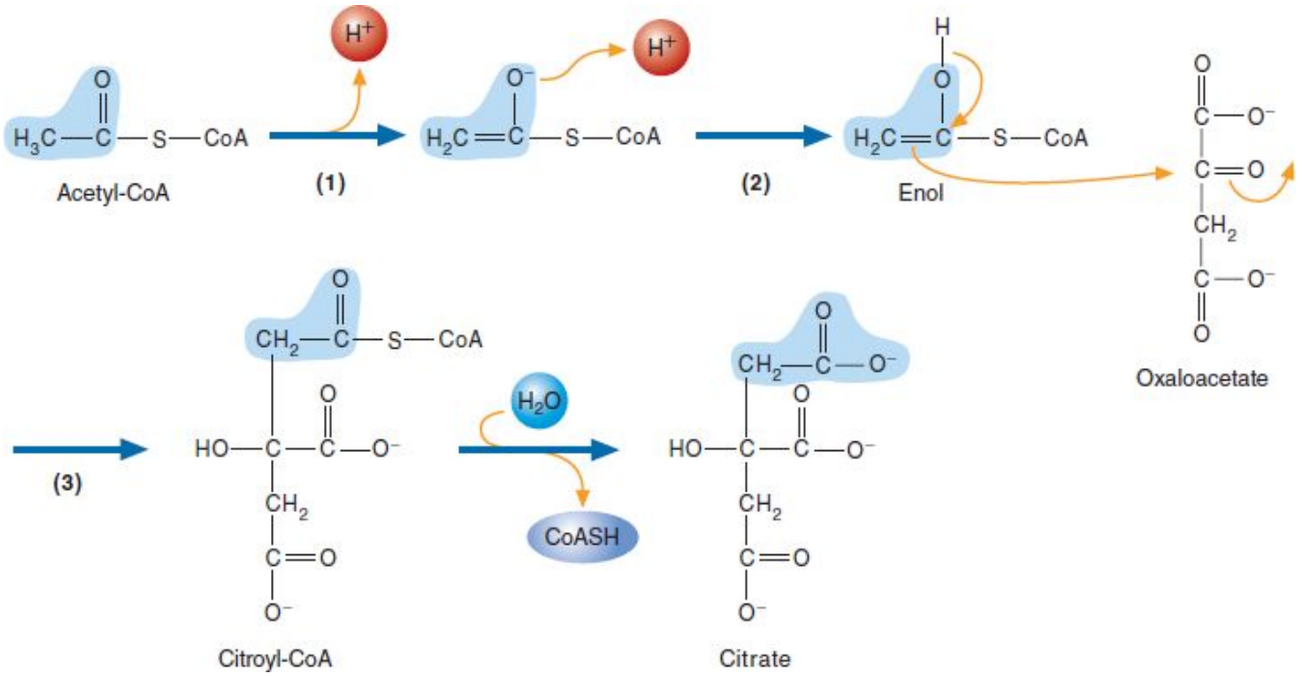
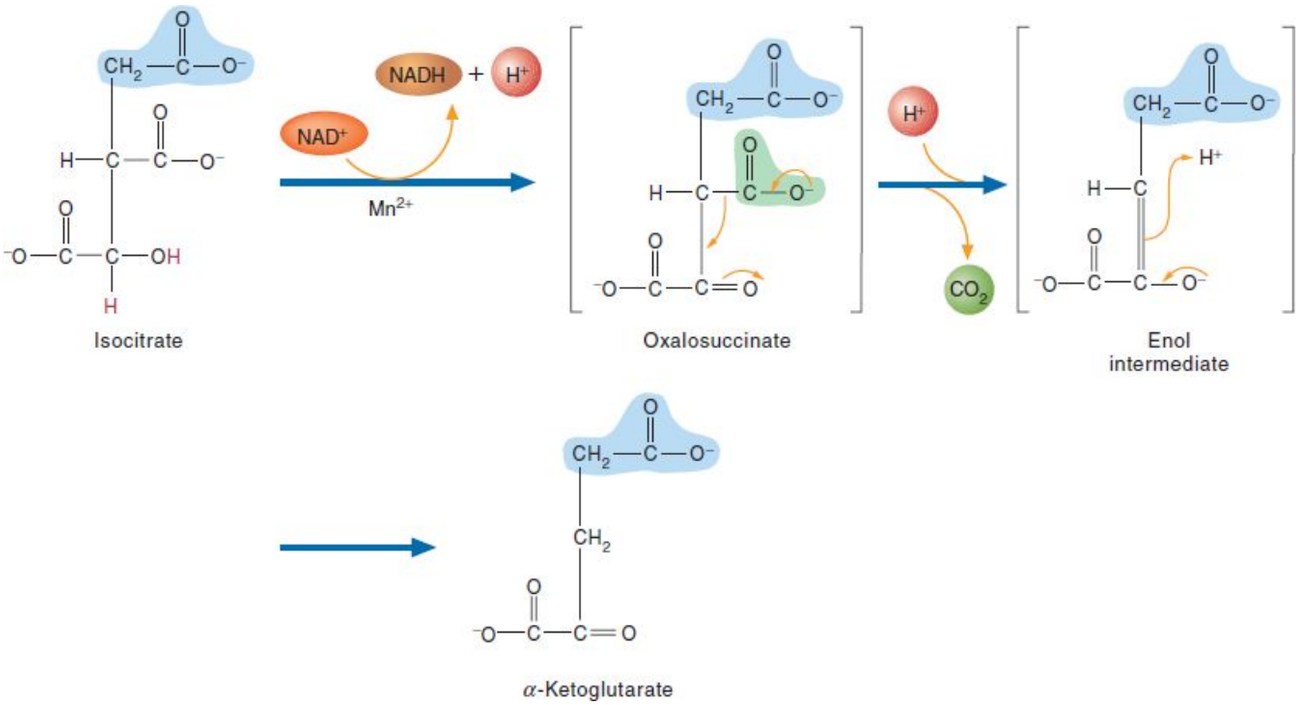


FIGURE 9.13

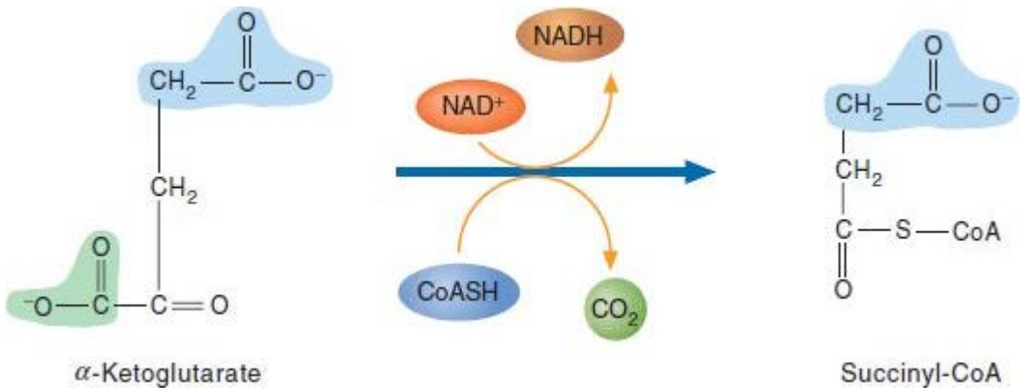
Citrate Synthesis

(1) A side chain carboxylate group (Asp) of the enzyme citrate synthase, acting as a base, removes a proton from the methyl group of acetyl-CoA to form an enol. (2) Simultaneously, a side chain NH group of a histidine residue donates a proton to the carbonyl oxygen, thus generating an enol intermediate. (3) The same histidine side chain then deprotonates the enol to produce an enolate anion that launches a nucleophilic attack on the carbonyl carbon of oxaloacetate. (4) The product, citroyl-CoA, is then hydrolyzed in a nucleophilic acyl substitution reaction when the oxygen of a nearby water molecule (deprotonated by a second histidine side chain) attacks the thioester bond, to yield citrate and CoASH.

3. **Isocitrate is oxidized to form α-ketoglutarate and CO_2.** The oxidative decarboxylation of isocitrate, catalyzed by isocitrate dehydrogenase, proceeds with a $\Delta G^{\circ\prime}$ of –8.4kJ/mol. There are three isoenzyme forms of isocitrate dehydrogenase in mammals. The NAD^+-requiring isozyme (IDH3) is found only within mitochondria. The other isozymes, IDH1 (cytoplasm) and IDH2 (mitochondria), use $NADP^+$ as a cofactor. NADPH is required in biosynthetic processes and quenching of reactive oxygen species (ROS). Note that the NADH produced in the conversion of isocitrate to α-ketoglutarate is the first link between the citric acid cycle and the ETC and oxidative phosphorylation. The reaction begins when the α-carbon (C2) hydroxy group of isocitrate is deprotonated to form the carbonyl group of the intermediate oxalosuccinate. As the carbonyl group forms, a hydride ion is extracted from the α-carbon and then donated to $NAD(P)^+$ to yield NAD(P)H. A manganese (Mn^{2+}) cofactor facilitates the polarization of the newly formed ketone group in oxalosuccinate. The decarboxylation of oxalosuccinate is initiated when a tyrosine side chain group polarizes the C-3 carboxyl group, causing the formation of the enol intermediate (i.e., electrons flow toward the α-carbon to form the double bond between carbons 2 and 3). The enol intermediate rapidly rearranges to form the more stable molecule α-ketoglutarate, an α-keto acid.



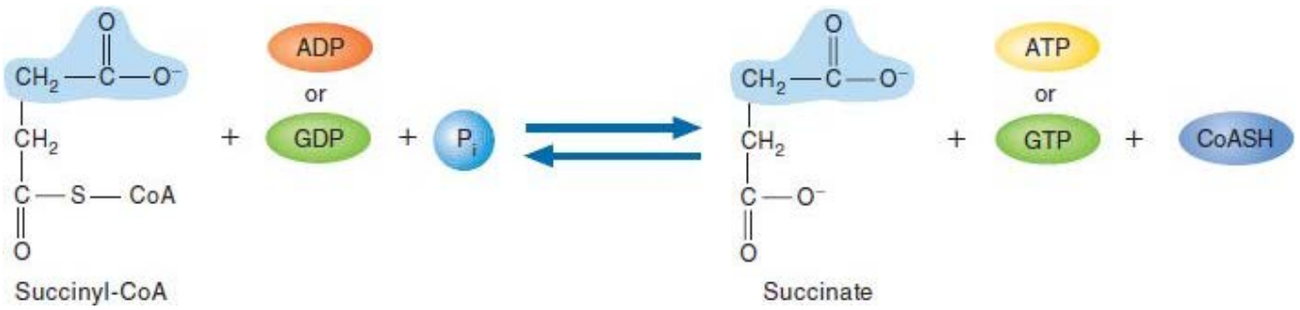
4. ***α*-Ketoglutarate is oxidized to form a second molecule each of NADH and CO_2.** The conversion of *α*-ketoglutarate to succinyl-CoA is catalyzed by the enzyme activities in the *α*-ketoglutarate dehydrogenase complex: *α*-ketoglutarate dehydrogenase, dihydrolipoyl transsuccinylase, and dihydrolipoyl dehydrogenase.



This highly exergonic reaction ($\Delta G^{\circ\prime}$ = −33.5 kJ/mol), an oxidative decarboxylation, is analogous to the conversion of pyruvate to acetyl-CoA catalyzed by the pyruvate dehydrogenase complex. Both reactions have energy-rich thioester molecules as products, that is, acetyl-CoA and succinyl-CoA. Other similarities between the two multienzyme complexes are that the same cofactors (TPP, CoASH, lipoic acid, NAD^+, and FAD) are required, and the same or similar allosteric effectors are inhibitors. *α*-Ketoglutarate dehydrogenase is inhibited by succinyl-CoA, NADH, ATP, and GTP. An important difference between the two complexes is that the control mechanism of the *α*-ketoglutarate dehydrogenase complex does not involve covalent modification.

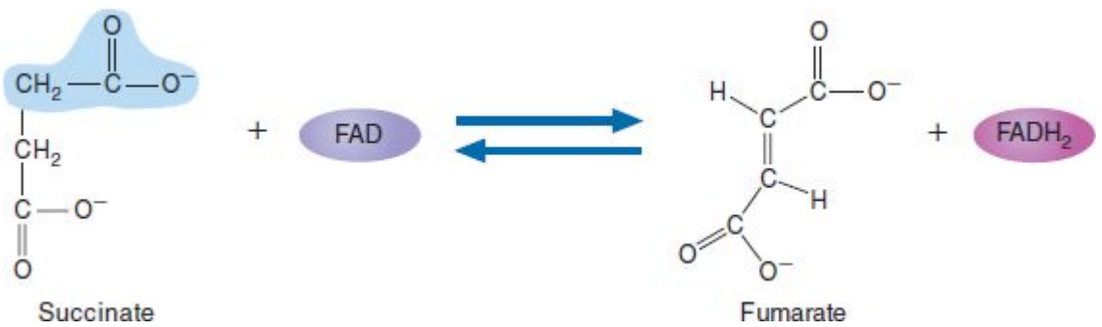
5. **The cleavage of succinyl-CoA is coupled to a substrate-level phosphorylation.** Succinyl-CoA contains a thioester bond with a free energy of hydrolysis value ($\Delta G^{\circ\prime}$) of –36 kJ/mol. The cleavage of the high-energy bond of succinyl-CoA to form succinate is a reversible reaction catalyzed by succinyl-CoA synthetase (also called succinate thiokinase). The reaction begins with the nucleophilic displacement of CoASH to yield succinyl phosphate.

The phosphate is subsequently removed by the enzyme with the formation of a covalent bond between a histidine side chain and the phosphate group. The phosphate group is then transferred to a nucleoside diphosphate. In mammals, the reaction is coupled to the substrate-level phosphorylation of ADP or GDP. Since the substrate level phosphorylation has a $\Delta G°'$ value of +32.2 kJ/mol, the overall reaction is reversible, with a net $\Delta G°'$ value of approximately −3.8 kJ/mol. There are two forms of succinyl-CoA synthetase: one is specific for ATP and the other for GTP. In many tissues, both enzymes are produced, although their relative amounts vary.



The direction of the reaction depends on the relative concentrations of nucleoside diphosphates (ADP and/or GDP) and nucleotide triphosphates (ATP and/or GTP). GTP is used for RNA, DNA, and protein synthesis within mitochondria. The phosphoryl group of GTP can also be donated to ADP in a reversible reaction catalyzed by nucleoside diphosphate kinase.

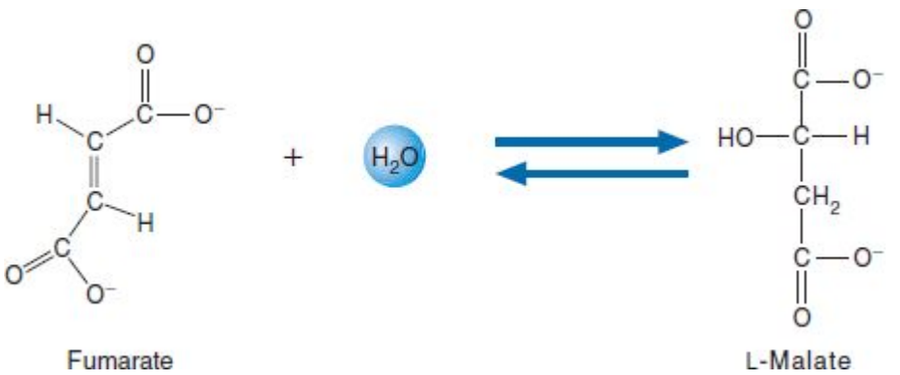
6. **The four-carbon molecule succinate is oxidized to form fumarate and $FADH_2$.** Succinate dehydrogenase catalyzes the reversible oxidation of succinate to form fumarate:



A general base within the enzyme's active site removes a proton from C-2 to form a C-2–C-3 double bond followed by the release of a C-3 hydride to FAD to yield $FADH_2$ and fumarate. Unlike the other citric acid cycle enzymes, succinate dehydrogenase is not found within the mitochondrial matrix. Instead, it is tightly bound to the inner mitochondrial membrane. Succinate dehydrogenase is a flavoprotein that uses FAD to drive the oxidation of succinate to fumarate. (Note that FAD is used instead of NAD^+ because it is a stronger oxidizing agent; in other words, FAD's more positive reduction potential [see Table 9.1] allows the oxidation of a carbon–carbon single bond to form a double bond.) Succinate dehydrogenase is composed of four subunits: (1) ShdA, which contains the succinate binding site and a covalently bound FAD; (2) ShdB, which possesses three iron-sulfur

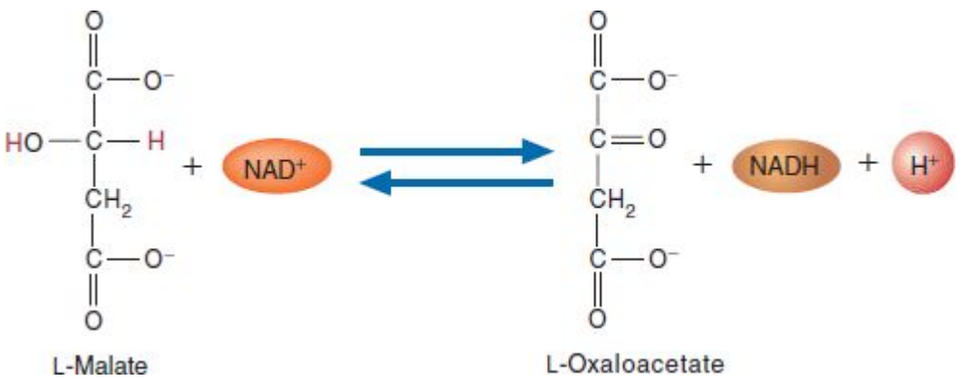
clusters (p. 364) that function as electron carriers between $FADH_2$ and coenzyme Q, a component of the ETC; and (3, 4) subunits ShdC and ShdD, which are hydrophobic molecules that anchor the enzyme complex into the inner membrane. The $\Delta G^{\circ\prime}$ for succinate oxidation is −5.6 kJ/mol. Succinate dehydrogenase is activated by high concentrations of succinate, ADP, and P_i and inhibited by OAA. The enzyme is also inhibited by malonate (see Figure 6.18 on p. 227), a structural analogue of succinate, whose accumulation redirects energy to fatty acid synthesis. Hans Krebs (1900–1981) used this inhibitor in his pioneering work on the citric acid cycle.

7. **Fumarate is hydrated.** Fumarate is converted to L-malate in a reversible stereospecific hydration reaction ($\Delta G^{\circ\prime}$ = −3.8 kJ/mol) catalyzed by fumarase (also referred to as fumarate hydratase):



Fumarase is a lyase that catalyzes a stereospecific hydration. A general base of the enzyme deprotonates a water molecule that then attacks the double bond to form a hydroxyl group at C-2 and a carbanion at C-3. Next, a general acid protonates the carbanion to yield L-malate.

8. **Malate is oxidized to form OAA and a third NADH.** Finally, OAA is regenerated with the oxidation of L-malate by malate dehydrogenase:



A histidine side chain in the active site of the enzyme removes a hydrogen from the hydroxyl group at C-2 of L-malate. Simultaneously, with the formation of the carbonyl group, a hydride ion is transferred to NAD^+ to yield NADH. Malate dehydrogenase uses NAD^+ as the oxidizing agent in a highly endergonic reaction ($\Delta G^{\circ\prime}$ = +29 kJ/mol). The reaction is pulled to completion because of the removal of oxaloacetate in the next round of the cycle. There are two isoenzyme forms of malate dehydrogenase. One is located in the mitochondrial matrix, whereas the other, found in the cytoplasm, is involved in the malate–aspartate shuttle (see Figure 10.21b on p. 382).

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on Hans Krebs and the citric acid cycle.

Fate of Carbon Atoms in the Citric Acid Cycle

In each turn of the citric acid cycle, two carbon atoms enter as the acetyl group of acetyl-CoA, and two molecules of CO_2 are released. A careful review of Figure 9.8 reveals that the two carbon atoms released as CO_2 molecules are not the same two carbons that just entered the cycle. Instead, the released carbon atoms are derived from OAA, which reacted with the incoming acetyl-CoA. The incoming carbon atoms subsequently form one-half of succinate. Because of succinate's symmetric structure, the carbon atoms derived from the incoming acetyl group become evenly distributed in all of the molecules derived from succinate. Consequently, incoming carbon atoms are released as CO_2 only after two or more turns of the cycle.

KEY CONCEPTS

- The citric acid cycle begins with the condensation of a molecule of acetyl-CoA with oxaloacetate to form citrate, which is eventually reconverted to oxaloacetate.
- During this process, two molecules of CO_2, three molecules of NADH, one molecule of $FADH_2$, and one molecule of GTP are produced.

QUESTION 9.5

Trace the labeled carbon in $CH_3{}^{14}C(=O)—SCoA$ through one round of the citric acid cycle. After examining **Figure 9.8**, show why more than two turns of the cycle are required before all the labeled carbon atoms are released as $^{14}CO_2$.

QUESTION 9.6

A mutated IDH1 isoenzyme is found in a high percentage of a type of brain cancer called glioblastoma. Instead of converting isocitrate to α-ketoglutarate, mutated IDH1 converts its substrate to 2-hydroxyglutarate, a circumstance that disrupts the citric acid cycle, among other effects. Review the structures of isocitrate and α-ketoglutarate and determine the structure of 2-hydroxyglutarate.

The Amphibolic Citric Acid Cycle

Amphibolic pathways can function in both anabolic and catabolic processes. The citric acid cycle is obviously catabolic: acetyl groups are oxidized to form CO_2, and energy is conserved in reduced coenzyme molecules. The citric acid cycle is also anabolic, since several citric acid cycle intermediates are precursors in biosynthetic pathways (**Figure 9.14**). For example, OAA is a

gluconeogenesis substrate (Chapter 8) and a precursor in the synthesis of the amino acids aspartate, lysine, threonine, isoleucine, and methionine (Chapter 14). α-Ketoglutarate also plays an important role in amino acid synthesis as a precursor of glutamate, glutamine, proline, and arginine. The synthesis of porphyrins such as heme (**Figure 5.35**) requires succinyl-CoA (Chapter 14). Finally, excess citrate molecules are transported into the cytoplasm, where they are cleaved to form OAA and acetyl-CoA. The latter molecule is used to synthesize fatty acids and steroid molecules such as cholesterol (Chapter 12).

Anabolic processes drain the citric acid cycle of the molecules required to sustain its role in energy generation. Several reactions, referred to as **anaplerotic reactions**, replenish them. One of the most important anaplerotic reactions is catalyzed by pyruvate carboxylase (p. 302). A high concentration of acetyl-CoA, an indicator of an insufficient OAA concentration, activates pyruvate carboxylase. As a result, OAA concentration increases. Other anaplerotic reactions include the synthesis of succinyl-CoA from certain fatty acids (Chapter 12) and the α-keto acids α-ketoglutarate and OAA from the amino acids glutamate and aspartate, respectively, via transamination reactions (Chapter 14).

KEY CONCEPTS

- The citric acid cycle is an amphibolic pathway; that is, it plays a role in both anabolism and catabolism.
- The citric acid cycle intermediates used in anabolic processes are replenished by several anaplerotic reactions.

QUESTION 9.7

Pyruvate carboxylase deficiency, a disease that is usually fatal, is caused when the enzyme that converts pyruvate to OAA is missing or defective. It is characterized by varying degrees of mental retardation and disturbances in several metabolic pathways, especially those involving amino acids and their degradation products. A prominent symptom of this malady is *lactic aciduria* (lactic acid in the urine). After reviewing the function of pyruvate carboxylase, explain why this symptom occurs.

Pyruvate Carboxylase Deficiency

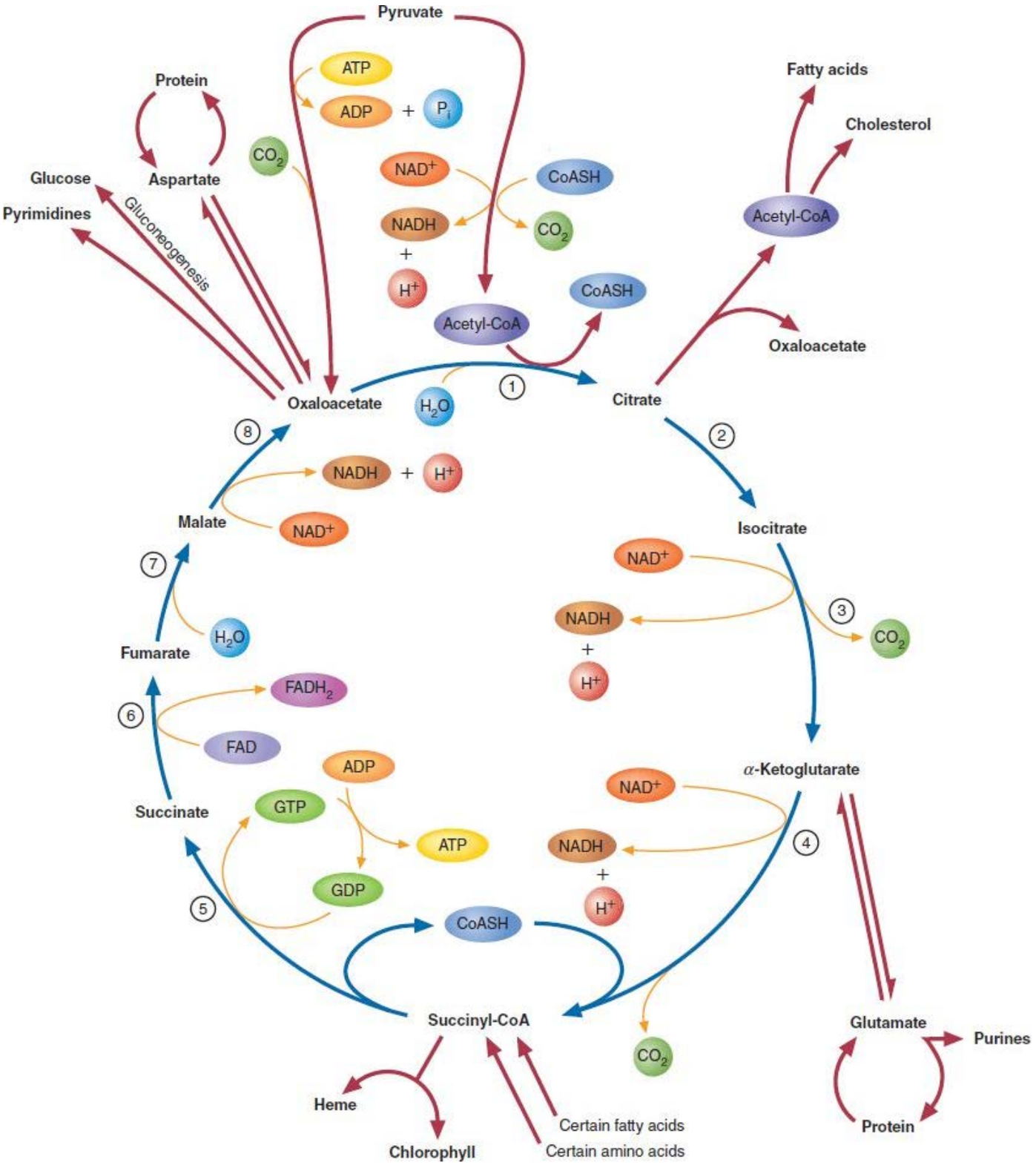


FIGURE 9.14

The Amphibolic Citric Acid Cycle

The citric acid cycle operates in both anabolic processes (e.g., the synthesis of fatty acids, cholesterol, heme, and glucose) and catabolic processes (e.g., amino acid degradation and energy production).

Citric Acid Cycle Regulation

The citric acid cycle is precisely regulated to meet the cell's energy and biosynthetic requirements (**Figure 9.15**). Regulation is achieved via control of three irreversible enzymes within the cycle:

citrate synthase, isocitrate dehydrogenase, and α-ketoglutarate dehydrogenase. These three enzymes operate far from equilibrium (i.e., with highly negative $\Delta G^{\circ\prime}$ values), and they also catalyze reactions that represent important metabolic branch points. Strategies of control include substrate availability, product inhibition, and competitive feedback inhibition. Increased matrix levels of Ca^{2+} also activate all three enzymes.

CITRATE SYNTHASE Citrate synthase, the first enzyme in the cycle, catalyzes the formation of citrate from acetyl-CoA and OAA. The concentrations of acetyl-CoA and OAA are low in mitochondria in relation to the amount of the enzyme. Any increase in substrate availability stimulates citrate synthesis. Citrate synthase is inhibited by its product, citrate. OAA is the product of an endergonic reaction. Therefore, its concentration in mitochondria is quite low relative to malate unless the $NADH/NAD^+$ ratio is low. In many Gram-negative bacteria (*E. coli*, for instance), ATP, NADH, and succinyl-CoA allosterically inhibit citrate synthase.

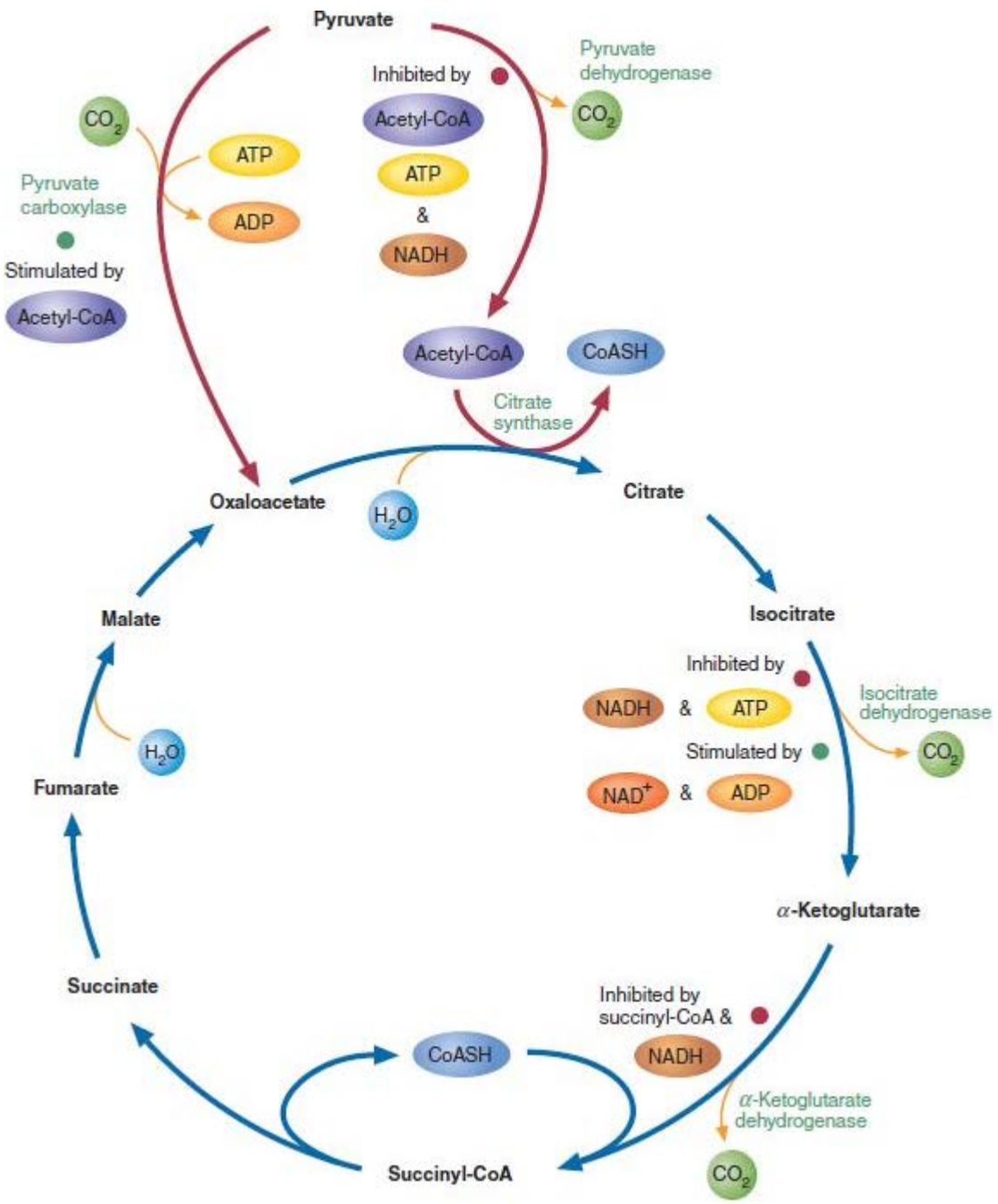


FIGURE 9.15

Control of the Citric Acid Cycle

The major regulatory sites of the cycle are indicated. Activators and inhibitors of regulated enzymes are shown in color: red sphere (inhibition) and green sphere (activation).

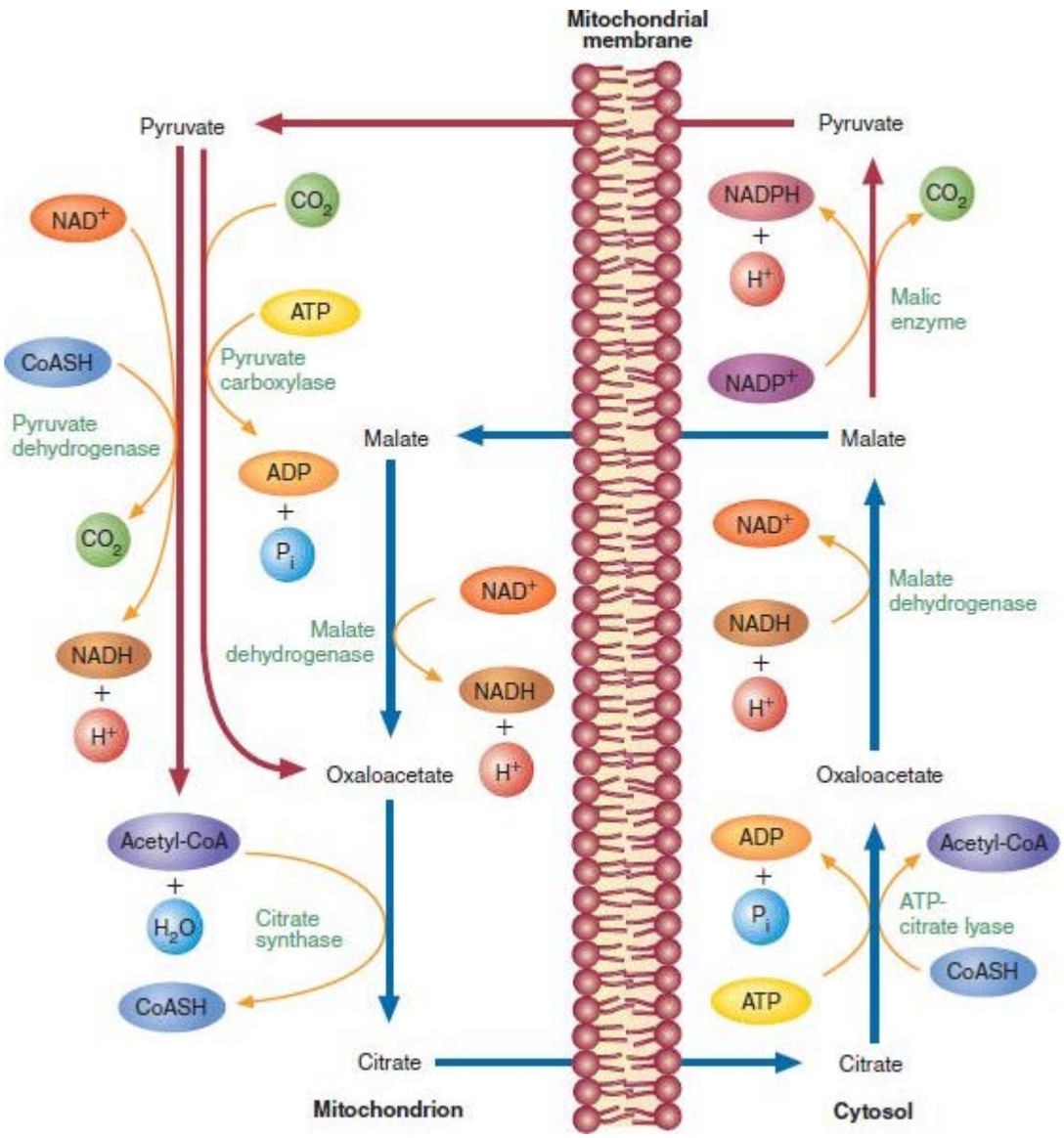


FIGURE 9.16

Citrate Metabolism

When citrate, a citric acid cycle intermediate, moves from the mitochondrial matrix into the cytoplasm, it is cleaved to form acetyl-CoA and oxaloacetate by citrate lyase. The citrate lyase reaction is driven by ATP hydrolysis. Most of the oxaloacetate is reduced to malate by malate dehydrogenase. Acetyl-CoA molecules can be used in biosynthetic pathways such as fatty acid synthesis. Malate may then be oxidized to pyruvate and CO_2 by malic enzyme. The NADPH produced in this reaction is used in cytoplasmic biosynthetic processes, such as fatty acid synthesis. Pyruvate enters the mitochondria, where it may be converted to oxaloacetate or acetyl-CoA. Malate may also reenter the mitochondria, where it is reoxidized to form oxaloacetate.

ISOCITRATE DEHYDROGENASE Isocitrate dehydrogenase's activity is stimulated by relatively

high concentrations of ADP and NAD and inhibited by ATP and NAD(P)H. Isocitrate dehydrogenase is closely regulated because of its important role in citrate metabolism (**Figure 9.16**). As described earlier, the conversion of citrate to isocitrate is reversible. An equilibrium mixture of the two molecules consists largely of citrate. (The reaction is driven forward because isocitrate is rapidly transformed to α-ketoglutarate.) Of the two molecules, only citrate can penetrate the mitochondrial inner membrane. When cellular energy demands are met, excess citrate molecules are transported out of mitochondria and into the cytoplasm. Citrate is then cleaved by ATP-citrate lyase to yield acetyl-CoA and oxaloacetate. The acetyl-CoA formed is used in the synthesis of fatty acids and other lipids. OAA is used in biosynthetic reactions, or it can be converted to malate. Malate either reenters the mitochondrion, where it is reconverted to OAA, or is converted in the cytoplasm to pyruvate by malic enzyme. Pyruvate then reenters the mitochondrion. In addition to being a precursor of acetyl-CoA and OAA in the cytoplasm, citrate also acts directly to regulate several cytoplasmic processes. Citrate is an allosteric activator of the first reaction of fatty acid synthesis. In addition, citrate metabolism provides some of the NADPH used in fatty acid synthesis. Finally, because citrate is an inhibitor of PFK-1, it inhibits glycolysis.

α-KETOGLUTARATE DEHYDROGENASE The activity of α-ketoglutarate dehydrogenase is strictly regulated because of the important role of α-ketoglutarate in several metabolic processes (e.g., amino acid metabolism). When a cell's energy stores are low, α-ketoglutarate dehydrogenase is activated, and α-ketoglutarate is retained within the cycle at the expense of biosynthetic processes. As the cell's supply of NADH rises, the enzyme is inhibited, and α-ketoglutarate molecules become available for biosynthetic reactions. The enzyme is also inhibited by its product succinyl-CoA and activated by AMP, a critical indicator of low energy charge.

Two enzymes outside the citric acid cycle profoundly affect its regulation. The relative activities of PDHC and pyruvate carboxylase determine the degree to which pyruvate is used to generate energy and biosynthetic precursors. For example, if a cell is using a cycle intermediate such as α-ketoglutarate in biosynthesis, the concentration of OAA falls and acetyl-CoA accumulates. Because acetyl-CoA is an activator of pyruvate carboxylase (and an inhibitor of PDHC), more OAA is produced from pyruvate, thus replenishing the cycle.

KEY CONCEPTS

- The citric acid cycle is closely regulated, thus ensuring that the cell's energy and biosynthetic needs are met.
- Allosteric effectors and substrate availability primarily regulate the enzymes citrate synthase, isocitrate dehydrogenase, α-ketoglutarate dehydrogenase, pyruvate dehydrogenase, and pyruvate carboxylase.

CALCIUM REGULATION The signal transduction mechanisms by which cells respond to a variety of stimuli (e.g., hormones, growth factors, and neurotransmitters) often involve transient increases in cytoplasm [Ca^{2+}], followed rapidly by increases in [Ca^{2+}] in the mitochondrial matrix (p. 60). A principal role of Ca^{2+} in the matrix is the stimulation of ATP synthesis by activating the enzymes that regulate the pace of the citric acid cycle. Calcium ions stimulate PDHC activity by activating the dephosphorylating enzyme PDP (p. 342). Both isocitrate dehydrogenase and α-ketoglutarate dehydrogenase are activated directly by Ca^{2+} when the ion binds to a regulatory site on each enzyme. The linkage of the cell's response to a stimulus-driven signal transduction pathway with the uptake of Ca^{2+} into the mitochondrial matrix thus serves to match energy demand with energy production.

The Citric Acid Cycle and Human Disease

Though rarely, several human diseases have been attributed to deficits in citric acid cycle enzymes. Because of the brain's high-energy requirements, the most commonly observed illnesses are severe forms of *encephalopathy* (brain dysfunction characterized by cognitive deficits, tremor, and seizures). For example, encephalopathies have been linked to mutations in the genes that code for α-ketoglutarate dehydrogenase, the A subunit of succinate dehydrogenase, fumarase, and succinyl-CoA synthetase. Several rare cancers are also caused by citric acid cycle enzyme deficits (p. 356). SHB and SHD mutations can cause a pheochromocytoma, an adrenal tumor that secretes excessive amounts of the hormone/neurotransmitter molecules epinephrine and norepinephrine. Symptoms include excessive heart rate and sweating, high blood pressure, and anxiety. A form of renal cell cancer is caused by mutations in fumarase. Several mutated versions of citric acid cycle enzymes contribute to the metabolic disorders in numerous cancers. (Refer to the Biochemistry in Perspective reading on carcinogenesis on p. 356.)

Encephalopathy and Cancers

The Glyoxylate Cycle

Plants and some fungi, algae, protozoans, and bacteria can grow using two-carbon compounds. (Molecules such as ethanol, acetate, and acetyl-CoA, derived from fatty acids, are the most common substrates.) The series of reactions responsible for this capability, referred to as the **glyoxylate cycle**, is a modified version of the citric acid cycle. In plants, the glyoxylate cycle occurs in organelles called glyoxysomes, a type of peroxisome (p. 61) found in germinating seeds. In the absence of photosynthesis, growth in germinating seed is supported by the conversion of oil reserves (triacylglycerol) to carbohydrate. In other eukaryotic organisms and in bacteria, glyoxylate enzymes occur in cytoplasm.

The glyoxylate cycle (**Figure 9.17**) consists of five reactions. The first two reactions (the synthesis of citrate and isocitrate) are familiar because they also occur in the citric acid cycle. However, the formation of citrate from OAA and acetyl-CoA and the isomerization of citrate to form isocitrate are catalyzed by glyoxysome-specific isozymes. The next two reactions are unique to the glyoxylate cycle. Isocitrate is split into two molecules (succinate and glyoxylate) by isocitrate lyase. (This reaction is an aldol cleavage.) Succinate, a four-carbon molecule, is eventually converted to malate by mitochondrial enzymes (**Figure 9.18**). The two-carbon molecule glyoxylate reacts with a second molecule of acetyl-CoA to form malate in a reaction catalyzed by malate synthase. The cycle is completed as malate is converted to OAA by malate dehydrogenase.

The glyoxylate cycle allows for the net synthesis of larger molecules from two-carbon molecules. The decarboxylation reactions of the citric acid cycle, in which two molecules of CO_2 are lost, are bypassed. Using two molecules of acetyl-CoA, the glyoxylate cycle produces one molecule each of succinate and OAA. The succinate product is used in the synthesis of metabolically important molecules such as glucose. The oxaloacetate product is used to sustain the glyoxylate cycle. In organisms, such as animals, that do not possess isocitrate lyase and malate synthase, gluconeogenesis substrates are always molecules with at least three carbon atoms. In these organisms, there is no net synthesis of glucose from fatty acids.

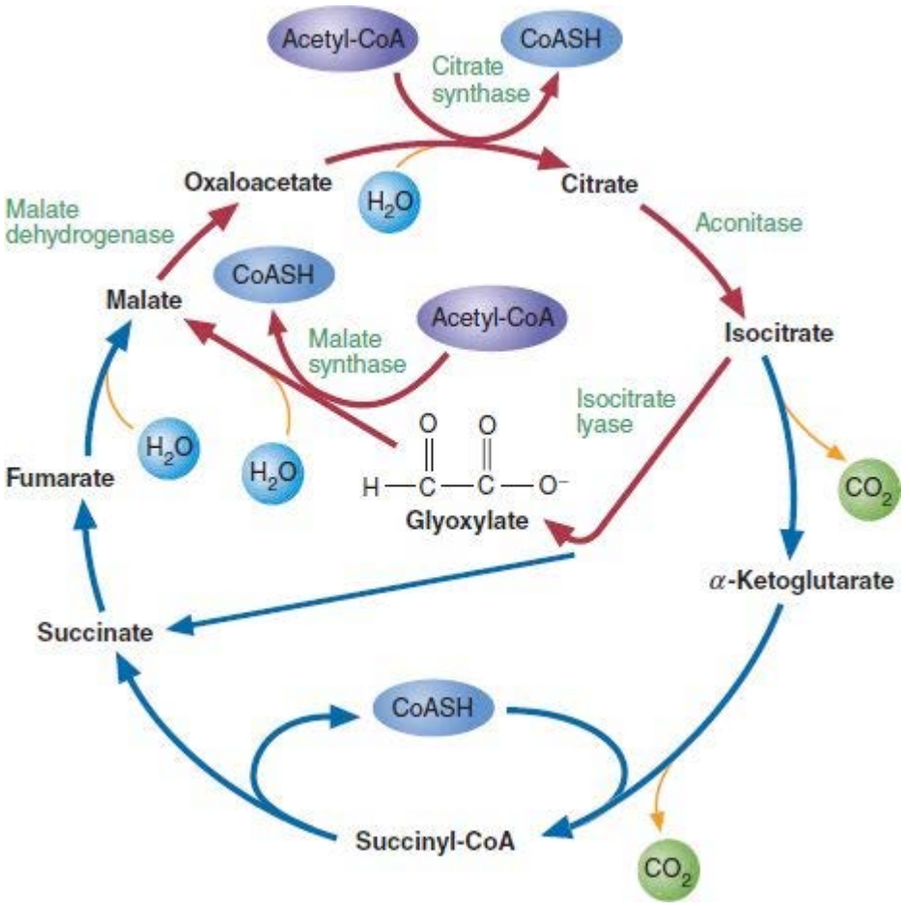


FIGURE 9.17

The Glyoxylate Cycle

Using some of the enzymes of the citric acid cycle, the glyoxylate cycle converts two molecules of acetyl-CoA to one molecule of oxaloacetate. Both decarboxylation reactions of the citric acid cycle are bypassed.

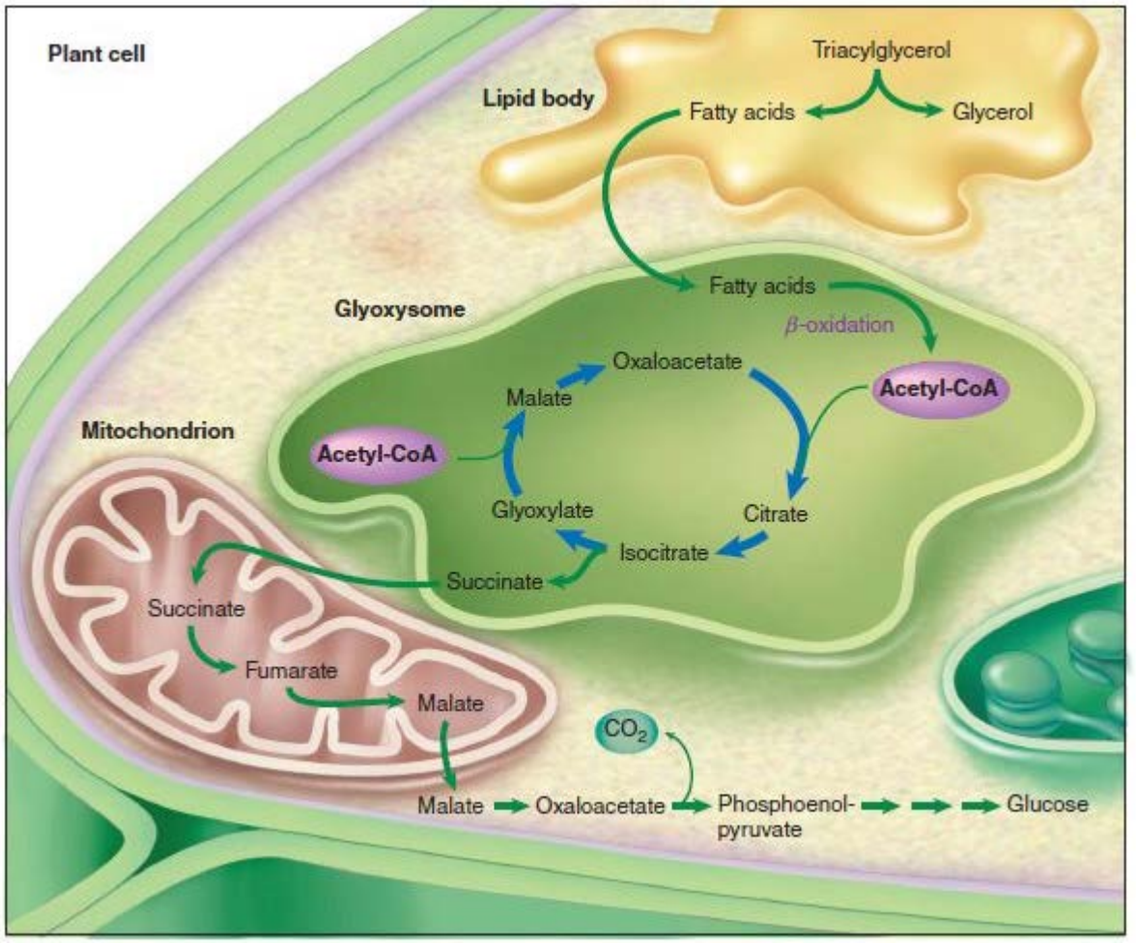


FIGURE 9.18

Role of the Glyoxylate Cycle in Gluconeogenesis

The acetyl-CoA used in the glyoxylate cycle is derived from the breakdown of fatty acids (*β*-oxidation, see Chapter 12). In organisms with the appropriate enzymes, such as some bacteria, glucose can be produced from two-carbon compounds such as ethanol and acetate. In plants, the reactions that convert fatty acids into glucose are localized in lipid bodies, glyoxysomes, mitochondria, and the cytoplasm.

KEY CONCEPTS

- Organisms in which the glyoxylate cycle occurs can use two-carbon molecules to sustain growth.
- In plants, the glyoxylate cycle is found in organelles called glyoxysomes.

Biochemistry IN PERSPECTIVE

Carcinogenesis: The Warburg Effect and Metabolic Reprogramming

What is aerobic glycolysis, and how does it impact cancerous cells? Cancer is a

group of diseases in which there is uncontrolled proliferation of abnormal cells. In the past, cancer research focused on investigations of mutated genes. More recently, cancer has been increasingly recognized as an intricate and heterogeneous set of metabolic disorders. The first abnormal metabolic cancer discovery is credited to Otto Warburg (1883–1970), who reported in 1924 that many tumor cells generate energy principally by glycolysis instead of aerobic respiration, even if O_2 is plentiful. He believed that this exception to the Pasteur effect (p. 301) in cancerous cells, now referred to as the *Warburg effect*, is caused by mitochondrial damage. Although there is mitochondrial damage in some forms of cancer, emerging research has revealed that "aerobic glycolysis" (the conversion of glucose to lactate) is more closely associated with altered cell signaling pathways that cause a reprogramming of metabolic processes. Such alterations give rise to selective advantages to cancerous cells in energy generation and biomolecule synthesis.

Cancerous cells require a large amount of energy and a large supply of metabolic precursor molecules to sustain rapid growth. Although glycolytic ATP synthesis is inefficient, the process of carcinogenesis involves metabolic changes that include markedly increased glucose uptake and glycolytic activity in "transformed" cells. Increased glycolytic flux (10 to 100 times faster than complete glucose oxidation) provides precursors for synthesis of the molecules required for cell division: amino acids, lipids, and nucleic acids. For example, glucose-6-phosphate is a precursor of ribose-5-phosphate; dihydroxyacetone phosphate is converted to glycerol, used in triacylglycerol synthesis; and glycerate-3-phosphate is a precursor of serine, glycine, and cysteine. Cancer cells also use the pentose phosphate pathway, among other processes, to synthesize large quantities of reducing power in the form of NADPH. The evasions of normal regulatory mechanisms that allow cells to rewire their metabolism involve mutations in a diverse group of genes. Prominent examples include c-myc, HIF-1, PKB, and p53. Although many cancers have normal oxidative phosphorylation, increased reliance on aerobic glycolysis in some cells can also be caused by tumor cell mitochondria with mutations in certain citric acid cycle enzymes.

MYC The *myc* gene codes for a transcription factor that regulates about 15% of all genes, including many that are involved in cell division and apoptosis (p. 59). Among myc's target genes are those that code for most of the glycolytic enzymes and several glucose transporters (GLUTs, p. 619). Tumors with mutated myc, called c-myc (representing approximately 70% of all human cancers), have increased lactic acid production and diversion of glycolytic intermediates into pathways that synthesize nucleotides, amino acids, and lipids. C-myc also promotes glutaminolysis, an alternative energy-generating pathway in which glutamine is transported into mitochondria and converted into the citric acid cycle intermediate α-ketoglutarate, which is subsequently converted to citrate, the source of acetyl CoA for fatty acid synthesis (p. 461). Cells also use IDH1 to synthesize NADPH. The glutamine transporter, glutaminase (the enzyme that converts glutamine to glutamate), LDHA (the gene that codes for the M subunit of lactate dehydrogenase), and the lactate plasma membrane transporter are a few of the molecules upregulated by c-myc.

HIF-1 HIF-1 (hypoxia inducible factor 1) is a transcription factor that responds to decreases in cellular oxygen levels by inducing the transcription of genes that promote survival under low-oxygen conditions. The most prominent examples are genes for glycolytic enzymes, glucose transporters, and PDK1, the enzyme that inactivates pyruvate dehydrogenase (p. 342), thereby diverting pyruvate away from mitochondrial oxidation. When oxygen is available, HIF-1 is inactivated by oxygen- and α-ketoglutarate-dependent prolyl hydroxylation. HIF-1 is stabilized in some tumors by mutations in succinate dehydrogenase and fumarase. Succinate or fumarate accumulations result in leakage of these molecules into cytoplasm, where they act as competitive

inhibitors of prolyl hydroxylase. In addition, mutated versions of IDH1 or IDH2 can result in the synthesis of 2-hydroxyglutarate, also an inhibitor of prolyl hydroxylase.

PKB Protein kinase B (PKB) (p. 612), also referred to as Akt, is a serine/threonine protein kinase that regulates glucose metabolism, among several other processes. Activated PKB enhances glucose uptake and glycolysis by inducing increased glucose transporter insertion into the plasma membrane, hexokinase activation, and PFK-2-dependent activation of PFK-1 (p. 298). PKB also promotes the diversion of glucose carbon into the biosynthetic pathways that produce fatty acids and cholesterol and other lipids required for rapid growth. PKB facilitates transport of pyruvate into mitochondria and its rapid conversion into acetyl-CoA. The subsequent increase in citrate results in its export into the cytoplasm where acetyl-CoA, the substrate for lipid synthesis, is regenerated by the PKB-activated enzyme ATP-citrate lyase. Mutated versions of PKB that cause it to be hyperactivated cause some malignancies, while those that decrease its activity cause others.

P53 With more than 100 target genes, p53 is one of the most important tumor suppressors. In addition to its well-known roles in DNA repair, cell cycle regulation, and apoptosis induction, p53 is now known as an important regulator of energy metabolism. P53 coordinates the regulation of glycolysis and oxidative phosphorylation. It does so by maintaining oxidative phosphorylation and down-regulating glycolysis (via decreases in glucose transporters and glycolytic enzymes). P53-induced regulation of energy metabolism is exerted through inducing the synthesis and/or activation of enzymes such as PKB, mTOR (p. 612), and TIGAR (TP53-induced glycolysis and apoptosis regulator), which decreases glycolytic flux by lowering the intracellular concentrations of fructose 2,6-bisphosphate (the PFK-1 allosteric activator). Aerobic glycolysis is one of many consequences of p53 inactivation.

SUMMARY Aerobic glycolysis is a process in tumor cells in which rapid glycolysis-generated ATP synthesis occurs even when O_2 is present. Loss of glycolysis regulation is one facet of carcinogenesis, the set of mechanisms whereby normal cells are gradually transformed into cancerous cells that are no longer responsive to the body's regulatory signals.

Chapter Summary

1. Aerobic organisms have an enormous advantage over anaerobic organisms, that is, a greater capacity to obtain energy from organic food molecules. To use oxygen to generate energy requires the following biochemical pathways: the citric acid cycle, the electron transport pathway, and oxidative phosphorylation.
2. Most reactions that capture or release energy are redox reactions. In these reactions, electrons are transferred between an electron donor (reducing agent) and an electron acceptor (oxidizing agent). In most biochemical reactions, hydride ions are transferred to NAD^+ or $NADP^+$ or hydrogen atoms are transferred to FAD or FMN. A substance's tendency to gain electron(s) is called its reduction potential. Electrons flow spontaneously from a substance with a less positive (more negative) reduction potential to a substance with a more positive (less negative) reduction potential. In favorable redox reactions $\Delta E^{\circ\prime}$ is positive and $\Delta G^{\circ\prime}$ is negative.
3. The citric acid cycle is a series of biochemical reactions that eventually completely oxidize organic substrates, such as glucose and fatty acids, to form CO_2, H_2O, and the reduced coenzymes NADH and

$FADH_2$. Pyruvate, the product of the glycolytic pathway, is converted to acetyl-CoA, the citric acid cycle substrate.

4. In addition to its role in energy generation, the citric acid cycle plays an important role in several biosynthetic processes, such as gluconeogenesis, amino acid synthesis, and porphyrin synthesis.
5. The glyoxylate cycle, found in plants and some fungi, algae, protozoans, and bacteria, is a modified version of the citric acid cycle in which two-carbon molecules, such as acetate, are converted to precursors of glucose.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on the citric acid cycle to help you prepare for exams.

Chapter 9 Review Quiz

Suggested Readings

Anderson NM, et al. 2017. The emerging role and targetability of the TCA cycle in cancer metabolism. Protein Cell doi:10.1007/s13238-017-0451-1.

Bilgen T. 2004. Metabolic evolution and the origin of life. In: Storey KB, editor. Functional metabolism: regulation and adaptation. Hoboken (NJ): Wiley-Liss. p. 557–82.

Colombini M. 2012. VDAC structure, selectivity and dynamics. Biochim Biophys Acta 1818(6):1457–65.

Kornberg H. 2000. Krebs and his trinity of cycles. Nat Rev Mol Cell Biol 1(3):225–7.

Krebs HA. 1970. The history of the tricarboxylic cycle. Perspect Biol Med 14:154–70.

Lane N. 2002. Oxygen: the molecule that made the world. Oxford (UK): Oxford University Press.

Lane N. 2015. Power, sex and suicide and the meaning of life. Oxford (UK): Oxford University Press.

Mailloux RJ. 2015. Still at the center of it all: novel functions of the oxidative Krebs cycle. Bioenergetics 4:122. doi:10.4172/2167-7662.1000122.

Potter M., et al. 2016. The Warburg effect: 80 years on. Biochem Soc Trans 44:1499–505.

Schmitt DL, An S. 2017. Spatial organization of metabolic enzyme complexes in cells. Biochem 56:3184–96.

Key Words

aerotolerant anaerobe, *331*

amphibolic pathway, *349*

anaplerotic, *349*

coenzyme A, *339*

conjugate redox pair, *333*

facultative anaerobe, *331*

flavin adenine dinucleotide (FAD), *335*

flavin mononucleotide (FMN), *335*
flavoprotein, *335*
glyoxylate cycle, *354*
lipoic acid, *342*
nicotinamide adenine dinucleotide (NAD), *335*
nicotinamide adenine dinucleotide phosphate (NADP), *335*
obligate aerobe, *331*
obligate anaerobe, *331*
reduction potential, *333*
standard reduction potential, *333*
thiamine pyrophosphate, *342*

Review Questions

SECTION 9.1

Comprehension Questions

1. Define the following terms:
    a. obligate anaerobe
    b. aerotolerant anaerobe
    c. facultative anaerobe
    d. obligate aerobe
    e. reactive oxygen species
2. Define the following terms:
    a. electron transport chain
    b. oxidation-reduction reactions
    c. conjugate redox pair
    d. reduction potential
    e. standard reduction potential
3. Define the following terms:
    a. aerobic metabolism
    b. aerobic respiration
    c. $\Delta G^{\circ\prime}$
    d. $\Delta E^{\circ\prime}$
    e. F
4. Define the following terms:
    a. FMN
    b. FAD
    c. NAD^+
    d. ETC
    e. salt bridge

Fill in the Blanks

5. The tendency for a substance to gain electrons is called the _____ potential.
6. _____ in the outer mitochondrial membrane allows the diffusion of small hydrophilic molecules.
7. Together an electron donor and an electron acceptor in a reaction is called a _____ .
8. A significant portion of the energy captured as electrons pass down the electron transport chain is used to synthesize _____.

Short-Answer Questions

9. Describe in general terms how the appearance of molecular oxygen in Earth's atmosphere about 3 billion years ago affected the history of living organisms.
10. List the biochemical processes required to obtain energy from glucose using O_2 as an electron acceptor.
11. Calculate the free energy changes that occur in the following reactions:

    a. $\frac{1}{2}\ O_2 + NADH + H^+ \rightarrow H_20 + NAD^+$

    b. $S + NADH + H^+ \rightarrow H_2S + NAD^+$

12. Referring to Question 11, calculate the difference in energy produced by the oxidation of NADH by O_2 and sulfur.

Critical-Thinking Questions

13. Determine the standard free energy ($\Delta G°'$) for the following reactions:

    a. $NADH + H^+ + \frac{1}{2}\ O_2 \rightarrow NAD^+ + H_2O$

    b. Cytochrome c (Fe^{2+}) + $\frac{1}{2}\ O_2$ → cytochrome c (Fe^{3+}) + H_2O

14. Using the data in Table 9.1 and the equation $\Delta G°' = -nF\Delta E°'$, calculate the free energy ($\Delta G°'$) produced by the reduction of sulfur to hydrogen sulfide and oxygen to water by NADH. How much more free energy is produced by the reduction of oxygen compared to that of sulfur?
15. Calculate the standard free energy change ($\Delta G°'$) for the following reaction:

    $\frac{1}{2}\ O_2 + FADH_2 \rightarrow H_2O + FAD$

16. The value of the standard reduction potential ($E°'$) for the redox half-reaction

    $Acetaldehyde + 2H^+ + 2e^- \rightarrow Ethanol$

    is −0.20 V. Calculate the free energy ($\Delta G°'$) for the following reaction:

    $Ethanol + NAD^+ \rightarrow Acetaldehyde + NADH + H^+$

SECTION 9.2

Comprehension Questions

17. Define the following terms:

    a. coenzyme A

b. TPP
c. lipoic acid
d. PDHC
e. PDK1

18. Define the following terms:

a. HETPP
b. nucleoside diphosphate kinase
c. amphibolic pathway
d. anaplerotic reaction
e. citric acid cycle

19. Define the following terms:

a. vitamin B_1
b. PDP
c. OAA
d. glyoxylate cycle
e. electron acceptor

Fill in the Blanks

20. _____ pathways can function in both anabolic and catabolic processes.
21. In eukaryotic cells, the citric acid cycle occurs in the _____.
22. The most pivotal event in the history of life on Earth, allowing for the emergence of eukaryotic cells, was the _____.
23. _____ produced in the citric acid cycle is an important carbon source for gluconeogenesis.
24. _____ reactions replenish the intermediate molecules of the citric acid cycle.
25. The acetyl group of acetyl-CoA is derived from the breakdown of lipids, carbohydrates, and certain _____.
26. Pyruvate is converted to acetyl-CoA by the enzymes of the _____ enzyme complex.
27. _____, also referred to as Akt, is a serine/threonine kinase that regulates glucose metabolism.

Short-Answer Questions

28. How do calcium ions regulate the citric acid cycle?
29. How do germinating seeds convert their triacylglycerol reserves to the glucose molecules required in the synthesis of complex carbohydrate such as cellulose?
30. How does pyruvate carboxylase deficiency result in lactic aciduria, an illness in which lactate appears in the urine?
31. Give examples of biochemical pathways that utilize citric acid cycle intermediates as precursor molecules.
32. The citric acid cycle operates only when O_2 is present, yet O_2 is not a substrate for the cycle. Explain.
33. Describe in detail the structure of the pyruvate dehydrogenase complex and the role of each enzyme, cofactor, and coenzyme.
34. If a small amount of [1–^{14}C] glucose is added to an aerobic yeast culture, where will the ^{14}C

label initially appear in citrate molecules?

35. Provide examples of molecules derived from citric acid intermediates.
36. Write the net equation for the citric acid cycle.
37. Discuss the mechanisms of control of the irreversible steps in the citric acid cycle.
38. Write balanced equations for each reaction in the citric acid cycle.
39. What coenzymes are required in each of the reactions in Question 38?
40. Each of the following molecules has a role in carcinogenic aerobic glycolysis: c-myc, HIF-1, PKB, and p53. Describe the function of each molecule.
41. Explain why animals cannot produce glucose from two carbon molecules, such as acetate or ethanol.
42. What steps of the citric acid cycle are regulated? How and why are they regulated?
43. The redox half reaction

    $$\text{Oxaloacetate} + 2H^+ + 2e^- \rightarrow \text{Malate}$$

    has a standard reduction potential value of 0.166 V. Calculate the free energy ($\Delta G°'$) for the following reaction:

    $$\text{Malate} + NAD^+ \rightarrow \text{Oxaloacetate} + NADH + H^+$$

44. Using the data in Table 9.1, calculate the free energy ($\Delta G°'$) for the synthesis of NADH in the conversion of isocitrate to α-ketoglutarate in the citric acid cycle.
45. Malonate (p. 227) poisons the citric acid cycle because it inhibits succinate dehydrogenase. After reviewing its structure, describe how the inhibitory effect of malonate can be overcome.

Critical-Thinking Questions

46. One consequence of ethanol addiction is fatty liver disease, an illness in which liver cells accumulate large amounts of triacylglycerols, the esters derived from glycerol and fatty acids. Ethanol is oxidized in the cytoplasm of liver cells by alcohol dehydrogenase and aldehyde dehydrogenase to yield acetate and 2 NADH. Acetate is then transported into the mitochondrion, where it is converted to acetyl-CoA and metabolized in the citric acid cycle. When alcohol is consumed in excessive quantities, the resulting high levels of NADH cause metabolic abnormalities, one of which is high levels of fatty acid synthesis. Fatty acid synthesis, also a cytoplasmic process, uses acetyl-CoA as a substrate and NADPH as a reducing agent. Speculate about how a high level of cytoplasmic NADH provides a source of NADPH for fatty acid synthesis.
47. Despite the absence of the glyoxylate cycle in mammals, when ^{14}C-labeled acetate is fed to lab animals, small amounts of the radioactive label later appear in glycogen stores. Explain.
48. One effect of chronic alcohol abuse is thiamine deficiency, caused by impaired absorption of the vitamin through the intestinal wall and diminished storage in a damaged liver. When thiamine levels are inadequate, cellular energy generation is diminished. List three enzymes involved in cellular metabolism that require thiamine. Describe the possible metabolic consequences of inadequate thiamine levels.
49. Pyruvate dehydrogenase deficiency is a fatal disease usually diagnosed in children. Symptoms include severe neurological damage. Elevated blood levels of lactate, pyruvate, and alanine are also seen. Explain how the deficiency of pyruvate dehydrogenase causes these elevated values.
50. The plant toxin fluoroacetate ($F\text{-}CH_2COO^-$) is easily converted to fluorocitrate when an animal ingests the plant. The enzyme aconitase has a high affinity for the 2-fluorocitrate molecule. Consider the purpose of this enzyme and speculate as to why the reaction to produce isocitrate

does not occur as planned. Why is fluoroacetate considered a poison? [*Hint*: Fluorine is more electronegative than oxygen.]

51. Systemic fungal infections (*Candida* and *Cryptococcus*) and tuberculosis (*Mycobacterium tuberculosis*) are on the rise. These organisms have high virulence in part because they fare better following macrophage phagocytosis than other microorganisms. Phagocytosis activates the glyoxylate cycle in fungi and mycobacteria, allowing them to use two-carbon substrates to sustain growth in an otherwise nutrient-poor environment of the phagolysosome. Suggest some likely molecules that could be processed through the glyoxylate cycle in this circumstance
52. What is the significance of substrate-level phosphorylation reactions? Which of the reactions in the citric acid cycle involve a substrate-level phosphorylation? Name another example from a biochemical pathway with which you are familiar.
53. You have just consumed a piece of fruit. Trace the carbon atoms in the glucose in the fruit through the biochemical pathways between their uptake into tissue cells and their conversion to CO_2.
54. Fatty acid degradation stimulates the citric acid cycle through the activation of pyruvate carboxylase by acetyl-CoA. Why would the activation of pyruvate carboxylase increase energy generation from fatty acids?
55. Dichloroacetate, which inhibits the enzyme pyruvate dehydrogenase kinase, has been used with limited results to treat lactic acidosis. The phosphorylation of the α-subunit of the pyruvate dehydrogenase component of the PDHC by pyruvate dehydrogenase kinase causes complete loss of enzymatic activity. Describe the hypothesis behind the clinical use of dichloroacetate.
56. Shock (failure of the circulatory system) is an abnormal condition in which blood flow is inadequate. The most common causes of shock are massive blood loss and obstruction of blood flow. As a result of shock, there is a failure to provide cells with oxygen and nutrients. In this circumstance, cells swell and lysosomal membranes rupture, among other effects. Describe how energy is generated during shock and why cell structure becomes destabilized.
57. Lactic acidosis occurs as a result of shock. Explain why low oxygen levels promote lactate production.
58. The large amount of energy used during aerobic exercise (e.g., running) requires large amounts of oxaloacetate. Explain why acetyl-CoA cannot be used to produce oxaloacetate in this circumstance. What is the likely source of oxaloacetate molecules during aerobic activity?

MCAT Study Questions

59. In which of the following reactions is the reactant oxidized?

    a. ATP → AMP
    b. NAD^+ → NADH
    c. NADH → NAD^+
    d. acetaldehyde → ethanol

60. Which of the following statements is true when glucose is oxidized in the presence of O_2?

    a. All ATP synthesis requires the presence of O_2.
    b. O_2 is the final electron acceptor.
    c. No water molecules are formed.
    d. The electron transport chain is responsible for all ATP synthesis.

61. Which of the following is true concerning the function of lipoic acid in pyruvate metabolism?

a. transfer of acetyl groups
b. pyruvate decarboxylation
c. conversion of pyruvate to lactate
d. oxidation of $FADH_2$ to yield NADH

62. Which of the following is not a product of the citric acid cycle?

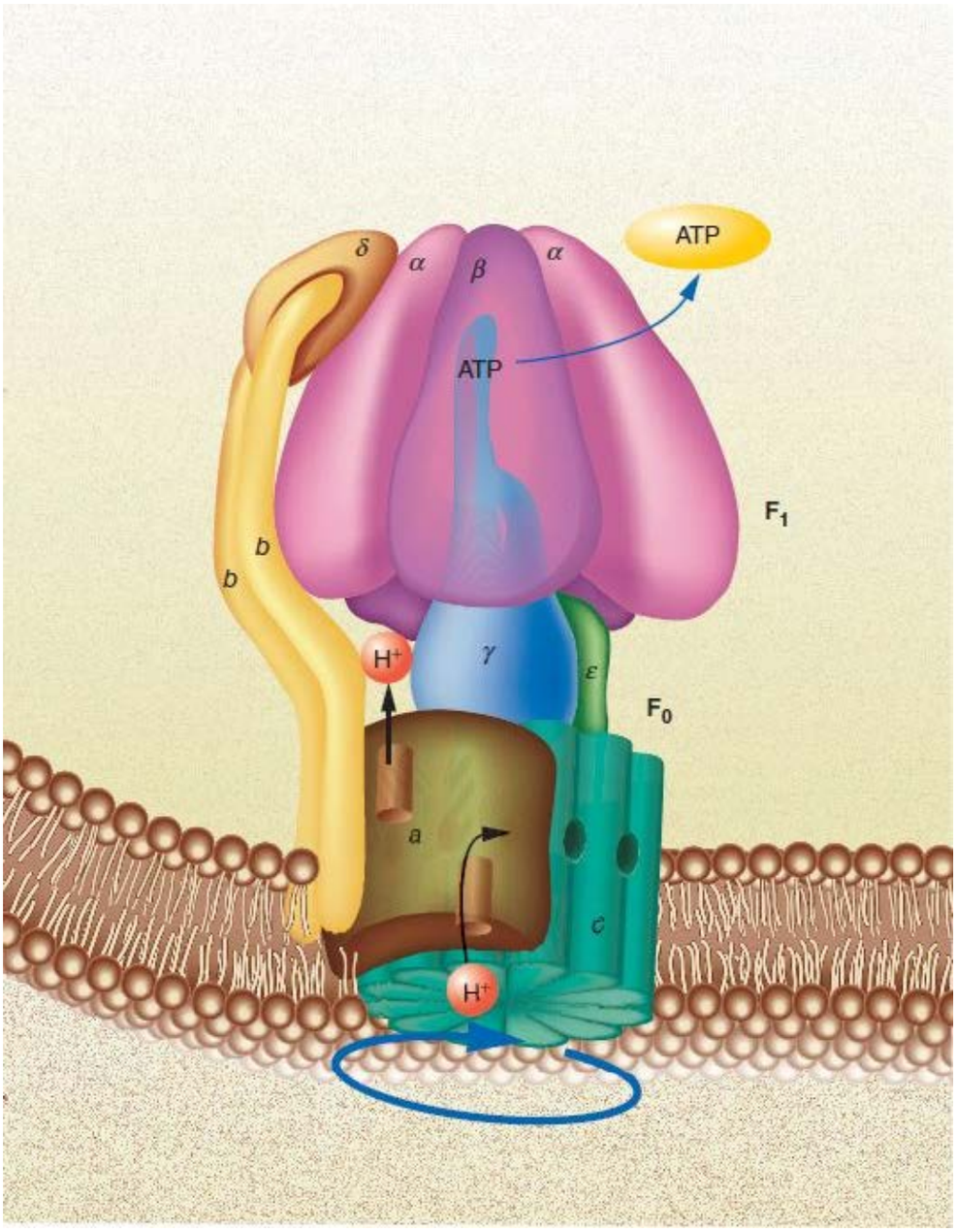
a. $FADH_2$
b. NADH
c. AMP
d. CO_2

63. The coenzymes required for the activity of the pyruvate dehydrogenase complex are

a. biotin, thiamine pyrophosphate, CoASH, and lipoic acid
b. lipoic acid, NAD^+, thiamine pyrophosphate and FAD
c. NAD^+, lipoic acid, tetrahydrofolate, CoASH
d. biotin, tetrahydrofolate, vitamin B6

CHAPTER 10

Aerobic Metabolism II: Electron Transport and Oxidative Phosphorylation



The ATP Synthase The ATP synthase is a rotating molecular machine that synthesizes ATP. This multiprotein complex is composed of two principal domains: the membrane-spanning F_0 component and the ATP-synthesizing F_1 component. The flow of protons through F_0, made possible by a gradient created by electron transport, generates a torque that forces the shaft (the γ subunit) to rotate. Rotational force within F_1 then triggers conformational changes that result in ATP synthesis.

OUTLINE

OXYGEN: A MOLECULAR PARADOX

10.1 ELECTRON TRANSPORT

Electron Transport and Its Components
Electron Transport: The Fluid State and Solid State Models
Electron Transport Inhibitors

10.2 OXIDATIVE PHOSPHORYLATION

The Chemiosmotic Theory
ATP Synthesis

Control of Oxidative Phosphorylation
Creatine Kinase and Creatine
The Complete Oxidation of Glucose
Uncoupled Electron Transport

10.3 OXYGEN, CELL SIGNALING, AND OXIDATIVE STRESS
Reactive Oxygen Species
Antioxidant Enzyme Systems
The Redox Proteome
Antioxidant Nutrients

Biochemistry in Perspective
Myocardial Infarct: Ischemia and Reperfusion

AVAILABLE ONLINE
Biochemistry in Perspective
Glucose-6-Phosphate Dehydrogenase Deficiency

Oxygen: A Molecular Paradox

Oxygen is a very dangerous molecule! It is a *diradical* (i.e., it has two unpaired electrons) that combines with most other elements to form unstable and highly reactive intermediates. As such, it can cause significant damage: Oxygen's reaction with iron to form rust, for example, costs the U.S. economy several tens of billions of dollars every year. But oxygen is also extremely useful. It is such a powerful oxidizing agent that liquefied oxygen and hydrogen served as fuel in the main engines of the Space Shuttle. Similarly, hydrocarbon fuel combustion is an oxygen-requiring process that under controlled conditions is used to heat homes and move vehicles. Considering oxygen's dangerous properties, why can living organisms use it to generate energy? The answer to this question begins with a comparison between combustion and aerobic respiration.

Combustion

The reaction of oxygen with hydrocarbons (RH) is a rapid and extremely exothermic radical chain reaction. This reaction is controllable because of an energy barrier that prevents the spontaneous oxidation of most organic molecules when they are in contact with air. Combustion begins only after the introduction of an energy source (e.g., a lighted match). The chain reaction begins with the abstraction of a hydrogen atom by the oxygen diradical to form a hydroxyperoxide radical (HOO·) and an alkyl radical (R·), both of which proceed to react with other hydrocarbon molecules. The combustion process generates heat as chemical bonds break. The chain reaction accelerates as newly formed radicals react with other fuel molecules. Other radicals formed during the chain reaction include ROO·, RO·, and HO·. Because radicals react with the first molecule in their path, the process is unstoppable until either the hydrocarbon fuel is expended or oxygen is excluded from the reaction. Complete combustion of hydrocarbons produces CO_2 and H_2O.

Aerobic Respiration: Controlled Combustion and the Oxygen Paradox

Aerobic respiration is similar to hydrocarbon combustion in that organic molecules (carbohydrates and fats) are converted into CO_2 and H_2O. Instead of releasing energy in a rapid burst of flames, however, aerobic organisms have evolved a rigorously controlled step-by-step process that extracts hydrogen atoms from fuel molecules while simultaneously protecting themselves from ROS (Section 10.3). As with hydrocarbon combustion, the diradical molecule O_2 accepts electrons one at a time: $O_2 + 4e^- + 4H^+ \rightarrow 2H_2O$. Released

energy is stored in ATP molecules. Although aerobic organisms usually keep ROS formation to a minimum, reactive intermediates can leak out and react with cell constituents. Despite antioxidant enzyme systems and repair processes, eventually the damage becomes so great that function is compromised. Oxidative damage has been linked to the aging process and numerous degenerative diseases. So this is the paradox of oxygen: aerobic organisms use oxygen to generate the vast amounts of energy required to maintain their metabolic processes, as they risk damage caused by this highly reactive molecule.

Overview

THE AEROBIC LIFESTYLE DEPENDS ON THE LARGE QUANTITIES OF ENERGY MADE POSSIBLE BY OXYGEN. OXYGEN IS ALSO REQUIRED DIRECTLY OR indirectly for 1000 biochemical reactions that cannot occur under anaerobic conditions. A high price is paid for oxygen's enormous benefits, however. Research efforts have revealed that aerobic organisms have evolved an array of mechanisms that provide protection from the toxic by-products of oxygen metabolism. Numerous enzymes and antioxidant molecules usually prevent most oxidative cell damage. Despite this protection, injury does occur. Oxygen metabolites are now known to contribute to an array of human disorders that include cancer and heart and neurological diseases.

Oxidative Cell Damage

Oxygen has several properties that in combination have made possible a highly favorable mechanism for extracting energy from organic molecules. First, oxygen is found almost everywhere on Earth's surface. In contrast, most other electron acceptors are relatively rare. Second, oxygen diffuses easily across cell membranes. This is not true of several other electron acceptors such as the charged species sulfate and nitrate. Finally, its diradical structure allows oxygen to readily accept electrons. This capacity is also responsible for its tendency to form highly destructive metabolites called *reactive oxygen species* (ROS).

Chapter 10 describes the basic principles of *oxidative phosphorylation*, the complex mechanism by which modern aerobic cells manufacture ATP. The discussion begins with a review of the electron transport system in which electrons are donated by reduced coenzymes to the electron transport chain (ETC). The ETC is a series of electron carriers in the inner membrane of the mitochondria of eukaryotes and the plasma membrane of aerobic prokaryotes. The chapter's next topic is *chemiosmosis*, the means by which the energy extracted from electron flow is captured and used to synthesize ATP. Chapter 10 ends with a discussion of the formation of toxic oxygen products and the strategies that cells use to protect themselves.

10.1 ELECTRON TRANSPORT

The mitochondrial ETC, also referred to as the electron transport system, is a series of electron carriers arranged in the inner membrane in order of increasing electron affinity; it is these molecules that transfer the electrons derived from the reduced coenzymes NADH and $FADH_2$ to oxygen. During this transfer, there is a decrease in reduction potential ($\Delta E°'$). When NADH is the electron

donor and oxygen is the electron acceptor, the change in standard reduction potential is +1.14 V (i.e., +0.82V − (−0.32V); see **Table 9.1**). The energy released during electron transfer is coupled to several endergonic processes, the most prominent of which is ATP synthesis. Other processes driven by electron transport transfer Ca^{2+} into the mitochondrial matrix via mitochondria associated membranes (MAMs) (p. 60) and generate heat in brown adipose tissue (described on p. 383). Reduced coenzymes, derived from glycolysis, the citric acid cycle, and fatty acid oxidation, are the principal sources of electrons.

Electron Transport and Its Components

The components of the ETC in eukaryotes, located in the inner mitochondrial membrane (**Figure 10.1**), are organized into four complexes. Each of these complexes, which function as oxidoreductase enzymes, consists of several proteins and prosthetic groups. The structural and functional properties of each complex are briefly described. The roles of coenzyme Q (ubiquinone, UQ) and the cytochromes are also described. **Figure 10.1** provides an overview of electron flow in the ETC.

Complex I, also referred to as the *NADH dehydrogenase complex*, catalyzes the transfer of electrons from NADH to UQ. The largest protein component in the inner membrane, NADH dehydrogenase is L-shaped with

1. a hydrophilic peripheral arm that possesses the NADH, FMN (**Figure 9.6**), UQ (**Figure 10.2**) binding sites, and a series of iron–sulfur clusters; and
2. a membrane domain composed of transmembrane helices that contain four proton-translocating channels.

The bovine heart mitochondrial enzyme has at least 45 subunits with a molecular mass of 1 MDa (1000 kDa). The bacterial enzyme complex has a molecular mass of about 550 kDa. Iron–sulfur clusters (**Figure 10.3**), most of which consist of two to four iron atoms complexed with the sulfurs of cysteine residues, mediate one-electron transfer reactions.

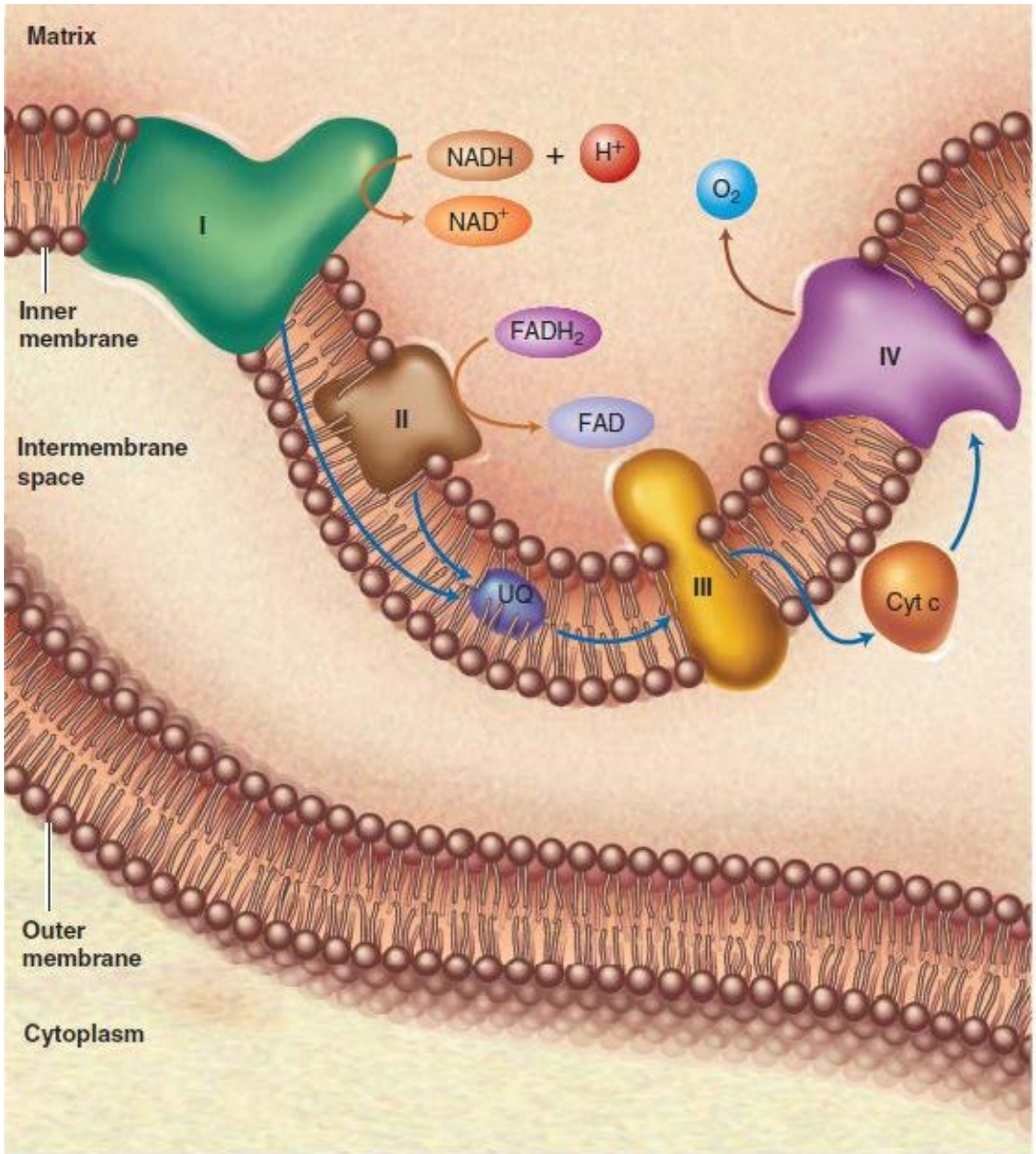


FIGURE 10.1

The Electron Transport Chain

Complexes I and II transfer electrons from NADH and succinate (via $FADH_2$), respectively, to UQ. Complex III transfers electrons from UQH_2 to cytochrome c. Complex IV transfers electrons from cytochrome c to O_2. The arrows represent the flow of electrons. The function of electron transport is to transfer protons from the matrix across the inner membrane into the intermembrane space. ATP synthesis is linked to proton transport.

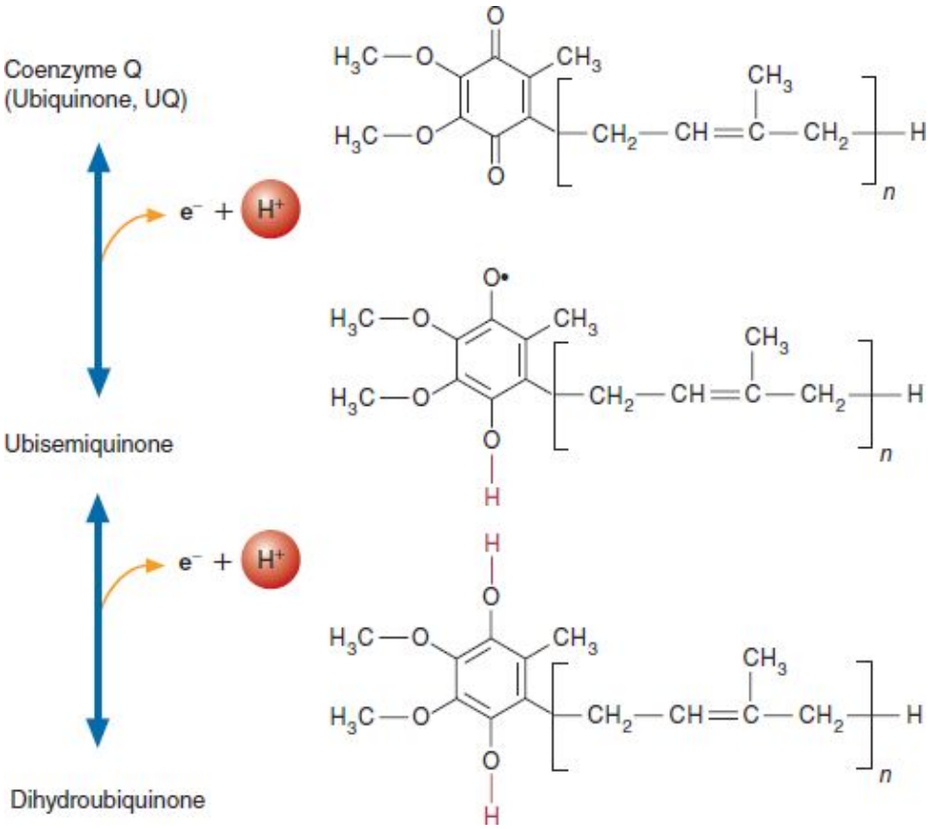


FIGURE 10.2

Structure and Oxidation States of Coenzyme Q

The length of the side chain varies among species. For example, some bacteria have six isoprene units. For mammals, however, $n = 10$, and such molecules are referred to as Q_{10}.

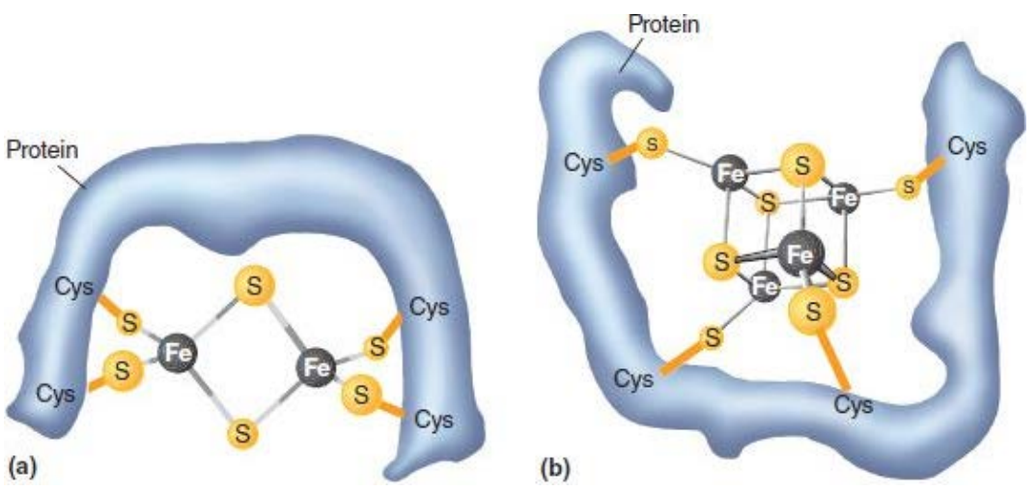


FIGURE 10.3

Two Iron–Sulfur Clusters

(a) An iron–sulfur cluster with 2 irons and 2 sulfurs. (b) A 4Fe–4S iron–sulfur cluster. In both cases, the cysteine residues are part of a polypeptide.

Figure 10.4 illustrates the transfer of electrons through complex I of the archaea *Thermus thermophilus*. When NADH is oxidized in its binding site in the distal segment of the peripheral arm,

its hydride ion is donated to FMN to yield $FMNH_2$. The two electrons removed from NADH are then transferred one at a time to the first of 95 Å-long sequences of seven iron–sulfur clusters, positioned with increasing electron affinity, with both electrons being donated to UQ to yield UQH_2. During each cycle, a minielectric current, which lasts about 5 ms (milliseconds), triggers conformational changes that facilitate the movement of four protons through separate channels in the membrane domain from the matrix into the intermembrane space. The significance of this phenomenon for ATP synthesis will be discussed.

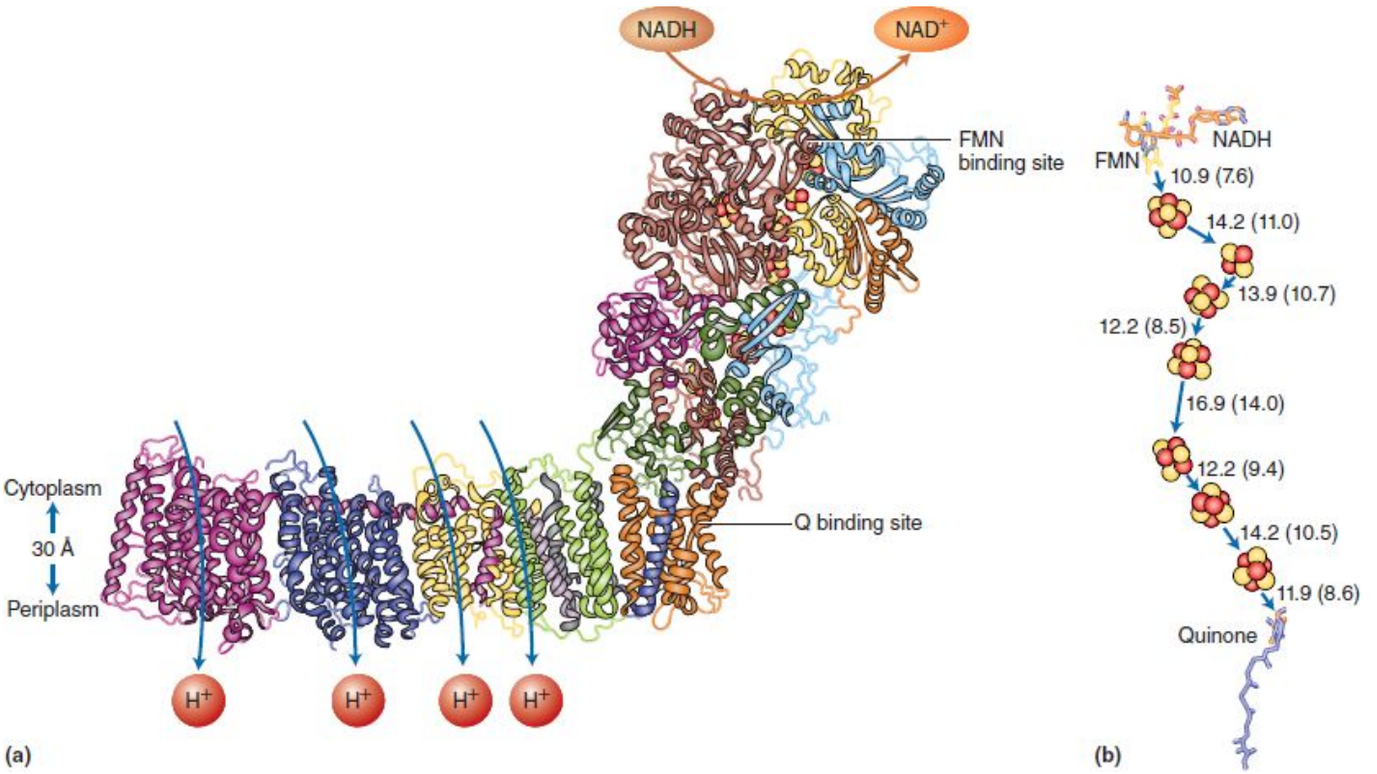


FIGURE 10.4

Transfer of Electrons through Complex I of *Thermus thermophilus*

(a) Electron transport within the peripheral arm of Complex I (the NADH-oxidizing dehydrogenase) begins with the reduction of FMN by NADH, a process that requires one proton from the cytoplasm. The two electrons of $FMNH_2$ are transferred, one at a time, to seven Fe-S clusters, after which both reduce UQ at the Q site to yield UQH_2. The movement of electrons through the peripheral arm causes conformational changes, involving altered pK_a's of specific side chains, which open proton channels in the membrane domain. (b) The iron sulfur clusters are depicted in their relative positions within the peripheral arm of Complex I in relation to the NADH binding site and the Q site. The distances between the clusters, measured in angstroms (Å), roughly the size of atoms, are center to center and, in parentheses, edge to edge. Note that the edge-to-edge distances are 14 Å or less. These short distances allow electron transfer via electron tunneling, a quantum-mechanical process in which particles with low-mass and wave properties such as electrons can move (i.e., tunnel) through activation barriers rather than over them. Electron tunneling, facilitated by the precise orientation of polypeptide side chains and internal water molecules, increases the rate of electron transfer to physiologically relevant values. (Refer to the online Biochemistry in Perspective reading *Quantum Tunneling and Catalysis* cited in Chapter 6.)

The *succinate dehydrogenase complex* (complex II), also called succinate ubiquinol reductase, consists of four protein subunits. Both ShdA, a flavoprotein with a succinate binding site and a covalently bound FAD, and ShdB, an iron–sulfur protein with three iron–sulfur clusters, extend into the matrix (p. 347). Subunits C and D are integral membrane proteins. The hydrophobic UQ binding site is in a crevice that is lined with side chains from ShdC and ShdD. The CD dimer also contains a binding site for a heme group. The role of the heme group is to suppress electron leakage from

complex II, a process that can result in oxygen radical formation.

Complex II mediates the transfer of electrons from succinate to UQ (**Figure 10.5**). After FAD is reduced (in the oxidation of the citric acid cycle intermediate succinate, p. 347), its electrons are transferred to a series of three iron–sulfur clusters and then to UQ. Unlike the other ETC complexes, succinate dehydrogenase does not translocate protons from the matrix to the intermembrane space. Other flavoproteins also donate electrons to UQ (**Figure 10.6**). In some cell types, glycerol-3-phosphate dehydrogenase, an enzyme located on the outer face of the inner mitochondrial membrane (IMM), transfers electrons from cytoplasmic NADH to the ETC (see p. 306). Several other $FADH_2$-producing dehydrogenases also donate electrons to UQ by utilizing a short electron transport mechanism involving electron transferring flavoprotein (ETF) and electron transferring flavoprotein ubiquinone oxidoreductase (ETF:QO). ETF is a small-matrix protein that functions as an electron acceptor for dehydrogenases in fatty acid oxidation, amino acid oxidation (e.g., proline, leucine, isoleucine, and valine), and pyrimidine synthesis (p. 559). ETF's covalently bound $FADH_2$ is then oxidized by ETF:QO, which then transfers the electrons to UQ. Acyl-CoA dehydrogenase, the first enzyme in fatty acid oxidation (Chapter 12), transfers electrons to UQ from the matrix side of the inner membrane.

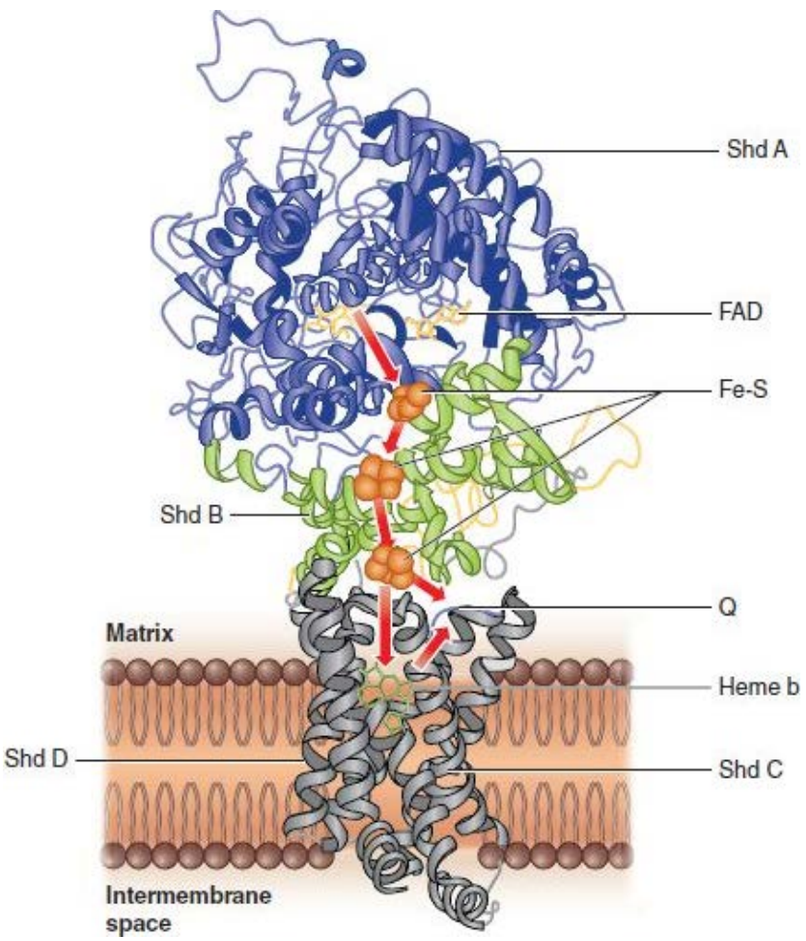


FIGURE 10.5

Transfer of Electrons through Mitochondrial Complex II

Succinate is oxidized by a covalently bound FAD within ShdA to yield fumarate and $FADH_2$. The two electrons accepted by the flavin cofactor (red arrow) are then donated one at a time to a series of three iron–sulfur clusters. The electrons are then quickly donated, again one at a time, to UQ in its binding site, located in a gap between ShdB, ShdC, and ShdD. The first electron reduction step yields ubisemiquinone. UQH_2 is formed with the acceptance of the second electron. Note that the heme group is not in the direct path of electron transfer, but may function as an electron sink (i.e., it suppresses ROS formation).

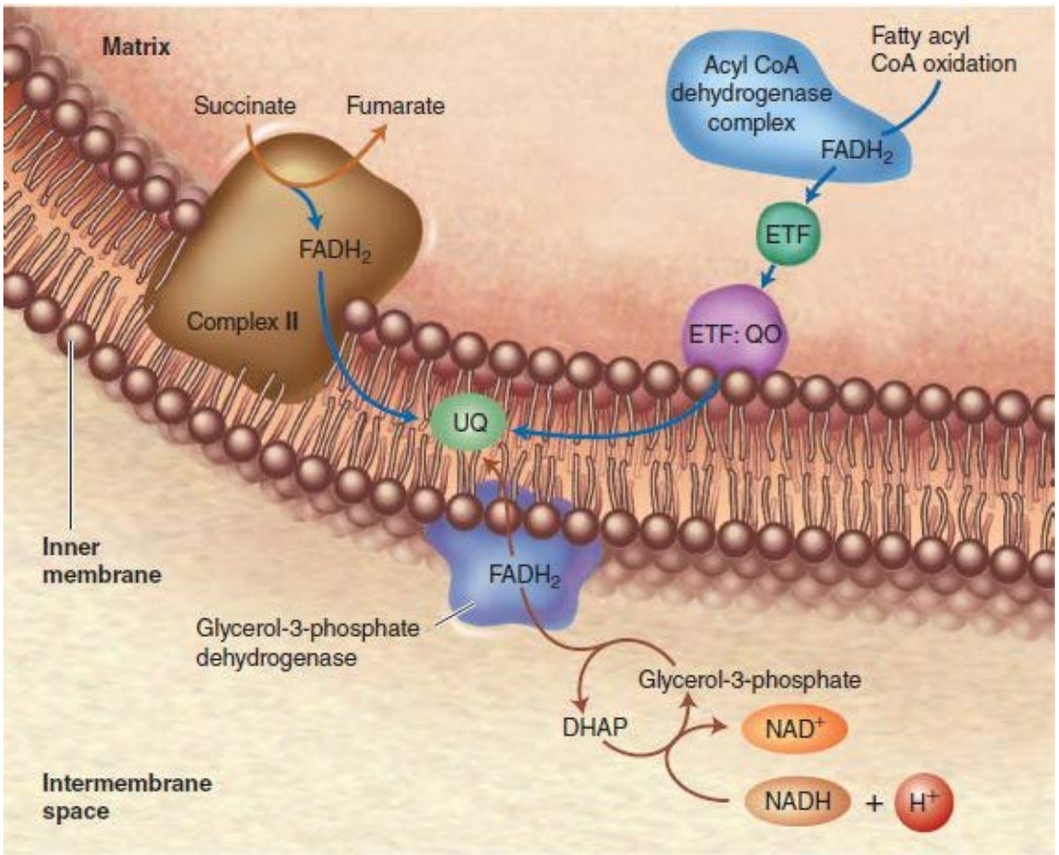


FIGURE 10.6

Electron Transport Pathways to UQ from $FADH_2$-producing Dehydrogenases

Electrons from succinate are transferred within Complex II to UQ to yield UQH_2. Electrons from several other flavoprotein dehydrogenases also donate electrons to UQ via $FADH_2$. Electrons from cytoplasmic NADH are transferred to UQ by glycerol-3-phosphate dehydrogenase (see p. 306). Acyl-CoA dehydrogenase is an example of a large group of dehydrogenases that donate their electrons to UQ via ETF and ETF:QO.

Complex III, also referred to as *cytochrome bc_1 complex*, is a dimer, with each monomer containing 11 subunits (**Figure 10.7**). Among these are three cytochromes (cyt b_L, cyt b_H, and cyt c_1) and one iron-sulfur cluster. The *cytochromes* are a series of electron transport proteins that contain a heme prosthetic group. Electrons are transferred by cytochromes one at a time in association with a reversible change in the oxidation state of a heme iron (i.e., between a reduced Fe^{2+} and an oxidized Fe^{3+}). The function of complex III is the transfer of electrons from reduced coenzyme Q (UQH_2) to a protein called cytochrome c (cyt c). Cyt c (**Figure 10.8**) is a mobile electron carrier that is loosely associated with the outer face of the inner mitochondrial membrane.

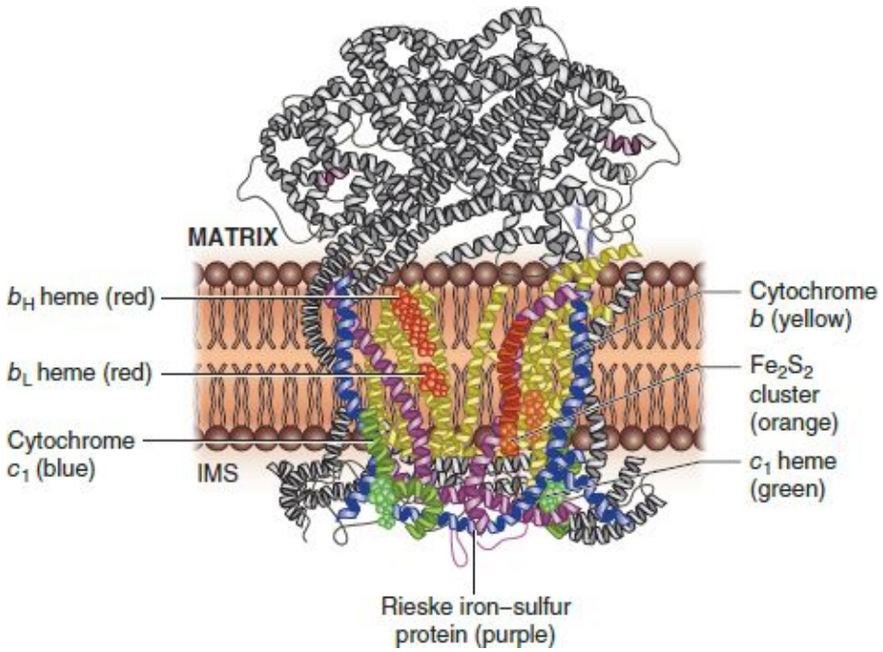


FIGURE 10.7

Complex III Structure

Cytochrome bc_1 oxidoreductase is a homodimer in the IMM that catalyzes the redox reaction between two mobile molecules: UQH_2 and cytochrome c. It has four cofactors, cyt c, cyt b_H, cyt b_L, and a 2Fe–2S cluster within a Rieske iron–sulfur protein subunit. The two substrates for this enzyme complex are UQH_2 and cyt c_{ox} (Fe^{3+}). Its products are UQ and cyt c_{red} (Fe^{2+}). Electron transfer through complex III causes the pumping of four protons across the IMM into the intermembrane space.

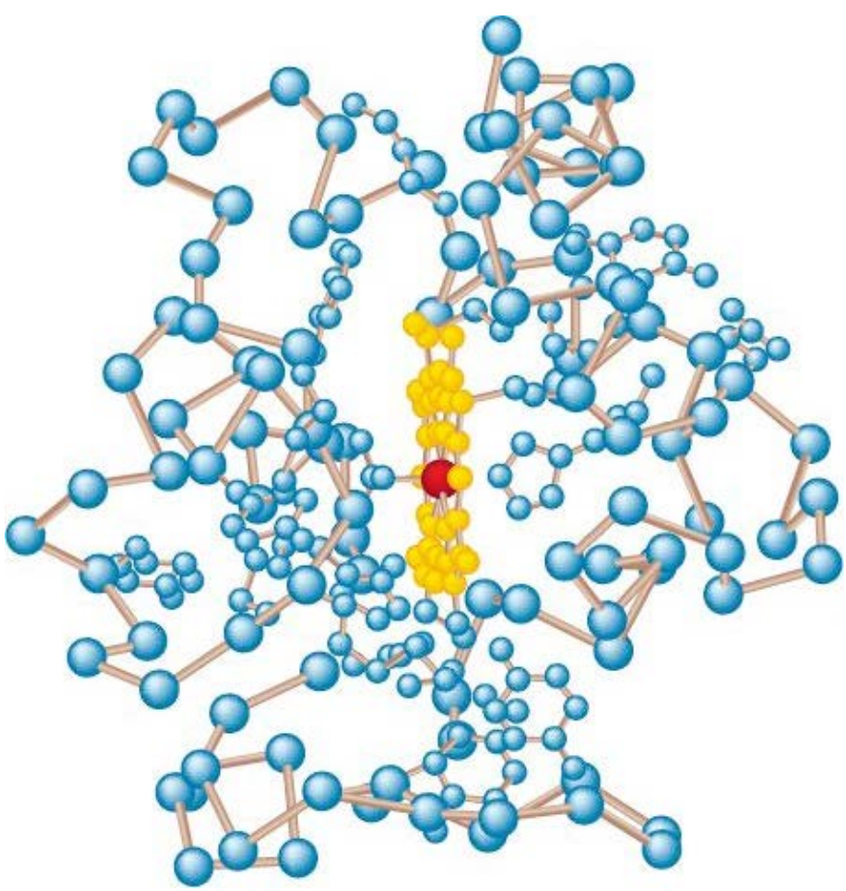


FIGURE 10.8

Structure of Cytochrome c

Cytochrome c is a member of the class of small proteins called cytochromes, each of which possesses a heme prosthetic group. During the electron transport process, the iron in the heme is alternately oxidized and reduced.

The passage of electrons through complex III (**Figure 10.9**), referred to as the **Q cycle**, is complicated. However, the overall reaction for the process whereby each UQH_2 donates two electrons to cyt c is straightforward:

$$UQH_2 + 2\ \text{cyt}\ c_{ox}\,(Fe^{3+}) + 2H^+_{matrix} \rightarrow UQ + 2\ \text{cyt}\ c_{red}\,(Fe^{2+}) + 4H^+_{IMS} \qquad (1)$$

During the Q cycle, coenzyme Q molecules diffuse within the inner membrane between the electron donors in complex I or II and the electron acceptor in complex III. UQH_2 donates its two electrons one at a time. One electron flows to the Fe–S protein, which transfers it to cyt c_1 and, subsequently, to cyt c. The products of this transfer are $UQ^{\bullet}$ and two protons from UQH_2, which are transferred to the intermembrane space. The $UQ^{\bullet}$ then transfers its second electron first to cyt b_L and then to cyt b_H. The product of the first round of the Q cycle is UQ, with one electron transferred to cyt c and the second electron to cyt b_H; two protons are transferred into the intermembrane space. The second round of the Q cycle involves a second molecule of UQH_2, which transfers one electron to cyt c in the same manner as in the first round. The product $UQ^{\bullet}$ then accepts the electron from cyt b_H and two protons from the mitochondrial matrix to form UQH_2. The net effect is that two electrons and four protons are fed off the intermembrane space side of the inner mitochondrial membrane to oxidize cyt c and contribute to the proton gradient. One molecule each of UQ and UQH_2 is formed within the membrane, with two protons contributed from the mitochondrial matrix.

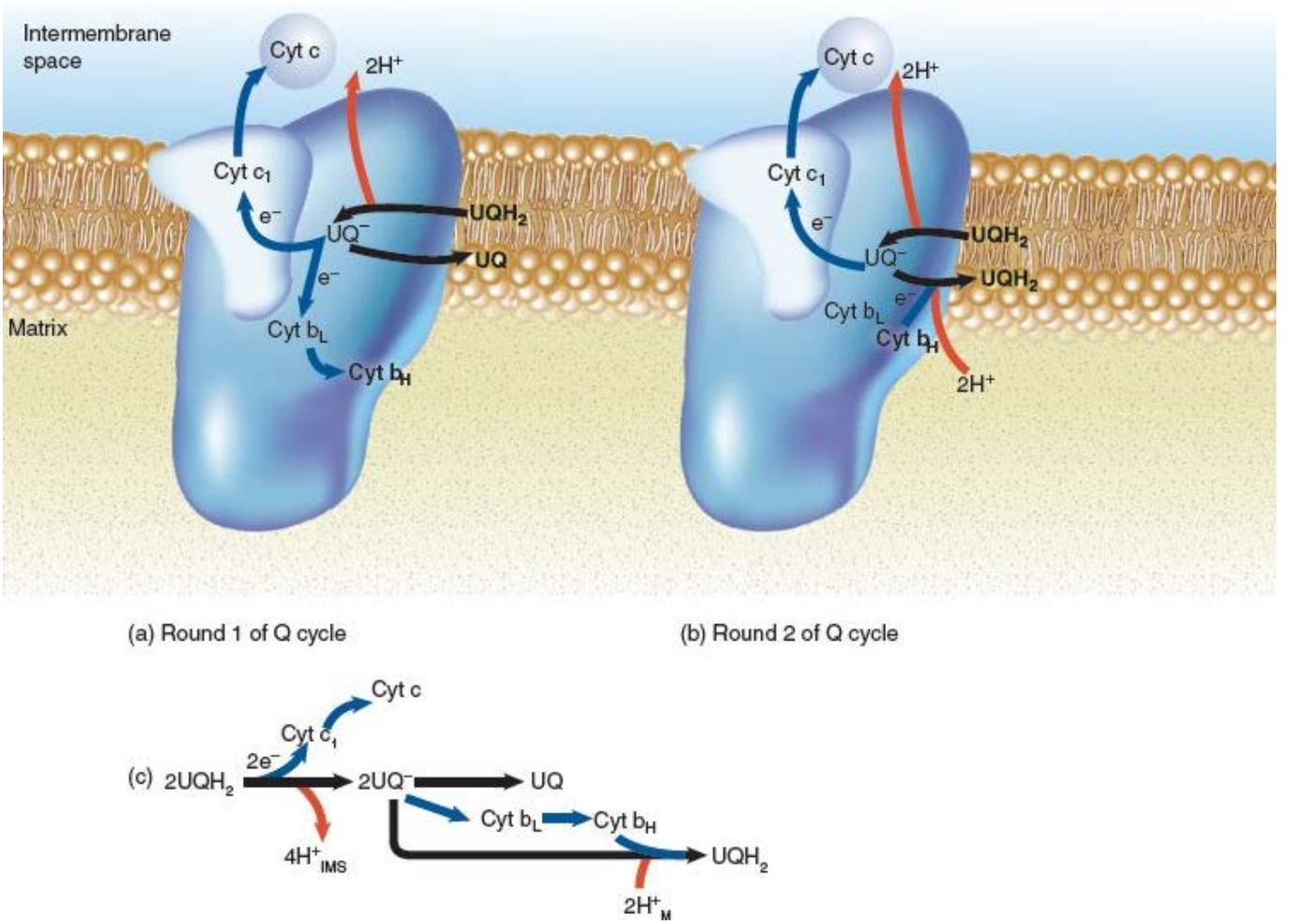


FIGURE 10.9

Electron Transport through Complex III

Two molecules of UQH_2 are oxidized sequentially to supply two electrons (e^-) to cyt c. The first electron from each UQH_2 is transferred to the Fe-S protein (not shown) and then to cyt c_1 and cyt c as two protons each are transferred to the intermembrane space. (a) The second electron from one $UQ^{\bullet-}$ is transferred to cyt b_L and then to cyt b_H. UQ is released into the inner mitochondrial membrane. (b) The second $UQ^{\bullet-}$ picks up the electron from cyt b_H and two protons from the mitochondrial matrix to form UQH_2. (c) The summary reaction shows that two UQH_2 molecules enter complex III, and UQ and UQH_2 are released from the complex. (Black arrows = reaction arrows; red arrows = proton transfer; blue arrows = electron flow.)

Cytochrome oxidase (complex IV) is a homodimer protein that catalyzes the four-electron reduction of O_2 to form H_2O. Note that O_2 can only accept one electron at a time because the two electrons in its outermost valence shell have the same spin quantum number (see p. P-4). (It is this property that can result in ROS formation.) Each monomer of the membrane-spanning complex (**Figure 10.10**) in mammals contains 14 cytochromes a and a_3 and three copper ions. Two copper ions form Cu_A, a binuclear Cu–Cu center, and heme a_3 and Cu_B form a binuclear Fe–Cu center. Both centers accept electrons, one at a time (**Figure 10.11**). Electrons flow from cytochrome c to Cu_A to cytochrome a and then to a_3-Cu_B and, finally, to O_2. Four electrons are shuttled through complex IV from the outer face of the inner mitochondrial membrane to the matrix for delivery to the cytochrome a_3-Fe(II)-bound dioxygen, which then reacts with four matrix protons to form two water molecules:

$$O_2 + 4H^+ + 4e^- \rightarrow 2H_2O \qquad (2)$$

The overall reaction is

$$4\ \text{cyt}\ c_{red}\ (Fe^{2+}) + 8H^{+}_{matrix} + O_2 \rightarrow 4\ \text{cyt}\ c_{ox}\ (Fe^{3+}) + 2H_2O + 4H^{+}_{IMS}$$

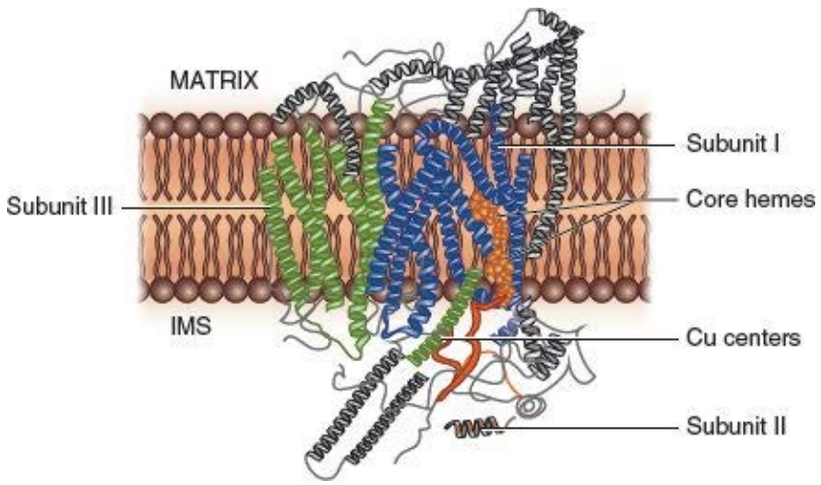


FIGURE 10.10

Structure of Complex IV (Cytochrome Oxidase)

Complex IV is a homodimer in which the monomers consist of 14 subunits. Subunit I (blue) has two hemes (a and a_3). The iron atom in heme a_3 and Cu_B form a binuclear Fe-Cu center. Subunit II (red) contains two Cu ions that form the binuclear center Cu_A. Subunit III (green) facilitates the transport of four protons from the matrix to the IMS. All other subunits are gray.

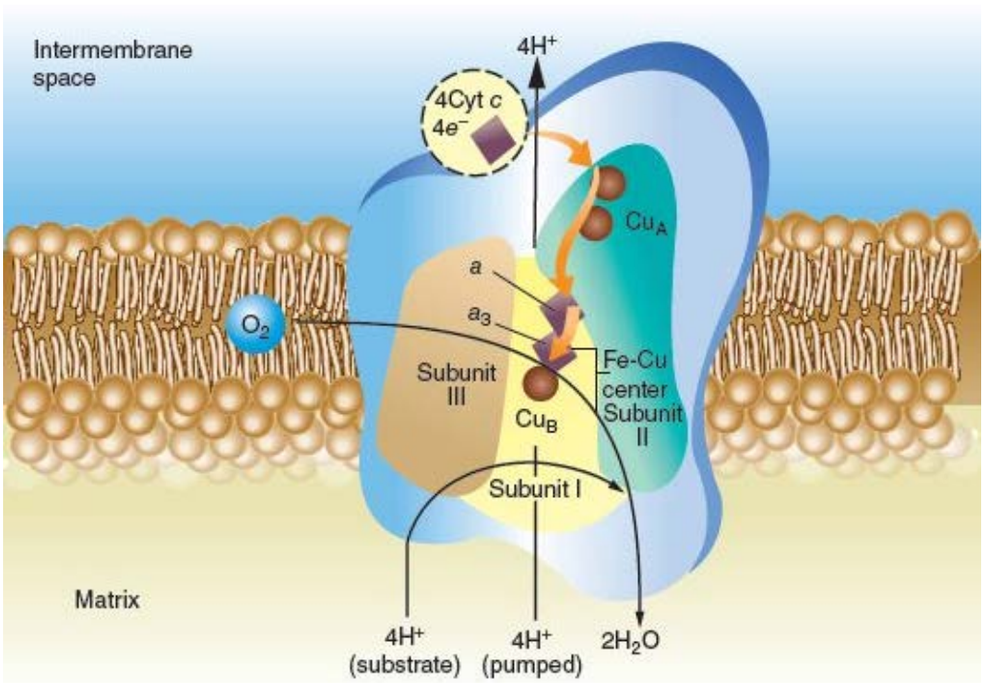


FIGURE 10.11

Electron Transport through Complex IV

The process begins with the sequential oxidation of four cytochrome c molecules by CuA. Each electron is transferred one at a time to cyt a and then to the binuclear a_3-Cu center, where O_2 undergoes reduction utilizing four matrix protons. Note that unlike other electron transport complexes, electrons do not leak out of complex IV.

The energy release from electron transport through Complex IV pumps four protons from the matrix to the IMS.

There are ATP-binding regulatory sites on cyt c and complex IV. When ATP concentrations are high, ATP, acting as an allosteric inhibitor, binds to these sites and causes decreased electron transport activity.

NADH oxidation results in the release of a substantial amount of energy, measured by decreased reduction potentials ($\Delta E^{\circ\prime}$), as electrons flow through complexes I, III, and IV (**Figure 10.12**). Approximately 2.5 molecules of ATP are synthesized for each pair of electrons transferred between NADH and O_2 in the ETC. Approximately 1.5 molecules of ATP result from the transfer of each pair donated by the $FADH_2$ produced by succinate oxidation. The mechanism by which ATP synthesis is coupled to electron transport, the proton gradient, is described on page 374. The components of the ETC are summarized in **Table 10.1**.

KEY CONCEPTS

The electron transport chain is a series of complexes consisting of molecular electron carriers located in the inner mitochondrial membrane of eukaryotic cells.

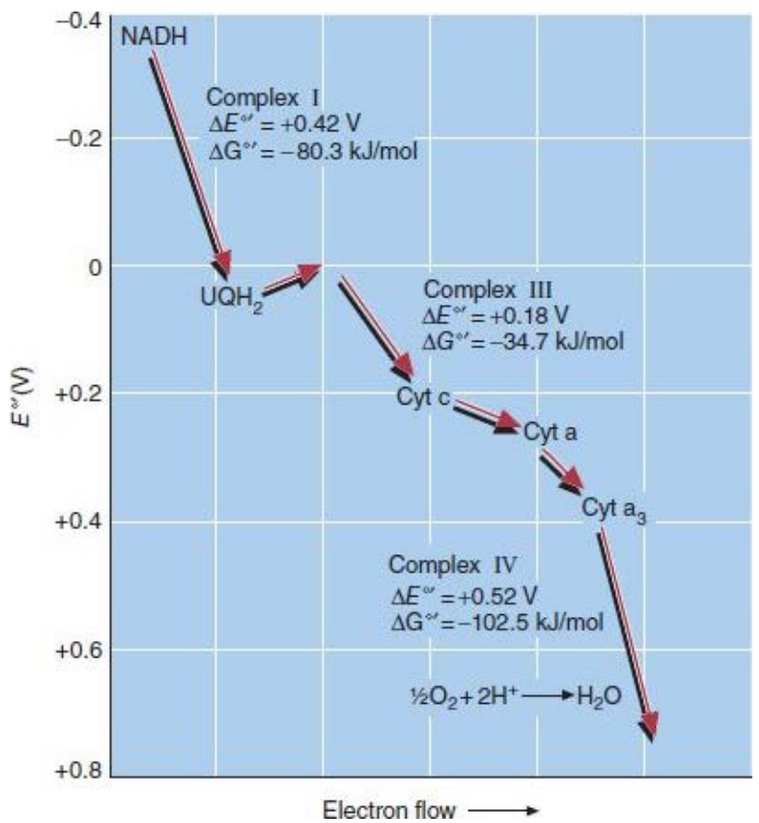


FIGURE 10.12

The Energy Relationships in the Mitochondrial Electron Transport Chain

Relatively large decreases in free energy occur in three steps. During each of these steps, sufficient energy is released to account for the synthesis of ATP.

TABLE 10.1 Supramolecular Components of the Electron Transport Chain

Enzyme Complex	Prosthetic Groups

Complex I (NADH dehydrogenase)	FMN, FeS
Complex II (succinate dehydrogenase)	FAD, FeS
Complex III (cytochrome bc_1 complex)	Hemes, FeS
Cytochrome c	Heme
Complex IV (cytochrome oxidase)	Hemes, Fe, Cu

Electron Transport: The Fluid State and Solid State Models

The fluid state model has been the conventional view of mitochondrial electron transport for several decades. Based in part on the experimental extraction and purification of the four ETC complexes in enzymatically active forms, the fluid state model describes electron transfer between ETC components. Electron transfer is seen as the result of random collisions among the ETC complexes and the mobile electron carriers UQ and cytochrome c. Within the past two decades, however, significant evidence has supported a solid state organization of the electron transport process instead. As a result of gentler methods of extraction and purification, several supercomplexes (SCs) composed of combinations of complexes I, III, and IV (e.g., I, III and III, $IV_{1–4}$, but not complex II) have been isolated and identified. The I, III_2, $IV_{1–2}$ supercomplex has been found in several animals, plants, and fungi. Now referred to as the **respirasome**, it is believed to be the functional respiration unit because it is the most active and stable version of the ETC. (The symbol III_2 represents the complex III dimer.)

According to the solid state model, electron transfer is efficient because of short diffusion distances for the mobile electron carriers. Structural studies of the bovine cardiac I, III_2, IV supercomplex indicate that the binding sites of the mobile electron carriers in the ETC complexes are in close proximity. For example, the UQ binding site in complex I directly faces the complex III binding site for reduced UQ. The IMM is exceptionally protein-dense (a 75%:25% protein-to-lipid ratio). As a result, random protein movement, as envisioned in the fluid state model, is severely restricted in the heavily folded inner membrane.

Electron Transport Inhibitors

Several molecules specifically inhibit the electron transport process (**Figure 10.13**). Used in conjunction with reduction potential measurements, inhibitors have been invaluable in determining the correct order of ETC components. In such experiments, electron transport is measured with an oxygen electrode, since O_2 consumption is a sensitive measure of this process. When electron transport is inhibited, oxygen consumption is reduced or eliminated. Oxidized ETC components accumulate on the O_2-reducing side of the site of inhibition. Reduced ETC components accumulate on the nonoxygen side of the site of inhibition. For example, antimycin A inhibits cyt b in complex III. If this inhibitor is added to a suspension of mitochondria, NAD^+, the flavins, and cyt b molecules become more reduced, and the cytochromes c_1, c, and a become more oxidized. Other prominent examples of ETC inhibitors include rotenone and amytal, which inhibit NADH dehydrogenase (complex I). Carbon monoxide (CO), azide (N_3^-), and cyanide (CN^-) inhibit cytochrome oxidase.

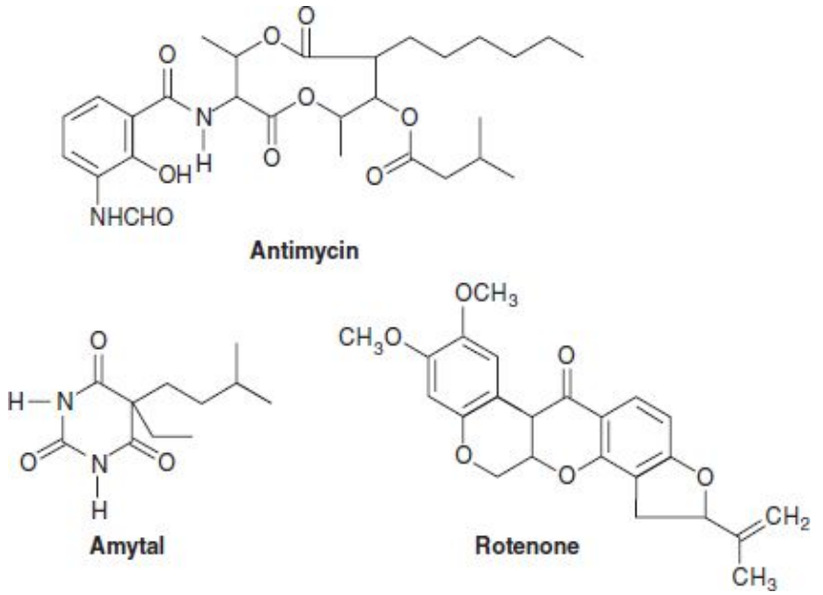


FIGURE 10.13

Several Inhibitors of the Mitochondrial Electron Transport Chain

Antimycin blocks the transfer of electrons from the b cytochromes. Amytal and rotenone block NADH dehydrogenase.

QUESTION 10.1

Which compound in each of the following pairs is the better reducing agent?

a. $NADH/H_2O$
b. $UQH_2/FADH_2$
c. Cyt c (reduced)/cyt b (reduced)
d. $FADH_2/NADH$
e. $NADH/FMNH_2$

10.2 OXIDATIVE PHOSPHORYLATION

Oxidative phosphorylation, the process whereby the energy generated by the electron transport chain is conserved by the phosphorylation of ADP to yield ATP, is explained by the **chemiosmotic coupling theory** proposed by Peter Mitchell in 1961. In *chemiosmosis*, the energy released by electron transport generates an electrochemical gradient across a membrane, which in turn drives ATP synthesis. Living organisms use an astonishing amount of ATP to sustain living processes. By one estimate, a single human cell uses about 10 million ATPs per second. As a result, ATP turnover per day for an average human is in the vicinity of 60 kg.

The Chemiosmotic Theory

The chemiosmotic coupling theory has the following features:

1. As electrons pass through the ETC, protons are transported from the matrix and released into the intermembrane space. The results are an electrical potential Ψ [a charge difference of 150 millivolts (mV) (negative inside)] and a proton gradient ΔpH of approximately 0.5 pH units across the inner membrane. The electrochemical proton gradient is sometimes referred to as

the **protonmotive force** Δp.

2. Protons, which are present in the intermembrane space in great excess as a result of the electron transport process, can pass through the inner membrane and back into the matrix down their concentration gradient only through special channels. (The inner membrane itself is impermeable to ions such as protons.) ATP is synthesized from ADP and P_i by a molecular machine called *ATP synthase*. Also referred to as complex V, it contains a proton channel. ATP synthesis occurs as a result of a thermodynamically favorable flow of protons through the channel.

Mitchell used the term *chemiosmotic* to emphasize that chemical reactions can be coupled to osmotic gradients. **Figure 10.14** presents an overview of the chemiosmotic model as it operates in the mitochondrion.

Examples of the evidence that supports the chemiosmotic theory include the following:

1. Actively respiring mitochondria expel protons. The pH of a weakly buffered suspension of mitochondria measured by an electrode drops when O_2 is added. In actively respiring mitochondria, the pH within the matrix is close to 8, and that of the intermembrane space is about 7.4.
2. ATP synthesis stops when the inner membrane is disrupted. For example, although electron transport continues, ATP synthesis stops in mitochondria placed in a hypertonic solution. Mitochondrial swelling results in proton leakage across the inner membrane.
3. A variety of ATP synthesis inhibitor molecules are now known to specifically dissipate the proton gradient (**Figure 10.15**). According to the chemiosmotic theory, a disrupted proton gradient dissipates the energy derived from food molecules as heat. **Uncouplers** such as dinitrophenol collapse the proton gradient by equalizing the proton concentration on both sides of membranes. (As they diffuse across the membrane, uncouplers pick up protons from one side and release them on the other.) **Ionophores** are hydrophobic molecules that dissipate osmotic gradients by inserting themselves into a membrane and forming a channel. For example, gramicidin is an antibiotic that forms a channel in membranes that allows the passage of H^+, K^+, and Na^+.

The proton gradients generated by electron transport systems are dissipated for two general purposes: ATP is synthesized as protons flow through the ATP synthase, and regulated proton leakage is used to drive several other types of biological work. Examples include heat generation and transport of substances such as phosphate and the adenine nucleotides ADP and ATP across the inner membrane (see p. 378). In the next section, descriptions of ATP synthesis and its regulation within mitochondria are followed by a brief overview of energy generation from glucose. Section 10.2 ends with a discussion of nonshivering thermogenesis, the mechanism by which the mitochondrial proton gradient in certain animal cells regulates body temperature.

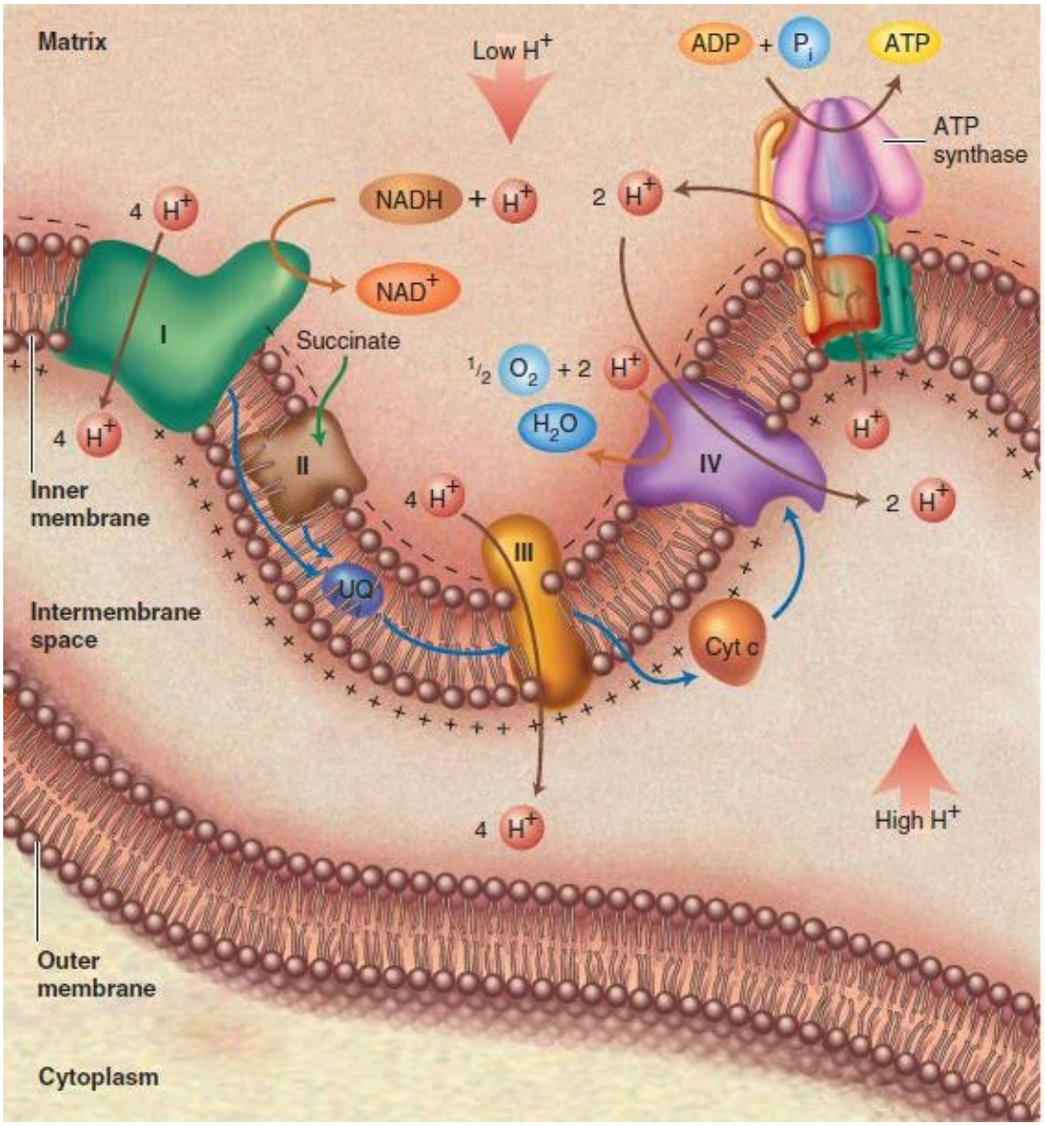


FIGURE 10.14

Overview of the Chemiosmotic Model

In Mitchell's model, protons are driven from the mitochondrial matrix across the inner membrane and into the intermembrane space by the electron transport mechanism. The energy captured from electron transport is used to create an electrical potential and a proton gradient. Because the inner membrane is impermeable to protons, they can traverse the membrane only by flowing through specific proton channels. The flow of protons through the ATP synthase drives the synthesis of ATP. (See **Figure 10.1** for brief descriptions of the roles of complexes I, II, III, and IV in electron transport.) Note that as a result of proton flow across the inner membrane, the pH of the matrix rises and the matrix side of the membrane becomes more negatively charged.

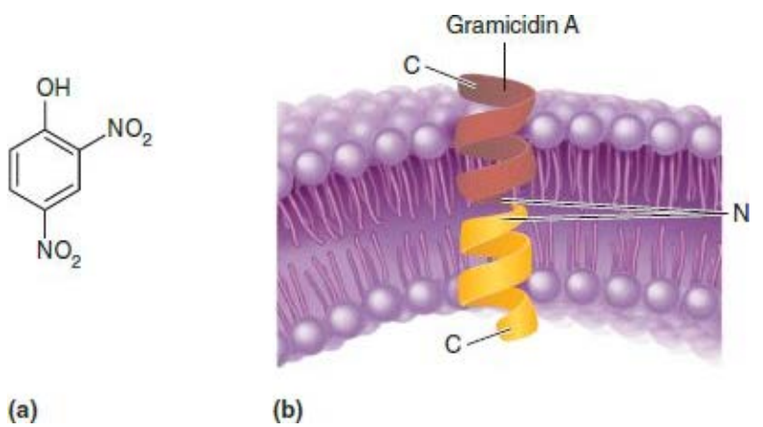


FIGURE 10.15

Uncouplers

(a) Dinitrophenol, which diffuses across the membrane, picks up protons on one side and releases them on the other. (b) Gramicidin A, an 11-residue peptide, forms an end-to-end dimer that creates a proton-permeable pore in the membrane: C = carboxy terminus; N = amino terminus.

QUESTION 10.2

Dinitrophenol (DNP) is an uncoupler that was used as a diet aid in the 1930s, until several people who had been taking it died. The dose effective for dieting is especially close to a fatal dose. Suggest why DNP consumption results in weight loss. The deaths caused by DNP were largely a result of liver failure. Explain. [*Hint:* Liver cells contain an extraordinarily large number of mitochondria.]

DNP and Liver Failure

ATP Synthesis

Early electron microscopic studies of mitochondria showed numerous lollipop-shaped structures studding the inner membrane on its inner surface (**Figure 10.16**). Experiments begun in the early 1960s revealed that each lollipop is a proton translocating ATP synthase. These investigations made use of *submitochondrial particles*, which are small membranous vesicles formed when mitochondria are subjected to sonication. Further work showed that the ATP synthase, which works like a hydroelectric turbine (**Figure 10.17**), consists of two major components. The F_1 unit, the active ATPase, possesses five different subunits present in the ratio α_3:β_3:γ:δ:ε. There are three nucleotide-binding catalytic sites on F_1. The F_0 unit, a transmembrane channel for protons, has three subunits present in the ratio a:b_2:$c_{10–12}$. The F_0 unit is so named because its function is inhibited by oligomycin, which is an antibiotic produced by *Streptomyces*, a group of Gram-positive bacteria. Oligomycin blocks the proton channel when it binds to subunit a.

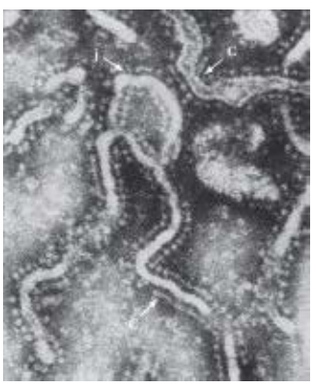


FIGURE 10.16

The ATP Synthase

An early electron micrograph of submitochondrial particles revealing the “lollipop”-like structures that would

eventually be identified as ATP synthase.

The ATP synthase consists of two rotors (rotary motors) linked together by a strong, flexible *stator* (a stationary component of a motor). In respiring organisms, the F_0 motor converts the protonmotive force into a rotational force that drives ATP synthesis catalyzed by the F_1 unit. The revolving component is the c ring, formed from the c subunits. The c ring is attached to a central shaft composed of the ε and γ subunits and rotates within the α,β hexamer of the F_1 unit. The stator (the b and δ subunits) prevents the α,β hexamer from rotating.

The synthesis of each ATP requires the translocation of approximately three protons through the ATP synthase. (The transfer of an additional proton is required for the transport of ATP and OH^- out of the matrix in exchange for ADP and P_i.) As protons flow through F_0, the rotation of the proton channel (c_{12}, also referred to as the c ring) is transmitted to the asymmetric γ subunit that projects into the core of the F_1 unit. The rotation of the central shaft, estimated at roughly 150 times per second, puts it in three possible positions relative to each α,β dimer. In effect, the protonmotive force induces three sequential 120° rotations of the α,β hexamer. As rotation proceeds, each of three nucleotide-binding sites undergoes a series of conformational changes that result in ATP synthesis. In certain circumstances, for example, *E. coli* in anaerobic conditions and fermentative lactic acid bacteria, the F_1 unit, acting as a motor, works in the opposite direction so that ATP is hydrolyzed. As a result, protons are pumped outward. The outward-moving proton gradient is then used to perform cellular work such as flagella rotation and nutrient transport.

The mechanism of ATP synthesis by the ATP synthase is as follows. In the F_0 unit, each c subunit in the c ring consists of two transmembrane antiparallel helices. The C-terminal helix of c subunits contains an essential aspartate (or glutamate) residue that upon protonation causes a swiveling motion that in turn triggers the rotation of the entire subunit. Deprotonation of this residue causes the C-terminal helix to return to its original conformation. Protons enter the c ring through the *a* subunit, which possesses two half-channels. The proton enters through the first half-channel (on the IMS side of the inner membrane) and then binds to the unprotonated carboxylate group within the proximal c subunit. This binding event causes a conformational change that displaces the proximal c subunit from the *a* subunit. As the next proton moves through the *a* subunit half-channel, the process repeats itself.

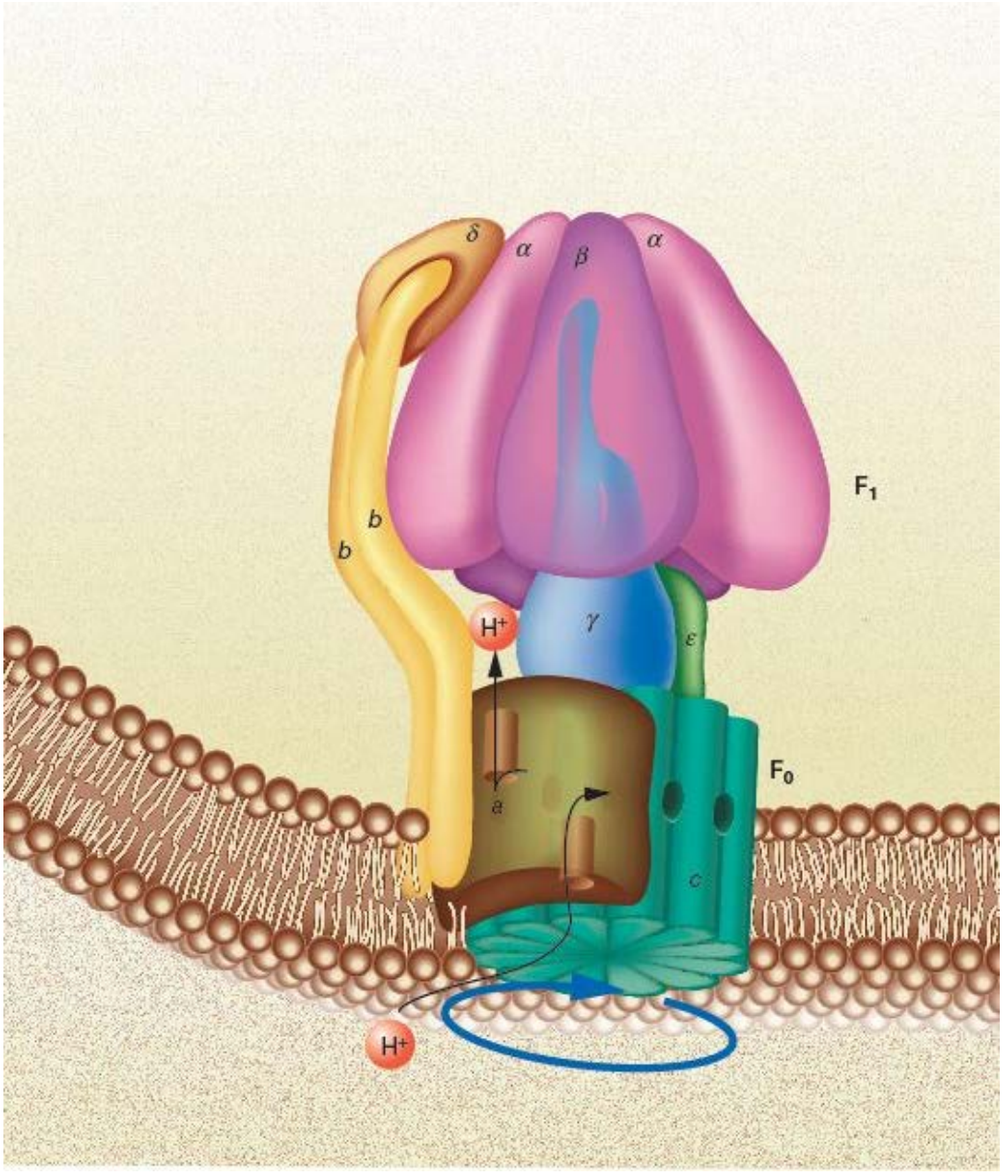


FIGURE 10.17

ATP Synthase from *Escherichia coli*

The rotor consists of ε, γ, and c_{12} subunits. The stator consists of a, b_2, δ, α_3, and β_3 subunits. The molecular components of ATP synthase are well conserved among bacteria, plants, and animals.

The net effect is the rotation of the c ring in a counterclockwise direction (as viewed from the membrane). As each *c* subunit reaches the other half-channel within the *a* subunit on the matrix, deprotonation occurs and the proton exits into the mitochondrial matrix. The *torque* (twisting force) generated by the c subunit rotation causes the asymmetric central shaft (composed of the ε and γ subunits) to rotate within a sleeve inside the α,β hexamer. The α,β hexamer catalytic sites are within the β subunits. They occur in three conformations in terms of affinity for adenine nucleotide ligands (**Figure 10.18**): *open* (O), which is inactive with low affinity; *tight* (T), which is active with high affinity; and *loose* (L), which is also inactive. Interconversion between these conformations is caused by interaction with the rotating γ subunit. There are essentially three steps in the ATP synthesizing process: (1) ADP and P_i bind to an L site; (2) ATP is synthesized when the L conformation is transformed to a T conformation; and (3) ATP is released as the T conformation converts to an O conformation. As the γ subunit rotates and sequentially interacts with each β subunit, each active site is forced through the O, T, and L conformations.

KEY CONCEPTS

- In aerobic organisms, the energy used to drive the synthesis of most ATP molecules is the protonmotive force.
- The protonmotive force is generated as free energy is released when electrons flow through the electron transport chain.

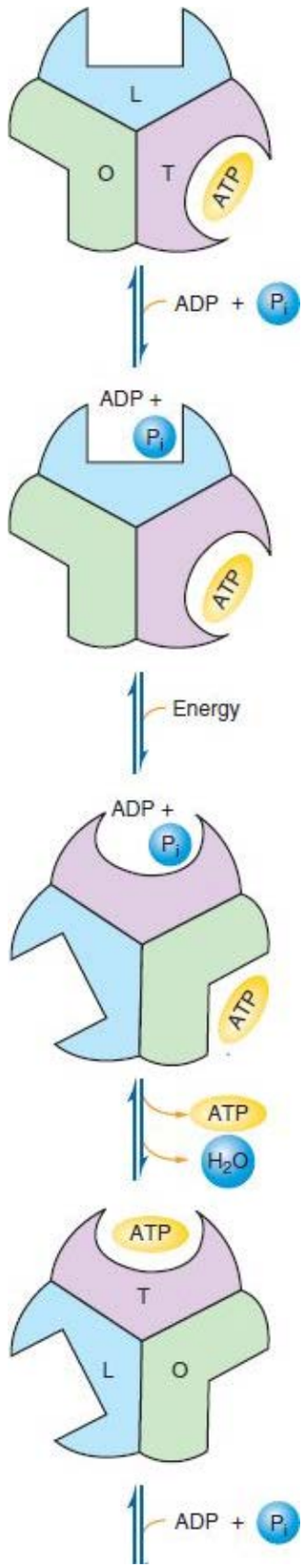


FIGURE 10.18

ATP Synthesis Model

The catalytic sites in the β subunits of ATP synthase are believed to have three conformations: open (O: inactive with low affinity for adenine nucleotides), tight (T: active with high affinity), and loose (L: also inactive). ATP synthesis begins with the binding of ADP and P_i to an L site. A conformational change driven by the rotation of the central shaft within the α,β hexamer converts the L site into a T conformation as ATP is synthesized. As the shaft continues to rotate, the T site is converted to an O site that releases the ATP molecule. ATP cannot be released from the O site unless ADP and P_i are bound to the adjacent β-subunit site in the T conformation.

QUESTION 10.3

A suspension of inside-out submitochondrial particles (i.e., the F1 unit is on the outside of the particle) is placed in a solution that contains ADP, P_i, and NADH. Will increasing the proton concentration of the solution result in ATP synthesis? Explain.

Control of Oxidative Phosphorylation

Control of oxidative phosphorylation allows cells to adapt to ever-fluctuating energy requirements. Recall that in normal circumstances, electron transport and ATP synthesis are tightly coupled. The value of the *P:O ratio* (the number of moles of P_i consumed for each oxygen atom reduced to H_2O) reflects the degree of coupling observed between the protonmotive force, created by electron transport, and ATP synthesis. The measured maximum ratio for the oxidation of NADH is 2.5. The maximum P:O ratio for $FADH_2$ is 1.5. If isolated mitochondria are provided with an oxidizable substrate (e.g., succinate), all of the ADP is eventually converted to ATP. At this point, oxygen consumption becomes greatly depressed. Oxygen consumption increases dramatically when additional ADP is supplied. The control of aerobic respiration by ADP is referred to as **respiratory control**.

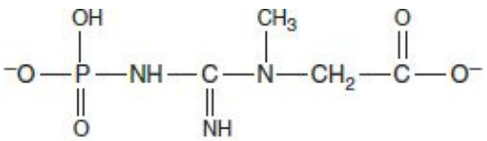
The formation of ATP appears to be strongly related to the ATP mass action ratio ([ATP]/[ADP][P_i]). In other words, ATP synthase is inhibited by a high concentration of its product (ATP) and activated when ADP and P_i concentrations are high. The relative amounts of ATP and ADP within mitochondria are controlled largely by two transport proteins in the inner membrane: the ADP–ATP translocator and the phosphate carrier. The *ADP–ATP translocator* (**Figure 10.19**), also referred to as the *adenine nucleotide translocator* (ANT), is a dimeric protein responsible for the 1:1 exchange of intramitochondrial ATP for the ADP produced in the cytoplasm. As previously described, there is a potential difference across the inner mitochondrial membrane (negative inside). Because ATP molecules have one more negative charge than ADP molecules, the outward transport of ATP and inward transport of ADP are favored. The transport of $H_2PO_4^-$, mediated by the phosphate translocase and also referred to as the $H_2PO_4^-/H^+$ symporter, is dependent on the proton gradient. (*Symporters* are transmembrane transport proteins that move solutes across a membrane in the same direction. See Section 11.2 for a discussion of membrane transport mechanisms.) The inward transport of four protons is required for the synthesis of each ATP molecule: three to drive the ATP synthase rotor and one to drive the inward transport of phosphate.

KEY CONCEPTS

- The P:O ratio reflects the coupling between electron transport and ATP synthesis.
- The measured maximum P:O ratios for NADH and $FADH_2$ are 2.5 and 1.5, respectively.

Creatine Kinase and Creatine

The sizes and complex structures of eukaryotes require an enormous ATP synthesizing capacity. This is especially true of muscle and brain cells, which have both high and fluctuating energy demands. ATP, however, presents problems that limit its capacity to provide adequate energy to drive all the energy-requiring processes in large heterogeneous eukaryotic cells. With a molecular mass of 507 Da, ATP diffuses slowly through the viscous interior of cells in relation to physiological demands. In addition, the products of ATP hydrolysis (ADP, P_i, and H^+) must be quickly removed so as to avoid inhibition of ATPases. Eukaryotes solve this problem with the creatine kinase/phosphocreatine (CK/PCr) shuttle system. With its lower molecular mass (211 Da), phosphocreatine diffuses more quickly within cells.



Phosphocreatine

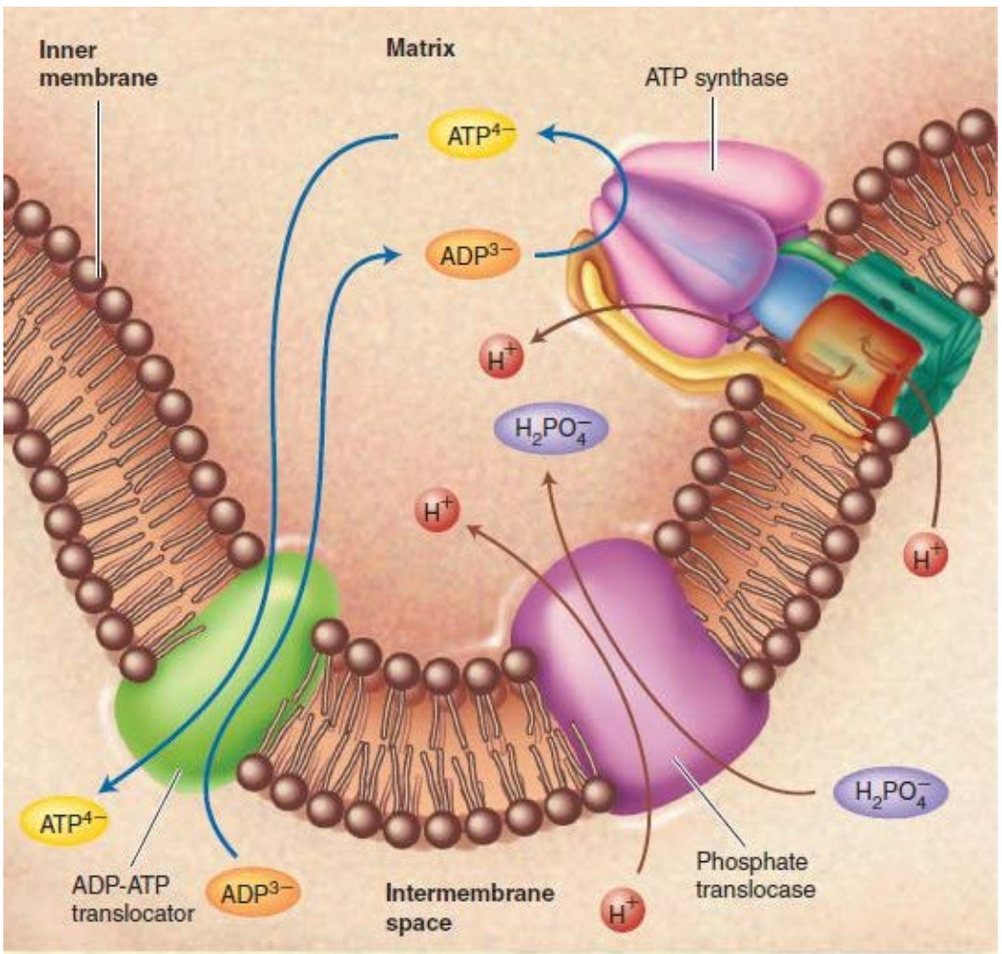


FIGURE 10.19

The ADP–ATP Translocator and the Phosphate Translocase

The transport of $H_2PO_4^-$ across the inner mitochondrial membrane by the phosphate translocase is driven by the

proton gradient. For every four protons that are transported out of the matrix, three drive the ATP synthase rotor and one drives the inward transport of phosphate. The simultaneous exchange of ADP^{3-} and ATP^{4-}, required for continuing ATP synthesis and mediated by the ADP–ATP translocator, is driven by the potential difference across the inner membrane. About 25% of the energy captured by electron transport or one proton per ATP/ADP exchange is used to regenerate the membrane potential.

THE CREATINE KINASE/PHOSPHOCREATINE SHUTTLE SYSTEM CK catalyzes the reversible reaction

$$PCr^{2-} + Mg\,ADP^{-} + H^{+} \rightarrow Cr + Mg\,ATP^{2-}$$

that cells use to maintain stable ATP/ADP ratios. [Note that PCr can be used to phosphorylate ADP because the phosphoryl group transfer potential value of phosphocreatine (Table 4.1) is −43.1 kJ/mol (−10.3 kcal/mol).] PCr is an otherwise inert molecule that is used to regenerate ATP hydrolyzed by ATPases dispersed throughout cells. Examples of ATPases that use exceptionally high ATP levels are ion pumps such as Na^{+}/K^{+}-ATPase, which maintains the resting potential of the plasma membrane and controls cell volume, and Ca^{2+}-ATPases in plasma membrane and ER membrane, which are crucial components of cell-signaling mechanisms (p. 609).

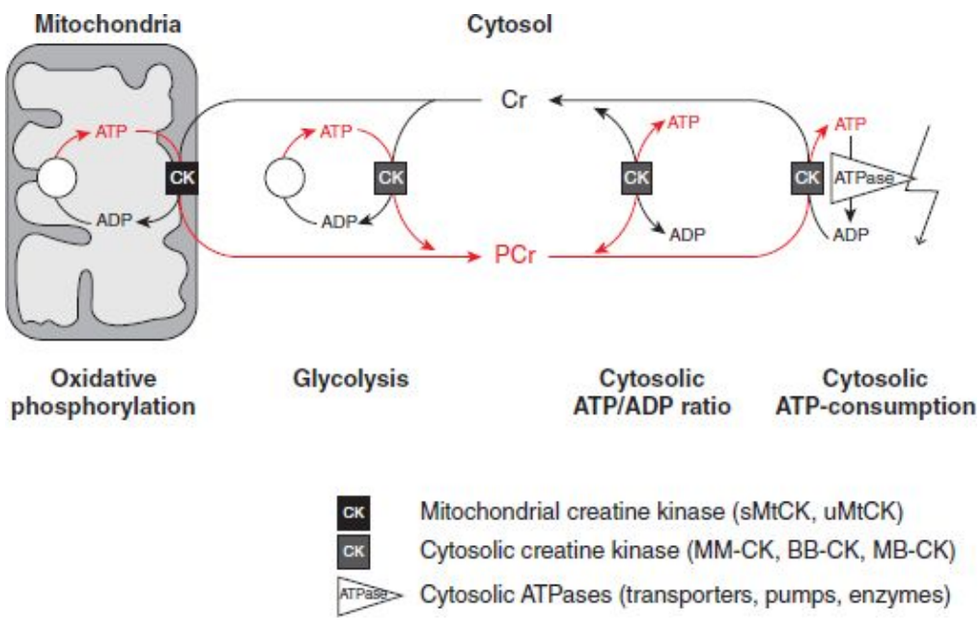


FIGURE 10.20

The Creatine Kinase/Phosphocreatine Shuttle System

The CK/PCr shuttling mechanism, which depends on microcompartment-specific CK isoenzymes, links ATP-synthesizing processes (oxidative phosphorylation and glycolysis) with ATP-consuming processes (ATPases) to maintain local ATP/ADP ratios. The high PCr levels in cells act as an energy buffering system.

There are four major creatine kinase (CK) isoenzymes: two cytoplasmic and two mitochondrial. Cytoplasmic CK is composed of dimers formed from either or both of two polypeptides: muscle type (M) and brain type (B). MMCK and BBCK homodimers occur in muscle cells and brain cells, respectively. Heart cells have MMCK, MBCK, and BBCK. There are two forms of mitochondrial CK (mtCK): ubiquitous mtCK, found in nonmuscle cells, and sarcomeric mtCK, which occurs in muscle cells. MtCKs are enzymatically active as octomers. They are located within the intermembrane space, where they are bound to cardiolipin (an IMM phospholipid; see p. 410) linked to ANT, and VDAC (voltage-dependent anion channel, a diffusion pore for small hydrophilic

molecules) in the OMM.

In the CK/PCr shuttle (**Figure 10.20**), mtCK catalyzes the transfer of high-energy phosphoryl groups from ATP emerging from ANT to creatine (Cr) to yield PCr. Once PCr moves through VDAC into the cytoplasm, it diffuses to cellular CK microcompartments (e.g., membrane ATPases and various ATP-requiring enzymes that are closely associated with CK). Some CKs are also specifically coupled to glycolytic enzymes. Once formed, Cr may either be phosphorylated by glycolysis-linked CK or diffuse to a nearby mitochondrion, where it is rephosphorylated by mtCK. The tight functional coupling between the CK/PCr shuttle and ATP synthesis ensures that the cell's energy requirements are met. In resting cells, the total Cr levels in a cell are typically two-thirds PCr and one-third Cr, with an ATP/ADP ratio of 100:1.

The Complete Oxidation of Glucose

Table 10.2 summarizes the sources of ATP produced from one molecule of glucose. ATP production from fatty acids, the other important energy source, is discussed in Chapter 12. Recall that two molecules of NADH are produced during glycolysis. When oxygen is available, the oxidation of this NADH by the ETC is preferable (in terms of energy production) to lactate formation. The inner mitochondrial membrane, however, is impermeable to NADH. Animal cells have evolved several shuttle mechanisms to transfer electrons from cytoplasmic NADH to the mitochondrial ETC. The most prominent examples are the glycerol phosphate shuttle and the malate–aspartate shuttle.

In the **glycerol-3-phosphate shuttle** (**Figure 10.21a**), DHAP, a glycolytic intermediate, is reduced by NADH to form glycerol-3-phosphate. This reaction is followed by the oxidation of glycerol-3-phosphate by mitochondrial glycerol-3-phosphate dehydrogenase. (The mitochondrial enzyme uses FAD as an electron acceptor.) Because glycerol-3-phosphate interacts with the mitochondrial enzyme on the outer face of the inner membrane, the substrate does not actually enter the matrix. The $FADH_2$ produced in this reaction is then oxidized by the ETC. FAD as an electron acceptor produces 1.5 ATP per molecule of cytoplasmic NADH.

TABLE 10.2 Summary of ATP Synthesis from the Oxidation of One Molecule of Glucose

	NADH	$FADH_2$	ATP
Glycolysis (cytoplasm)			
Glucose → glucose-6-phosphate			−1
Fructose-6-phosphate → fructose-1,6-bisphosphate			−1
Glyceraldehyde-3-phosphate → glycerate-1,3-bisphosphate	+2		
Glycerate-1,3-bisphosphate → glycerate-3-phosphate			+2
Phosphoenolpyruvate → pyruvate			+2
Mitochondrial Reactions			
Pyruvate → acetyl-CoA	+2		
Citric acid cycle			
Oxidation of isocitrate, α-ketoglutarate, and malate	+6		
Oxidation of succinate		+2	

GDP → GTP	+1.5*
Oxidative Phosphorylation	
2 Glycolytic NADH	+4.5† (3)‡
2 NADH (pyruvate to acetyl-CoA)	+5
6 NADH (citric acid cycle)	+15
2 $FADH_2$ (citric acid cycle)	+3
	31 (29.5)

* This number reflects the price of transport into the cytoplasm.

† Assumes the malate–aspartate shuttle.

‡ Assumes the glycerol-3-phosphate shuttle.

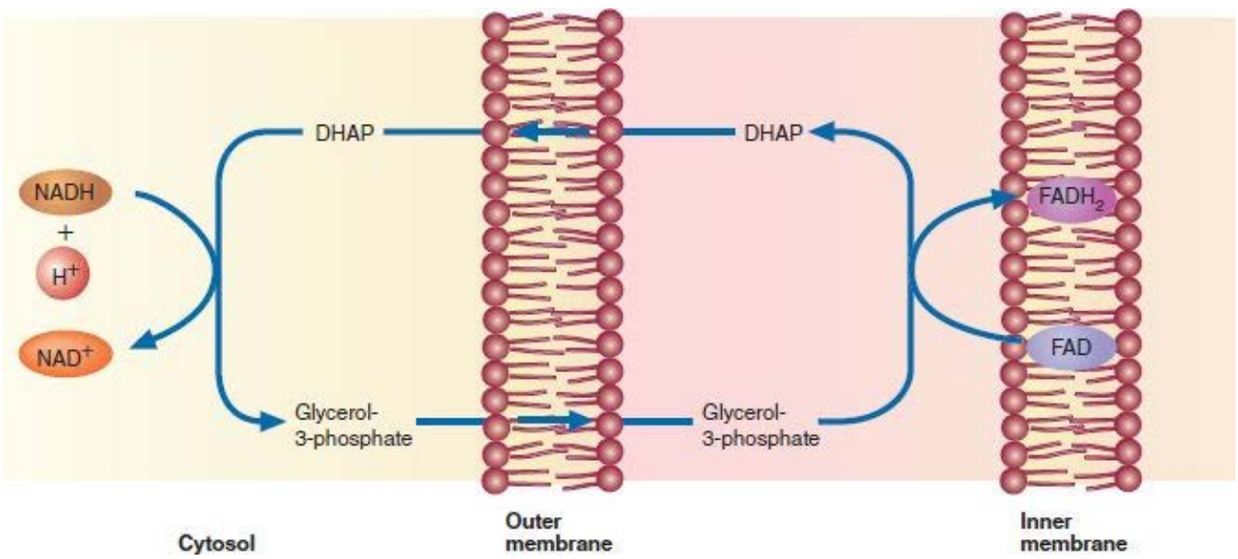


FIGURE 10.21a

Shuttle Mechanisms That Transfer Electrons from Cytoplasmic NADH to the Respiratory Chain

(a) The glycerol-3-phosphate shuttle. Dihydroxyacetone phosphate (DHAP) is reduced to form glycerol-3-phosphate. Glycerol-3-phosphate is reoxidized by mitochondrial glycerol-3-phosphate dehydrogenase and FAD is reduced to $FADH_2$. (b) The aspartate–malate shuttle. Oxaloacetate is reduced by NADH to form malate. Malate is transported into the mitochondrial matrix, where it is reoxidized to form oxaloacetate and NADH. Because oxaloacetate cannot penetrate the inner membrane, it is converted to aspartate in a transamination involving glutamate. Two inner membrane carriers are required for this shuttle mechanism: the glutamate–aspartate transport protein and the malate-α-ketoglutarate transport protein.

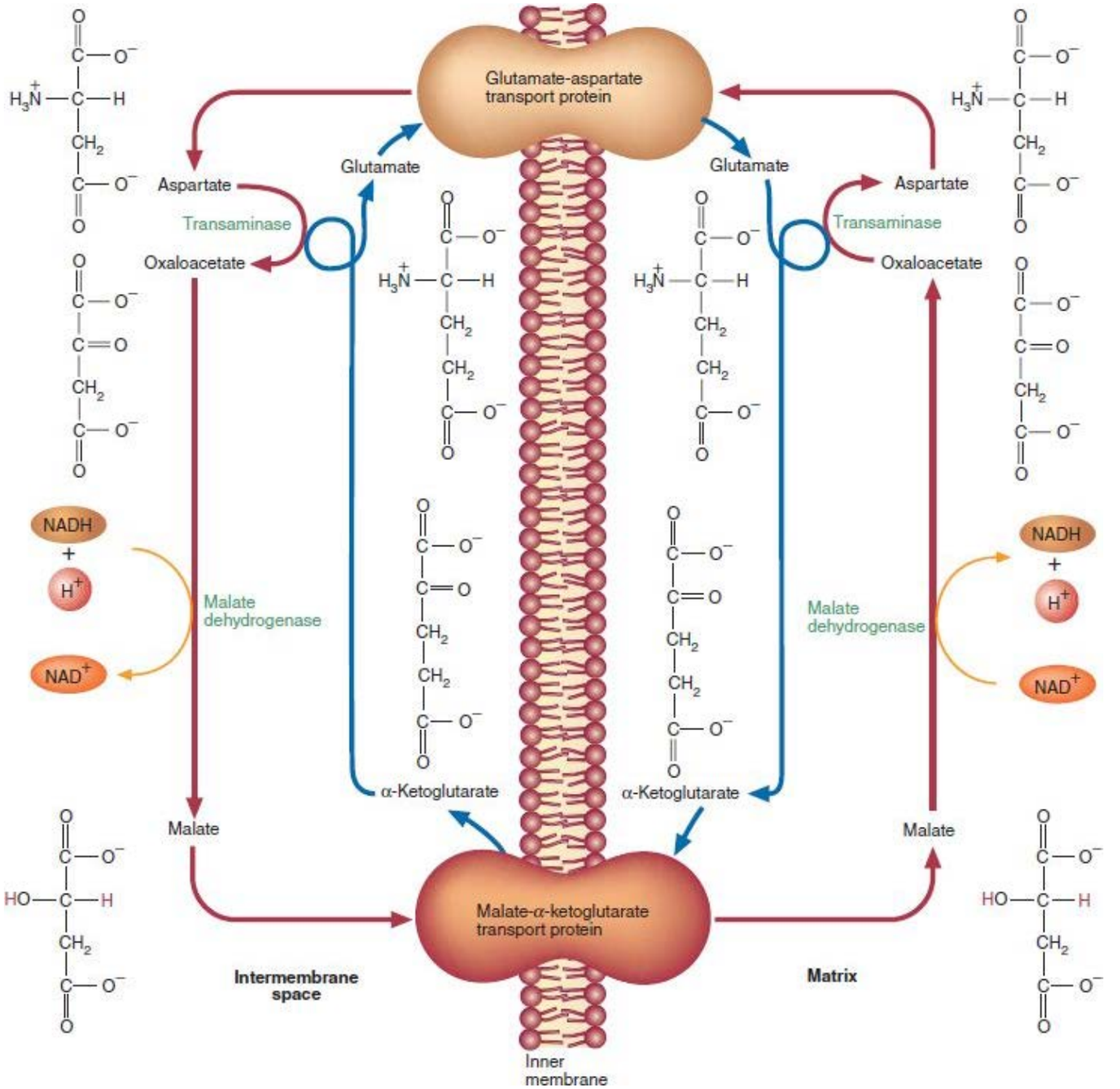


FIGURE 10.21b

Although the **malate–aspartate shuttle** (**Figure 10.21b**) is a more complicated mechanism than the glycerol phosphate shuttle, it is more energy-efficient. The shuttle begins with the reduction of cytoplasmic OAA to malate by NADH. After its transport into the mitochondrial matrix, malate is reoxidized. The NADH produced is then oxidized by the ETC. For the shuttle to continue, OAA must be returned to the cytoplasm. Because the inner membrane is impermeable to OAA, it is converted to aspartate in a transamination reaction (Chapter 14) involving glutamate.

The aspartate is transported to the cytoplasm in exchange for glutamate (via the glutamate–aspartate transport protein), where it can be converted to OAA. The α-ketoglutarate is transported to the cytoplasm in exchange for malate (via the malate-α-ketoglutarate transport protein), where it can be converted to glutamate. The glutamate–aspartate transporter requires moving a proton into the matrix. Therefore, the net ATP synthesis using this mechanism is somewhat reduced. Instead of generating 2.5 molecules of ATP for each NADH molecule, the yield is approximately 2.25 molecules of ATP.

One final issue concerned with ATP synthesis from glucose remains. Recall that two molecules of ATP are produced in the citric acid cycle (from GTP). The price for their transport into the cytoplasm, where they will be used, is the uptake of two protons into the matrix. Therefore, the total amount of ATP produced from a molecule of glucose is reduced by about half a molecule of ATP.

Depending on the shuttle used, the total number of molecules of ATP produced per molecule of glucose varies (approximately) from 29.5 to 31. Assuming that the average amount of ATP produced is 30 molecules, the net reaction for the complete oxidation of glucose is as follows:

$$C_6H_{12}O_6 + 6O_2 + 30\ ADP + 30\ P_i \rightarrow 6CO_2 + 6H_2O + 30\ ATP \quad (3)$$

The number of ATP molecules generated during the complete oxidation of glucose is in sharp contrast to the two molecules of ATP formed by glycolysis. Obviously, organisms that use oxygen to oxidize glucose have a substantial advantage.

KEY CONCEPTS

The aerobic oxidation of glucose yields between 29.5 and 31 ATP molecules.

QUESTION 10.4

Traditionally, the oxidation of each NADH and $FADH_2$ by the ETC was believed to result in the synthesis of three molecules of ATP and two molecules of ATP, respectively. As noted, recent measurements, which have considered such factors as proton leakage across the inner membrane, have reduced these values somewhat. Use the earlier values to calculate the number of ATP molecules generated by the aerobic oxidation of a glucose molecule. First, assume that the glycerol-3-phosphate shuttle is operating. Then, assume that the malate–aspartate shuttle is transferring reducing equivalents into the mitochondrion.

QUESTION 10.5

Calculate the maximum number of ATP that can be generated from a mole of sucrose.

Uncoupled Electron Transport

Certain proteins, called **uncoupling proteins**, partially dissipate oxidative energy by translocating protons across the mitochondrial inner membrane. As a result ATP synthesis is reduced. Five uncoupling proteins have been identified in humans:

- Uncoupling protein 1 (UCP1), the best-characterized example, is a dimer that forms a proton channel. Also known as *thermogenin*, UCP1 is found exclusively in the mitochondria of brown fat, a specialized form of adipose tissue present in newborn babies (to help normalize body temperature), some human adults, and hibernating mammals (to warm their bodies at the end of winter). (The characteristic color of brown adipose tissue results from the large number of mitochondria it contains.)
- UCP2, found in a wide variety of tissue cells, is used to control ROS formation.
- UCP3, restricted to skeletal muscle and brown adipose tissue, is believed to have a role in the regulation of fatty acid oxidation and decreasing ROS formation.
- UCP4 and UCP5 are expressed in central nervous system neurons where their functions have not

been resolved.

UCP1, which may constitute about 10% of the protein in the mitochondrial inner membrane, is activated when it is bound to fatty acids. As a result of the reduction of the proton gradient by UCP1, the energy captured during electron transport is dissipated as heat. The entire process of heat generation from brown fat, called *nonshivering thermogenesis*, is regulated by norepinephrine. (In shivering thermogenesis, heat is produced by nonvoluntary muscle contraction.) Norepinephrine, a neurotransmitter released from specialized neurons that terminate in brown adipose tissue, initiates a cascade mechanism that ultimately hydrolyzes fat molecules. The fatty acid products of fat hydrolysis activate the uncoupler protein. Fatty acid oxidation continues until the norepinephrine signal is terminated or the cell's fat reserves are depleted.

10.3 OXYGEN, CELL SIGNALING, AND OXIDATIVE STRESS

All living organisms utilize redox reactions in a wide range of biochemical processes, which include anabolic and catabolic metabolism and the cell-signaling mechanisms that control responses to metabolic and environmental cues (e.g., growth factors, hormones, or nutrient availability). In aerobic organisms, these reactions take place within a redox environment that must be rigorously controlled so that metabolic processes can function without oxidative damage. The reduction potential and reducing capacity of linked redox pairs such as NAD(P)H/$NAD(P)^+$ work together with low-molecular-mass reducing agents such as glutathione (GSH/GSSG, reduced and oxidized forms), linked with large numbers of redox-sensitive proteins, to maintain redox homeostasis.

The advantage of using oxygen as a terminal electron acceptor to generate large quantities of energy coexists with oxygen's dangerous property: oxygen can accept single electrons to form unstable derivatives, referred to as **reactive oxygen species (ROS)**. In healthy cells, ROS formation is kept to a minimum by antioxidants and antioxidant enzymes, which react with ROS before nearby biomolecules are damaged.

Despite their potential danger, some ROS function in small amounts as cell-signaling devices by altering the redox status and therefore the functional properties of target proteins such as metabolic enzymes, cytoskeletal components, cell cycle regulators, and transcription and translation factors. It is noteworthy that ROS function together with **antioxidants** (molecules that react easily with ROS, some by transiently becoming radicals themselves) and antioxidant enzymes to form an adaptive system that is necessary for promoting health. For example, athletes in training depend on increased ROS production to regulate the signaling pathways that promote muscle adaptation in response to increased physical activity.

Under certain conditions, referred to collectively as **oxidative stress**, antioxidant mechanisms are overwhelmed, ROS levels rise, and some damage may occur. Damage results primarily from enzyme inactivation, polysaccharide depolymerization, DNA breakage, and membrane destruction. Examples of circumstances that may cause serious oxidative damage include infection, inflammation, certain metabolic abnormalities, the overconsumption of certain drugs or exposure to intense radiation, and repeated contact with certain environmental contaminants (e.g., tobacco smoke). In addition to contributing to the aging process, oxidative damage has been linked to at least 100 human diseases. Examples include cancer, cardiovascular disorders such as atherosclerosis, myocardial infarction, and hypertension, as well as neurological disorders such as amyotrophic lateral sclerosis (ALS, or Lou Gehrig's disease), Parkinson's disease, and Alzheimer's disease.

Reactive Oxygen Species

Dioxygen is a diradical because it possesses two unpaired electrons. (A **radical** is an atom or a group of atoms that contains one or more unpaired electrons.) For this reason, when it reacts, dioxygen can accept only one electron at a time.

Recall that during mitochondrial electron transport, H_2O is formed as a consequence of the sequential transfer of four electrons to O_2. Cytochrome oxidase traps these reactive intermediates within its active site until all four electrons have been transferred to oxygen. However, several other electron transport complexes leak electrons that react with O_2 to form ROS (**Figure 10.22**).

The first ROS formed during the reduction of oxygen is the superoxide radical $O_2^{\cdot-}$. Most superoxide radicals are produced by electrons derived from the flavoprotein NADH dehydrogenase (complex I) and from complex III. However, $O_2^{\cdot-}$ acts as a nucleophile and (in specific circumstances) as either an oxidizing agent or a reducing agent. Because of its solubility properties, $O_2^{\cdot-}$ causes considerable damage to the phospholipid components of membranes. When it is generated in an aqueous environment, $O_2^{\cdot-}$ can react with itself to produce O_2 and hydrogen peroxide (H_2O_2):

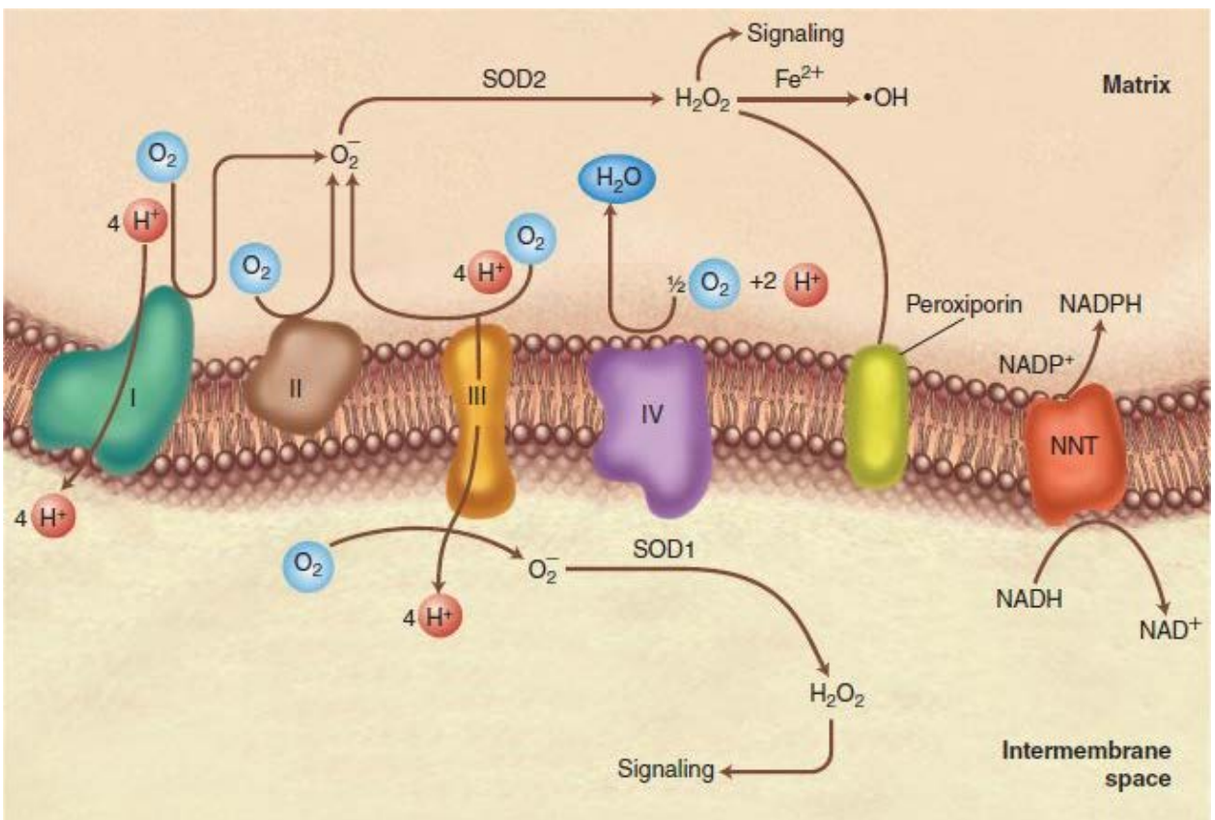


FIGURE 10.22

ROS Formation in the Mitochondrion

The mitochondrial electron transport mechanism leaks electrons to form superoxide radicals. The largest sources are complexes I and III, although complex II releases superoxide when the reoxidation of UQH_2 is prevented (p. 367). The majority of superoxide radicals are converted to H_2O_2 by superoxide dismutase (SOD: see p. 389), SOD1 in the intermembrane space, and SOD2 in the matrix. Within the matrix, H_2O_2 stimulates certain signaling mechanisms (e.g., those affecting nutrient oxidation and mitochondrial morphology). It can also react with Fe^{2+} to yield the hydroxyl radical (HO•), which can damage any molecule it collides with such as membrane lipids and

mitochondrial DNA. The matrix is protected from excess ROS by redox control enzymes (not shown; see pp. 389–90) and NNT (nicotinamide nucleotide transhydrogenase; see p. 391), an IMM protein that couples the proton gradient to the transfer of a hydride ion from NADH to $NADP^+$ to yield the antioxidant molecule NADPH. Since hydrogen peroxide is less reactive than superoxide radicals, it can easily diffuse across the IMM using channels called peroxiporins (a type of aquaporin: see p. 430) into the intermembrane space, where it can be converted to a water molecule by an antioxidant enzyme (not shown; see p. 390). H_2O_2 may also diffuse across the outer membrane into the cytoplasm, where it regulates diverse processes such as autophagy, cell proliferation, and apoptosis via redox-regulated transcription factors.

$$2H^+ + 2O_2^{\bullet-} \rightarrow O_2 + H_2O_2 \quad (4)$$

Most superoxide radicals, however, are converted to H_2O_2 by superoxide dismutase (SOD) (p. 389). Other smaller sources of superoxide/hydrogen peroxide include the pyruvate and α-ketoglutarate dehydrogenases and ETF:QO. Since H_2O_2 does not have any unpaired electrons, it is not a radical. The limited reactivity of H_2O_2 allows it to function as an important signal molecule and to cross membranes and become widely dispersed. The subsequent reaction of H_2O_2 with Fe^{2+} (or other transition metals) results in the production of the hydroxyl radical (•OH), a highly reactive species:

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + {}^{\bullet}OH + OH^- \quad (5)$$

The hydroxyl radical diffuses only a short distance before it reacts with whatever biomolecule it collides with. The hydroxyl radical is especially dangerous because it can initiate an autocatalytic radical chain reaction (**Figure 10.23**). Singlet oxygen ($^1O_2$), an excited state of dioxygen in which the unpaired electrons have become paired, can form from superoxide:

$$2O_2^{\bullet-} + 2H^+ \rightarrow H_2O_2 + {}^1O_2 \quad (6)$$

or from peroxides:

$$2ROOH \rightarrow 2ROH + {}^1O_2 \quad (7)$$

Singlet oxygen formed from certain reactions of H_2O_2 and during photosynthetic light harvesting can react with double bonds in biomolecules. It is particularly damaging to molecules with aromatic rings or conjugated carbon-carbon double bonds

As mentioned earlier (see p. 384), ROS are generated during several other cellular activities besides the reduction of O_2 to form H_2O. These include the biotransformation of xenobiotics within the smooth endoplasmic reticulum (SER) of liver and lung cells (pp. 485–86), the **respiratory burst** within white blood cells (**Figure 10.24**), and disulfide bridge formation in newly synthesized proteins within the RER (p. 779). Another such reaction, carbonylation, is a nonenzymatic protein modification resulting from the oxidation of amino acid side chains (i.e., Thr, Lys, Arg, or Pro) or the reaction of Cys, Lys, or His side chains with reactive carbonyl radicals.

There are also several nitrogen-containing radicals. Because their synthesis is linked to that of ROS, **reactive nitrogen species** (RNS) are often classified as ROS. Among the most important examples are nitric oxide ($^{\bullet}NO$), nitrogen dioxide ($^{\bullet}NO_2$), and peroxynitrite ($ONOO^-$). Nitric oxide ($^{\bullet}NO$) is a highly reactive gas that is an important signal molecule produced throughout the mammalian body.

Nitric Oxide Damage

Physiological functions in which $^{\bullet}NO$ is now believed to play a role include blood pressure regulation, blood-clotting inhibition, and macrophage-induced destruction of foreign, damaged, or cancerous cells. The disruption of the normally precise regulation of $^{\bullet}NO$ synthesis has been linked to numerous pathological conditions that include stroke, migraine headache, male impotence, septic shock, and several neurodegenerative diseases such as Parkinson's disease. $^{\bullet}NO$ can damage proteins with sulfhydryl groups, such as glyceraldehyde-3-phosphate dehydrogenase (GAPDH; p. 288), by converting SH groups into nitrosothiol (—SNO) derivatives. GAPDH nitrosylation, in response to apoptosis stimuli, results in movement of the protein into the nucleus, where it plays a critical role in cell death. $^{\bullet}NO$ also damages iron–sulfur proteins. Some of the damage attributed to $^{\bullet}NO$ is, in fact, caused by its oxidation products, $^{\bullet}NO_2$ ($2\ ^{\bullet}NO + O_2 \rightarrow 2\ ^{\bullet}NO_2$) and peroxynitrite ($^{\bullet}NO + O_2 \rightarrow ONOO^-$).

KEY CONCEPTS

- Reactive oxygen species form because oxygen is reduced by accepting one electron at a time.
- ROS formation is a normal by-product of metabolism and the result of conditions such as exposure to radiation.
- RNS are often classified as ROS because the synthesis of these species is linked.

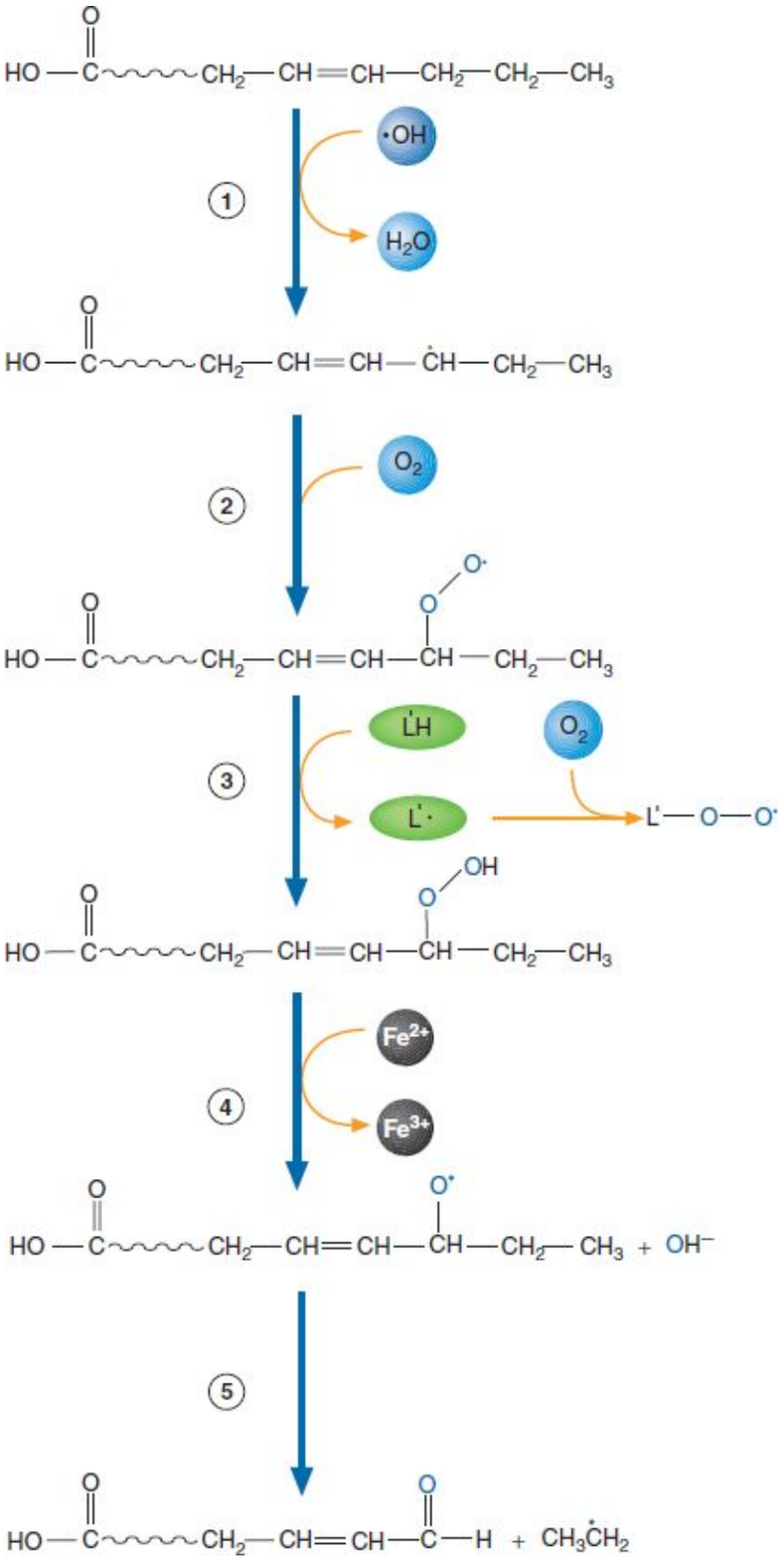


FIGURE 10.23

Radical Chain Reaction

Step 1: Lipid peroxidation reactions begin after the extraction of a hydrogen atom from an unsaturated fatty acid (LH → L•). Step 2: The lipid radical (L•) then reacts with O_2 to form a peroxyl radical (L• + O_2 → L−O—O•). Step 3: The radical chain reaction begins when the peroxyl radical extracts a hydrogen atom from another fatty acid molecule (L—O—O• + L′H → L—O—OH + L′•). Step 4: The presence of a transition metal such as Fe^{2+} initiates further radical formation (L—O—O—H + Fe^{2+} → LO• + HO^- + Fe^{3+}). Step 5: One of the most serious consequences of lipid peroxidation is the formation of *α*,*β*-unsaturated aldehydes, which involves a radical cleavage reaction. The chain reaction continues as the free radical product then reacts with a nearby molecule.

Reactive carbonyl products are also products of this process.

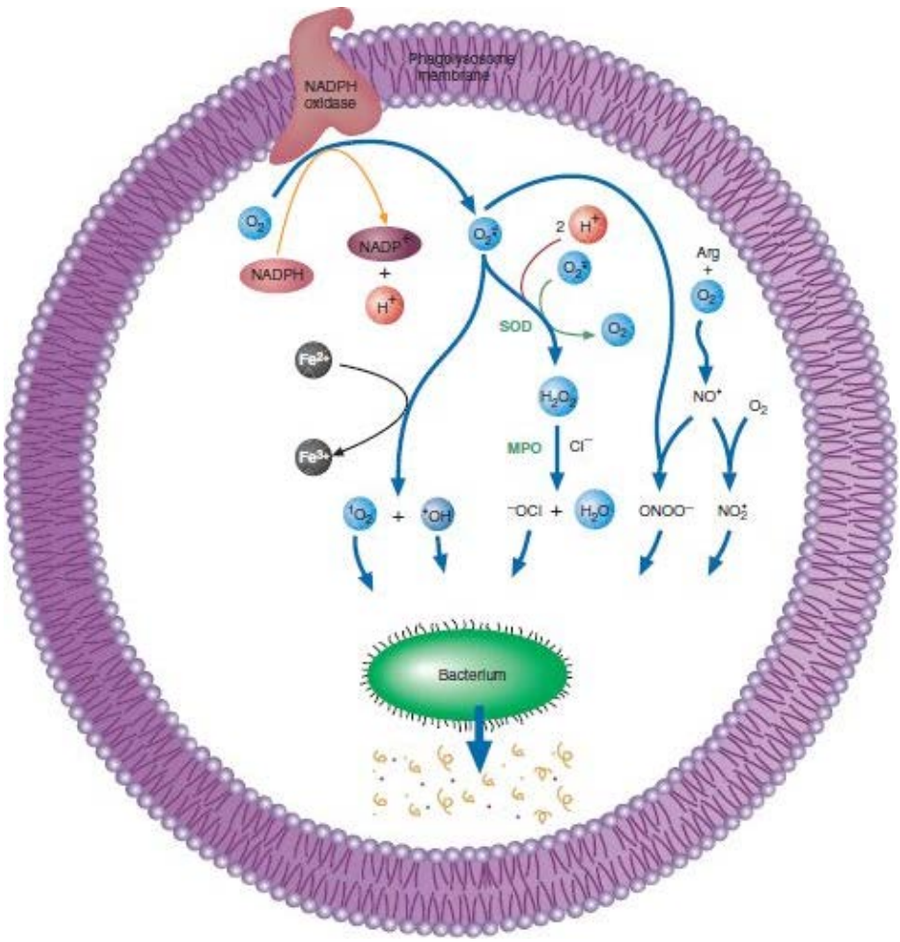


FIGURE 10.24

The Respiratory Burst

The respiratory burst provides a dramatic example of the destructiveness of ROS. Within seconds after a phagocytic cell such as a macrophage binds to a bacterium (or other foreign structure), its oxygen consumption increases nearly 100-fold. During endocytosis, the bacterium is incorporated into a large vesicle called a phagosome. Phagosomes then fuse with lysosomes to form phagolysosomes. Two destructive processes then ensue: the respiratory burst and digestion by lysosomal enzymes. The respiratory burst is initiated when the phagolysosomal membrane enzyme NADPH oxidase converts O_2 to $O_2^{\cdot-}$. Two molecules of $O_2^{\cdot-}$ combine in a reaction catalyzed by SOD (superoxide dismutase) to form H_2O_2. H_2O_2 is next converted to several types of *bactericidal* (bacteria-killing) molecules by myeloperoxidase (MPO), an enzyme found in abundance in phagocytes. For example, MPO catalyzes the oxygenation of halide ions (e.g., Cl^-) to form hypohalides. Hypochlorite (ClO^-, the active ingredient in household bleach) is extremely bactericidal. In the presence of Fe^{2+}, $O_2^{\cdot-}$ and H_2O_2 react to form •OH and $^1O_2$ (singlet oxygen), both of which are extremely reactive. Nitric oxide, synthesized by nitric oxide synthase from arginine and O_2 (p. 552), reacts with superoxide to form peroxynitrite and with molecular oxygen to form nitrogen dioxide. In addition to the damage inflicted by various types of ROS, these species, along with MPO products, activate proteases, which degrade microbial proteins. Proteases are themselves protected from oxidative injury by MPO, which also possesses a catalase activity. After the disintegration of the bacterial cell, lysosomal enzymes digest the fragments that remain.

Antioxidant Enzyme Systems

To protect themselves from oxidative stress, living organisms have developed several preventive strategies that include melanin, the skin pigment that protects against photooxidative stress, and chromatin, DNA/protein packaging in eukaryotic cells that provides some protection for DNA from oxidative attack. Organisms protect themselves from ROS once they are formed at higher than physiological levels by using powerful antioxidant enzymes. Oxidative biomolecular damage can be repaired by DNA repair mechanisms and autophagy, which are described on pp. 700–04 and 574–75, respectively. Since no biological process is perfect, unrepaired oxidative damage does accumulate over long periods, often contributing to a diverse group of degenerative diseases.

The major enzymatic defenses against oxidative stress are provided by four enzymes: superoxide dismutase, glutathione peroxidase, peroxiredoxin, and catalase. The wide distribution of these enzymatic activities underscores the ever-present problem of oxidative damage.

The superoxide dismutases (SODs) are a class of enzymes that catalyze the formation of H_2O_2 and O_2 from the superoxide radical:

$$2O_2^{\cdot-} + 2H^+ \rightarrow H_2O_2 + O_2 \qquad (8)$$

There are three major forms of SOD in humans. (1) SOD1 is a Cu–Zn isoenzyme that occurs in cytoplasm and the mitochondrial intermembrane space. (2) SOD3, also requiring Cu and Zn, is an extracellular enzyme. (3) A manganese-containing isozyme, SOD2, is found in the mitochondrial matrix. About 20% of inherited ALS are caused by a mutation in the gene that codes for SOD1, allowing it to accumulate in the intermembrane space where its H_2O_2 synthesis disrupts mitochondrial function. ALS is a fatal degenerative disease in which motor neurons are destroyed.

ALS

Glutathione peroxidase (GPx), a selenium-containing enzyme, is a key component in an enzymatic system most responsible for controlling cellular peroxide levels. Recall that this enzyme catalyzes the reduction of a variety of substances by the tripeptide reducing agent GSH (**Table 5.3**). In addition to reducing H_2O_2 to form water, glutathione peroxidase transforms organic peroxides into alcohols:

$$2\,GSH + R{-}O{-}O{-}H \rightarrow GSSG + R{-}OH + H_2O \qquad (9)$$

In humans, the cytoplasmic enzyme GPx1 reduces H_2O_2. GPx4 acts on lipid hydroperoxides. Several ancillary enzymes support glutathione peroxidase function (**Figure 10.25**). GSH is regenerated from GSSG by glutathione reductase:

$$GSSG + NADPH + H^+ \rightarrow 2\,GSH + NADP^+ \qquad (10)$$

The NADPH required in the reaction is provided primarily by two reactions of the pentose phosphate pathway (Chapter 8). Recall that NADPH is also produced by the reactions catalyzed by isocitrate dehydrogenase (p. 345) and malic enzyme (p. 352).

The *peroxiredoxins* (PRX) are a class of thiol-containing enzymes that detoxify peroxides. Their catalytic mechanism involves the oxidation of a redox-active cysteine side chain sulfhydryl group by the peroxide substrate to form sulfenic acid (RSOH). The sulfenic acid residue of peroxiredoxin is subsequently reduced by a thiol-containing protein such as thioredoxin. *Thioredoxin* (TRX) is involved in redox reactions mediated by the peroxiredoxin/thioredoxin reductase (TR) system (sometimes referred to as the TRX-centered system) (**Figure 10.26**). TR and TRX also reduce other oxidized cellular proteins, including many transcription factors, to their functional reduced sulfhydryl

form; this change is accomplished by means of an enzyme-catalyzed shift of electrons from reduced thioredoxin [TRX-$(SH)_2$] to the target protein. Thioredoxin reductase reduces oxidized thioredoxin [TRX(S_2)], with electrons delivered to it through a mobile NADPH and a bound $FADH_2$. TRX also serves as an electron shuttle for other nonantioxidant enzyme systems such as that of ribonucleotide reductase (p. 560).

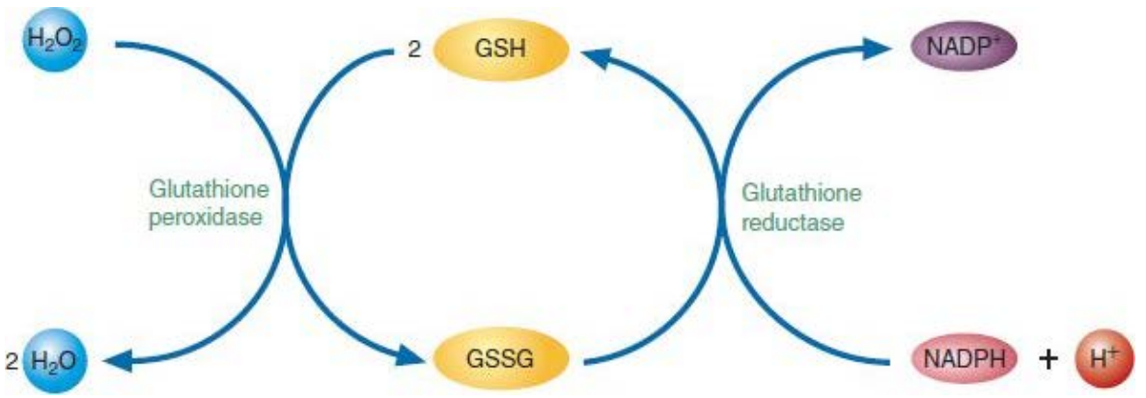


FIGURE 10.25

The Glutathione-Centered System

Glutathione peroxidase uses GSH to reduce the peroxides generated by cellular aerobic metabolism. GSH is regenerated from its oxidized form, GSSG, by glutathione reductase. NADPH, the reducing agent in this reaction, is supplied by the pentose phosphate pathway and several other reactions.

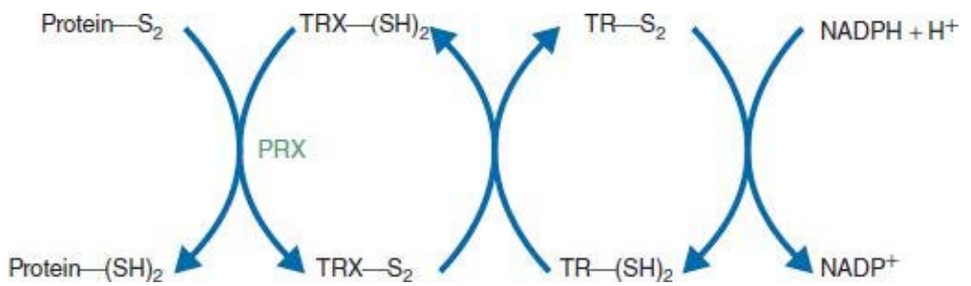


FIGURE 10.26

The Thioredoxin-Centered System

A substrate protein that contains function-altering disulfides is reduced by electron transfer from thioredoxin [TRX-$(SH)_2$] in a reaction mediated by the peroxiredoxin enzyme (PRX). The oxidized thioredoxin (TRX-S_2) is returned to its reduced form by thioredoxin reductase (TR). The electrons required to restore the TR to its reduced form come from NADPH via an enzyme-bound FAD.

Catalase is an enzyme, usually located in peroxisomes, whose primary role is to catalyze the conversion of H_2O_2 to water and dioxygen. (For every two molecules of H_2O_2 that are detoxified, one is oxidized to form O_2 and the other is reduced to H_2O.) Catalase has a heme-Fe(III) prosthetic group with the following mechanism:

$$\text{Step 1: } H_2O_2 + \text{Fe(III)—enzyme} \rightarrow H_2O + \text{O=Fe(IV)—enzyme} \qquad (11)$$

$$\text{Step 2: } H_2O_2 + \text{O=Fe(IV)—enzyme} \rightarrow O_2 + H_2O + \text{Fe(III)—enzyme}$$

This peroxidative reaction of catalase can also convert organic molecules such as alcohols and aldehydes to an often less toxic form:

$$R\text{—}CH_2\text{—}OH + \text{O=Fe(IV)—enzyme} \rightarrow R\text{—}CH_3 + O_2 + \text{Fe(III)—enzyme} \qquad (12)$$

KEY CONCEPTS

- The major enzymatic defenses against oxidative stress are superoxide dismutase, glutathione peroxidase, peroxiredoxin, and catalase.
- The pentose phosphate pathway produces the reducing agent NADPH.

The respiratory burst of macrophages, for example, generates H_2O_2 as a microbicide primarily through the incomplete oxidation of fatty acids. Catalase is found in high quantity in erythrocytes and peroxisomes of phagocytic white blood cells. Outside the peroxisome, the primary H_2O_2 generator is SOD1.

QUESTION 10.6

Selenium is generally considered a toxic element. (It is the active component of loco weed.) However, there is growing evidence that selenium is also an essential trace element. Because glutathione peroxidase activity is crucial to the protection of red blood cells against oxidative stress, a selenium deficiency can damage red blood cells. Although sulfur is in the same chemical family as selenium, it cannot be substituted. Can you explain why? [*Hint*: Selenium is more easily oxidized than sulfur.] Is sulfur or selenium a better scavenger for oxygen when this gas is present in trace amounts?

QUESTION 10.7

Ionizing radiation is believed to damage living tissue by producing hydroxyl radicals. Drugs that protect organisms from radiation damage, which must be taken *before* radiation exposure, usually have —SH groups. How do such drugs protect against radiation damage? Can you suggest any type of nonsulfhydryl group–containing molecule that would protect against hydroxyl radical-induced damage?

Radiation Damage

The Redox Proteome

The redox proteome is a large, diverse group of proteins that undergo reversible redox reactions linking electron transfers to metabolic processes. Of the three amino acid side chains that undergo reversible reactions (cysteine, methionine, and selenocysteine), proteins with ionizable cysteine thiol groups are the best researched. Such proteins are referred to collectively as the *cysteine proteome*. Although the pK_a of the thiol group of cysteine is approximately 8, redox active cysteine residues occur in protein microenvironments where thiol pK_a values are as low as 4 or 5, thereby promoting the formation of the thiolate anion (Cys-S^-). Thiol oxidation can occur with the transfer of one or two electrons to form the thiyl radical (Cys-S•) or sulfenic acids (Cys-SOH) and disulfides (Cys-S-S-R′). Higher oxidation states such as sulfinic acids (Cys-SO_2H) and sulfonic acids (Cys-SO_3H) may not be reversible. Current understanding of the relationship between the cysteine proteome, metabolic processes, and oxidative state of cells can be described by the redox code.

The **redox code** is a set of four principles that together describe the relationships between $NAD(P)^+$/NAD(P)H, GSH/GSSG, the cysteine proteome, and other redox molecules in space (i.e., subcellular structures) and time, in processes such as cell differentiation, organism development, and adaptation to environmental change:

1. Reversible high-flux electron accepting and donating by NAD and NADP is the central organizing feature of metabolism: all substrate redox reactions are linked directly or indirectly to the reduction of either of these molecules. The NADH system is involved in catabolism and energy capture, whereas the NADPH system drives anabolic pathways such as fatty acid synthesis and functions as an antioxidant in the control of oxidative reactions. Cellular redox homeostasis is maintained in part by nicotinamide nucleotide transhydrogenase (NNT), an IMM enzyme that uses the protonmotive force to transfer electrons from NADH to $NADP^+$ to yield NADPH.
2. Metabolic processes are linked to protein structure and function by kinetically controlled *redox switches*; in the cysteine proteome, these are sulfur redox switches. For example, a family of protein tyrosine phosphatases regulates the activity of tyrosine kinase receptors (cell-surface receptors for growth factors such as epidermal growth factor and hormones such as insulin). The ligand-binding process causes phosphorylation of tyrosines in the intracellular domains of these receptors, which initiates intracellular signal transduction pathways. Such signals are terminated when the phosphatase catalyzes dephosphorylation reactions initiated by its active-site cysteine thiolate anion. When H_2O_2 levels rise in proximity to the phosphatase, the active-site cysteine is oxidized, thereby inactivating the enzyme and causing signal transduction to continue until the active-site cysteine oxidation is reversed.
3. Cells depend on the capacity to sense (identify the presence of) oxidant molecules in both developmental processes (e.g., activation/deactivation cycles of H_2O_2 in complex processes such as mammalian development) and oxidative stress conditions. The best-characterized example of a *redox sensor* is OxyR, a transcriptional regulator in bacteria. When H_2O_2 levels begin to rise, the oxidation of a specific cysteine residue in OxyR by a hydrogen peroxide molecule yields a sulfenic acid derivative, which then forms a disulfide bridge with another specific cysteine side chain in the protein. The resulting change in the protein's structure allows binding to DNA, thereby initiating the transcription of antioxidant genes and the subsequent synthesis of antioxidant proteins.
4. Living organisms possess complex redox networks that support adaptation to environmental change (e.g., nutrient availability) and stress (e.g., physical damage or exposure to radiation). In the early stage of the inflammatory phase of normal wound healing, for example, large amounts of ROS (superoxide and peroxynitrite), produced by immune cells, destroy invading bacteria. Low levels of H_2O_2 are then used in various redox-sensitive processes required in wound healing such as blood coagulation, angiogenesis (new blood vessels), and new skin cell growth to form a barrier between the wound and the environment. Enzymes such as SOD, GPX, and PRX detoxify excess H_2O_2 to prevent excessive inflammation.

Antioxidant Nutrients

Living organisms use exogenous antioxidant molecules to protect themselves from radicals. Humans obtain α-tocopherol (vitamin E), ascorbic acid (vitamin C), and β-carotene (**Figure 10.27**) from their diet.

***α*-Tocopherol**, a potent radical scavenger, belongs to a class of compounds referred to as *phenolic antioxidants*. Phenols are effective antioxidants because the radical products of these molecules are resonance-stabilized and thus are relatively unreactive:

(13)

Because vitamin E (found in vegetable and seed oils, whole grains, and green, leafy vegetables) is lipid-soluble, it plays an important role in protecting membranes from lipid peroxyl radicals.

***β*-Carotene**, found in yellow-orange and dark green fruits and vegetables such as carrots, sweet potatoes, broccoli, and apricots, is a member of a class of plant pigment molecules referred to as the *carotenoids*. In plant tissue, the carotenoids absorb some of the light energy used to drive photosynthesis and protect against the ROS that form at high light intensities. In animals, *β*-carotene is a precursor of retinol (vitamin A) and an important antioxidant in membranes. (Retinol is a precursor of retinal, the light-absorbing pigment found in the rod cells of the retina.)

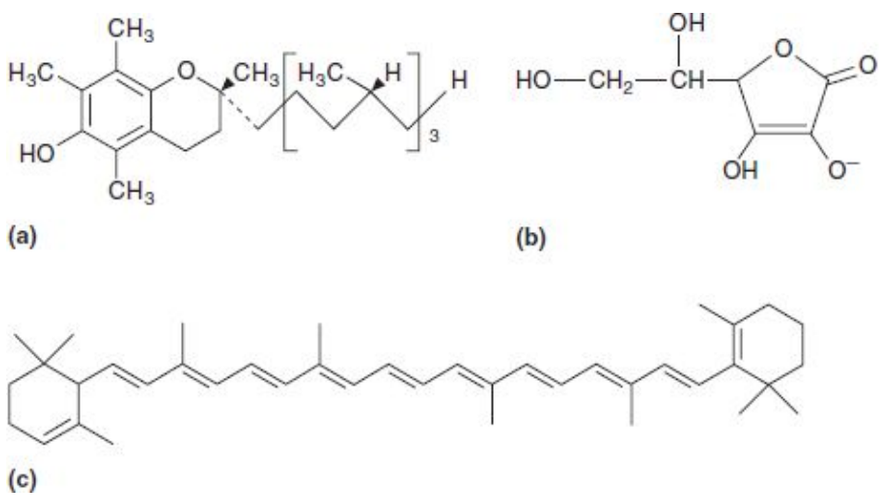
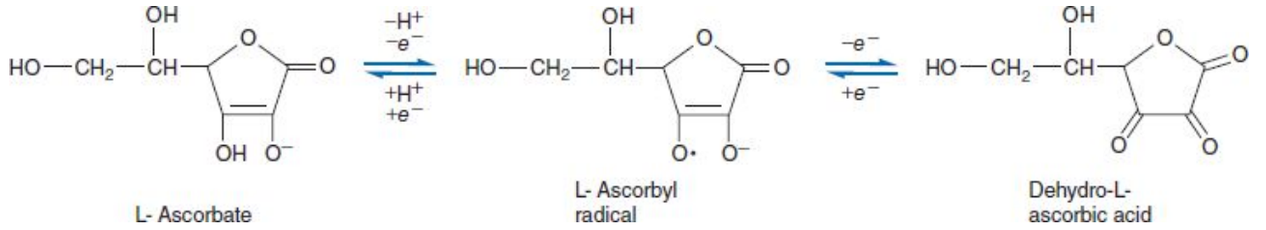


FIGURE 10.27

Selected Antioxidant Molecules

(a) *α*-tocopherol (vitamin E). (b) Ascorbate (vitamin C). (c) *β*-carotene.

Ascorbic acid is an efficient antioxidant. Present largely as ascorbate, this water-soluble molecule scavenges a variety of ROS within the aqueous compartments of cells and in extracellular fluids. Ascorbate is reversibly oxidized as shown:



Ascorbate protects membranes through two mechanisms. First, ascorbate prevents lipid peroxidation by reacting with peroxyl radicals formed in the cytoplasm before they can reach the membrane. Second, ascorbate enhances the antioxidant activity of vitamin E by regenerating reduced α-tocopherol from the α-tocopheroxyl radical (**Figure 10.28**). Ascorbate is then regenerated by reacting with GSH.

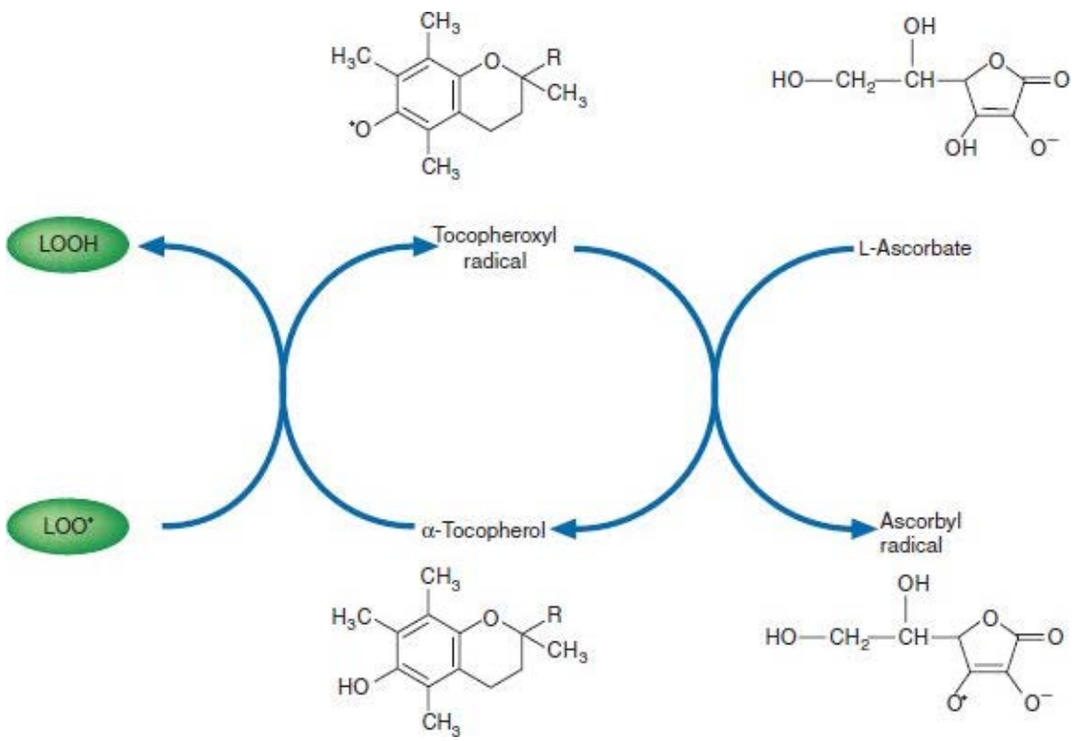


FIGURE 10.28

Regeneration of α-Tocopherol by L-Ascorbate

L-ascorbate, a water-soluble molecule, protects membranes from oxidative damage by regenerating α-tocopherol from an α-tocopheroxyl radical. The ascorbyl radical formed in this process is reconverted to L-ascorbate during a reaction with GSH.

It is noteworthy that in well-nourished individuals, the consumption of excessive amounts of antioxidant supplements renders the body's cells more vulnerable to oxidative stress. In small quantities, ROS act as signal molecules. When cells are experiencing oxidative stress (e.g., infection or inflammation), ROS levels begin to rise. Early in this process, ROS oxidize and/or covalently modify the sulfur groups of transcription factors, thereby triggering the expression of dozens of genes that strengthen the antioxidant defenses of the cell. In addition to increased concentrations of catalase, SOD, and other antioxidant enzymes, other stress proteins are produced. If cells contain excessive amounts of antioxidant molecules, obtained from dietary supplements, ROS-triggered defense mechanisms are compromised.

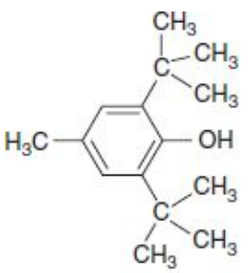
KEY CONCEPTS

- Antioxidant molecules protect cell components from oxidative damage.
- Prominent antioxidants include GSH and the dietary components α-tocopherol, β-carotene, and ascorbic acid.

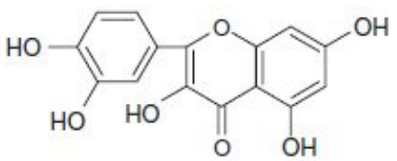
BHT as Preservative

QUESTION 10.8

The antioxidant BHT (butylated hydroxytoluene) is widely used as a food preservative. Quercitin is a member of a large group of potent antioxidants found in fruits and vegetables called the flavonoids.



BHT



Quercitin

What structural characteristic of these molecules is responsible for their antioxidant properties?

Biochemistry IN PERSPECTIVE

Myocardial Infarct: Ischemia and Reperfusion

How are heart cells damaged by the inadequate nutrient and oxygen flow caused by blood clots, and why does the reintroduction of O_2 cause further damage? The initial tissue damage that occurs during a myocardial infarction (heart attack) is caused by *ischemia*, a condition in which there is inadequate blood flow. Heart attacks are usually caused by atherosclerosis accompanied by blood-clot formation in the carotid artery. In atherosclerosis, soft masses of fatty material called *plaques* are formed in the linings of blood vessels. Unlike skeletal muscle, which is fairly resistant to ischemic injury, the cardiac muscle of the heart is extremely sensitive to hypoxic (low-oxygen) conditions. The shift from oxygen-requiring fatty acid oxidation to anaerobic glycolysis, which leads to lactate production and acidosis, is an early response of cells to ischemia. Under hypoxic conditions, fatty acid oxidation, which normally provides at least half of the heart's energy, is depressed. As a result, flux through glycolysis is increased because of lower levels of modulators such as citrate (pp. 298 and 342). Because energy production by glycolysis is inefficient, ATP levels begin to fall. As they do so, adenine nucleotides are degraded to form hypoxanthine (Chapter 15).

Without sufficient ATP, cardiac cells cannot maintain appropriate intracellular ion concentrations. For example, cytoplasmic calcium levels rise as a result of the depressed activity of the plasma membrane's Ca^{2+} ATPase. One consequence of this circumstance is the activation of calcium-dependent proteases such as the calpains, non-lysosomal Ca^{2+}-dependent cysteine proteases (enzymes with an active-site nucleophilic cysteine thiol) that cleave numerous cytoskeletal and membrane proteins.

ER stress is another important feature of hypoxic tissue. In normal circumstances, the ER is an oxidizing environment that promotes protein folding and disulfide bond formation. In addition, a high level of calcium ions supports the functions of the ER molecular chaperones. Under hypoxic and ER Ca^{2+} depletion conditions, protein folding is compromised and the unfolded protein response (UPR, p. 49) is triggered. If oxygen deprivation is not prolonged, UPR-triggered stress genes may allow an adaptive response that protects affected cells from some of the damage caused by the reintroduction of O_2. Severe or prolonged ER stress, however, may result in cell death.

The reoxygenation of an ischemic tissue, a process referred to as *reperfusion*, can be a life-saving therapy. For example, the use of streptokinase to digest artery-occluding clots in heart attack patients, accompanied by administration of oxygen, has saved many lives. However, depending on the duration of the hypoxic episode, the reintroduction of oxygen to ischemic tissue may also result in further damage.

Reperfusion injury, cellular damage caused by the reestablishment of the blood supply, results from a combination of numerous factors. Among the most important of these are ROS production and the opening of the mitochondrial permeability transition pore (mPTP). mPTP, a nonspecific channel that forms where the inner and outer mitochondrial membranes meet, allows the passage of molecules smaller than 1500 Da. Upon the reintroduction of O_2 and nutrients, ROS are first generated by reenergized mitochondrial ETC and soon by NADPH oxidase and xanthine oxidase (from neutrophils attracted to the injury site). The release of iron from cell components such as myoglobin, which can result from ROS-inflicted damage, can cause additional production of •OH. Finally, the acidosis caused by lactate accumulation in compromised heart muscle cells unloads abnormally high amounts of oxygen from hemoglobin. This latter condition greatly facilitates further ROS synthesis. Reperfusion also promotes the synthesis of large amounts of $NO^{\bullet}$, which react with superoxide to produce damaging peroxynitrite. $NO^{\bullet}$ also modifies oxidizable cysteine residues in the Ca^{2+} channels of the sarcoplasmic reticulum (SR, smooth ER in muscle cells), another factor contributing to ER stress. mPTP opening, triggered by ROS in combination with low ATP levels and high calcium levels (caused in part by ROS-inflicted damage of the SR Ca^{2+}-ATPase), leads to mitochondrial membrane potential collapse and mitochondrial swelling (caused by osmotic pressure). The release of cyt c from damaged mitochondria contributes to the initiation of apoptosis (p. 58).

SUMMARY Damage to heart cells as a result of oxygen deprivation originates with inefficient energy production, followed by osmotic pressure increases, lysosomal breakage, ER stress, and high cytoplasmic calcium levels. The reperfusion of damaged cells with O_2 leads to ROS formation, causing further damage.

Chapter Summary

1. Dioxygen (O_2), generally referred to as oxygen, is used by aerobic organisms as a terminal electron acceptor in energy generation. Several physical and chemical properties of oxygen make it suitable for this role. In

addition to its ready availability (it occurs almost everywhere on the Earth's surface), oxygen diffuses easily across cell membranes. Oxygen is a reactive diradical and an excellent oxidizing agent, readily accepting electrons from other species.

2. The NADH and $FADH_2$ molecules produced in glycolysis, the β-oxidation pathway, and the citric acid cycle generate usable energy in the electron transport pathway. The pathway consists of a series of redox carriers that receive electrons from NADH and $FADH_2$. At the end of the pathway the electrons, along with protons, are donated to oxygen to form H_2O.
3. During the oxidation of NADH, there are three steps in which the energy loss is sufficient to account for ATP synthesis. These steps occur within complexes I, III, and IV of the ETC.
4. Oxidative phosphorylation is the mechanism by which electron transport is coupled to the synthesis of ATP. According to the chemiosmotic theory, the creation of a proton gradient that accompanies electron transport is coupled to ATP synthesis.
5. The complete oxidation of a molecule of glucose results in the synthesis of 29.5 to 31 molecules of ATP, depending on whether the glycerol-3-phosphate shuttle or the malate–aspartate shuttle transfers electrons from cytoplasmic NADH to the mitochondrial ETC.
6. The use of oxygen by aerobic organisms is linked to the production of ROS. These species form because the diradical oxygen molecule accepts electrons one at a time. Examples of ROS include the superoxide radical, hydrogen peroxide, the hydroxyl radical, and singlet oxygen. Prominent RNSs include nitric oxide, nitrogen dioxide, and peroxynitrite.
7. The redox proteome is a diverse group of proteins that undergo reversible redox reactions that link electron transfer reactions to metabolic processes. Proteins with ionizable cysteine thiol groups are referred to as the cysteine proteome. The redox code is a set of principles that describe the relationships between NAD(P)+/NAD(P)H, GSH/GSSG, the cysteine proteome, and other redox molecules involved in numerous cellular processes.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on electron transport and oxidative phosphorylation to help you prepare for exams.

Chapter 10 Review Quiz

Suggested Readings

Bianchi VE. 2016. Reactive oxygen species: health and longevity. AIMS Mol Sci 3(4):479–504.

Frezza C. 2017. Mitochondrial metabolites: undercover signaling molecules. Interface Focus [Royal Society] 7(2):20160100.

Jones DP, Sies, H. 2015. The redox code. Antioxidants Redox Signaling 23(9):734–46.

Krengel U, Tornroth-Horsefield S. 2015. Coping with oxidative stress. Science 347:125–6.

Moyer MW. 2013. The myth of antioxidants. Sci Amer 308:62–7.

Nicholls DG, Ferguson SJ. 2013. Bioenergetics 4. London (UK): Academic Press.

Ursini F, et al. 2016. Redox homeostasis: the golden mean of healthy living. Redox Biology 8:205–15.

Wallimann T, Harris R. 2016. Creatine: a miserable life without it. Amino Acids 48:1739–50.

Key Words

antioxidant, *384*

β-carotene, *392*

chemiosmotic coupling theory, *374*
glycerol phosphate shuttle, *380*
ionophore, *374*
malate–aspartate shuttle, *382*
oxidative phosphorylation, *374*
oxidative stress, *384*
protonmotive force, *374*
Q cycle, *369*
radical, *385*
reactive nitrogen species (RNS), *386*
reactive oxygen species (ROS), *384*
redox code, *391*
respirasome, *372*
respiratory burst, *386*
respiratory control, *378*
α-tocopherol, *392*
uncoupler, *374*
uncoupling protein, *383*

Review Questions

SECTION 10.1

Comprehension Questions

1. Define the following terms:
    a. Fe–S clusters
    b. electron transport
    c. CoQ
    d. ubisemiquinone
    e. UQH_2
2. Define the following terms:
    a. ETF
    b. ETF:QO
    c. ShdA
    d. acyl CoA dehydrogenase complex
    e. cyt c_{ox}
3. Define the following terms:
    a. complex I
    b. complex II
    c. complex III
    d. complex IV
    e. ETC
4. Define the following terms:

a. respirasome
b. antimycin
c. CN^-
d. IMS
e. IMM

Fill in the Blanks

5. Complexes I and II of the electron transport complex transfer electrons from NADH and succinate, respectively, to _____.
6. The functional respiration unit within the mitochondria of plants and animals is called the _____.
7. _____ is a small matrix protein that functions as an electron acceptor for dehydrogenases in fatty acid oxidation.
8. _____ catalyzes the redox reaction between UQH_2 and cytochrome c.
9. Electron transport within complex I begins with the reduction of _____ by NADH.

Short-Answer Questions

10. What are the principal sources of electrons for the electron transport pathway?
11. List several reasons why oxygen is widely used by living organisms in energy production.
12. List the protein complexes in the mitochondrial ETC and describe their functions.
13. The electron transport complex consists of a series of oxidations rather than one reaction. Why is this an important feature of energy capture?
14. What metabolites accumulate when azide is added to actively respiring mitochondria?

Critical-Thinking Questions

15. The reduction potentials of iron in each of the cytochromes in the electron transport complexes vary from 0.1 V to 0.39 V. Explain why these different values are necessary for the operation of this process.
16. Suppose that the cytochrome complexes were not embedded within the mitochondrial inner membrane. According to the chemiosmotic theory, what would be the consequences?
17. Explain why an inhibitor of complex I will cause an increase not only in the ratio of NADH to NAD^+, but also in the UQ:UQH_2 ratio.
18. If nitrate is used as the terminal electron acceptor in the electron transport system of a bacterium, how many ATPs could be synthesized? [*Hint:* Refer to Table 9.1.]
19. Describe the movement of the electrons of NADH through complex I to CoQ.

SECTION 10.2

Comprehension Questions

20. Define the following terms:

a. protonmotive force
b. chemiosmotic coupling theory
c. uncoupler
d. ionophore

e. stator

21. Define the following terms:

    a. ATP synthase
    b. submitochondrial particle
    c. α,β-hexamer
    d. torque
    e. rotor

22. Define the following terms:

    a. respiratory control
    b. glycerol-3-phosphate shuttle
    c. uncoupling protein
    d. oligomycin
    e. malate–aspartate

23. Define the following terms:

    a. ANT
    b. symporter
    c. phosphate translocase
    d. CK
    e. CK/PCr shuttle

Fill in the Blanks

24. Heat generation by brown adipose tissue cells is referred to as _____.
25. The control of aerobic respiration by ADP is referred to as_____.
26. Hydrophobic polypeptides that dissipate osmotic gradients by inserting themselves into a membrane to form a channel are called _____.
27. _____ are small molecules that collapse a proton gradient by equalizing the proton gradient on both sides of a membrane.
28. An electrochemical proton gradient across a membrane is referred to as the _____ force.

Short-Answer Questions

29. What is shivering thermogenesis?
30. The antifungal drug nystatin kills fungal cells in part through formation of membrane pores that cause K^+ leakage. To what class of membrane-inserting molecules does nystatin belong?
31. Describe the processes that are driven by mitochondrial electron transport.
32. How does dinitrophenol inhibit ATP synthesis?
33. Four protons are required to drive the phosphorylation of ADP. Account for the function of each proton in this process.
34. Compare the amount of energy captured from a mole of glucose from glycolysis alone with the energy captured from the conversion of glucose to CO_2 and H_2O.
35. Describe the mechanism whereby uncoupling agents disrupt phosphorylation.
36. Valinomycin is an ionophore antibiotic that renders biological membranes permeable to K^+. Its side effects in patients with bacterial infections include a rise in body temperature and sweating. Explain.
37. What is the minimum voltage drop for individual electron transfer events in the mitochondrial electron transport system that is necessary for ATP synthesis?

38. Describe the role of UCP1 in nonshivering thermogenesis.

Critical-Thinking Questions

39. Ethanol is oxidized in the liver to form acetate, which is then converted to acetyl-CoA. Determine how many molecules of ATP are produced from 1 mol of ethanol. (Note that 2 mol of NADH are produced when ethanol is oxidized to form acetate.)
40. Glutamine is degraded to form NH_4^+, CO_2, and H_2O. How many molecules of ATP can be generated from 1 mol of this amino acid? (Removal of the amino group yields one molecule of NADH.)
41. Consumption of dinitrophenol by animals results in an immediate increase in body temperature. Explain the phenomenon. Why is dinitrophenol not used as a diet aid?
42. During an experiment, $^{14}CH_3$-COOH is fed to microorganisms. Trace the ^{14}C label through the citric acid cycle.
43. How many ATP molecules can be generated from one mol of $^{14}CH_3$-COOH? (The conversion of acetate to acetyl-CoA requires the consumption of 2 ATP.)
44. Explain the role of creatine kinase in cell metabolism.

SECTION 10.3

Comprehension Questions

45. Define the following terms:
    a. reactive oxygen species
    b. reactive nitrogen species
    c. superoxide dismutase
    d. antioxidant
    e. oxidative stress
46. Define the following terms:
    a. peroxiredoxin
    b. thioredoxin
    c. ascorbate
    d. respiratory burst
    e. glutathione
47. Define the following terms:
    a. β-carotene
    b. ischemia
    c. reperfusion injury
    d. MPTP
    e. α-tocopherol
48. Define the following terms:
    a. GSH
    b. SOD
    c. redox code
    d. RNS
    e. radical

Fill in the Blanks

49. The major enzymatic defense against oxidative stress is provided by superoxide dismutase, glutathione, peroxidase, peroxiredoxin, and _____.
50. A chemical species with an unpaired electron is called a_____.
51. Macrophages utilize an oxygen-consuming process called the _____ in which reactive oxygen species are generated to kill microorganisms or damaged cells.
52. Under conditions of _____ the antioxidant mechanisms of cells are overwhelmed.

Short-Answer Questions

53. When consumed in small amounts, vitamin E can protect against ROS. However, when it is taken in excessive amounts, vitamin E can make the body more susceptible to ROS. Explain.
54. Provide the reaction equations that illustrate the synthesis of ROS from electrons leaking from electron transport complexes.
55. Describe the types of cellular damage caused by ROS.
56. Describe the enzymatic activities used by cells to protect themselves from oxidative damage.
57. List some of the causes of reperfusion-triggered cardiac cell damage.
58. Explain the principles of the redox code. Distinguish between redox switches and redox sensors.

Critical-Thinking Questions

59. Dehydroascorbate is unstable at pH values greater than 6 and decomposes to form tartrate and oxalate. Cells use GSH to reduce the loss of ascorbate. What is the reaction pathway for the regeneration of ascorbate from dehydroascorbate?
60. In a major class of peroxiredoxins, referred to as 2-cys PRXs, two cysteines in the active site are involved in the catalytic mechanism. During the reaction, one of the cysteines is oxidized to a sulfenic acid by the substrate peroxide to yield an alcohol product. The sulfenic acid side chain then reacts with the sulfhydryl group of the other cysteine residue to yield a disulfide bond and a molecule of H_2O. Provide the pathway of this reaction. Include the regeneration of the two reduced cysteine residues in your answer.
61. The glutaridoxins are a class of small antioxidant enzymes that use GSH as a reducing agent. They have functions similar to those of the thioredoxins. Describe the pathway whereby an organic peroxide is reduced by a glutaridoxin. Include the redox cycle of GSH in your answer.

MCAT Study Questions

62. Which of the following molecules remains oxidized when the electron transport inhibitor antimycin A is added to a suspension of actively respiring mitochondria?

    a. cytochrome c
    b. NAD
    c. cytochrome a
    d. UQ

63. Which of the following ATP synthase components prevents the α,β hexamer from rotating?

    a. γ
    b. c
    c. b
    d. ε

64. The control of aerobic respiration by _____ is called respiratory control.

    a. ATP
    b. ADP
    c. AMP
    d. GDP

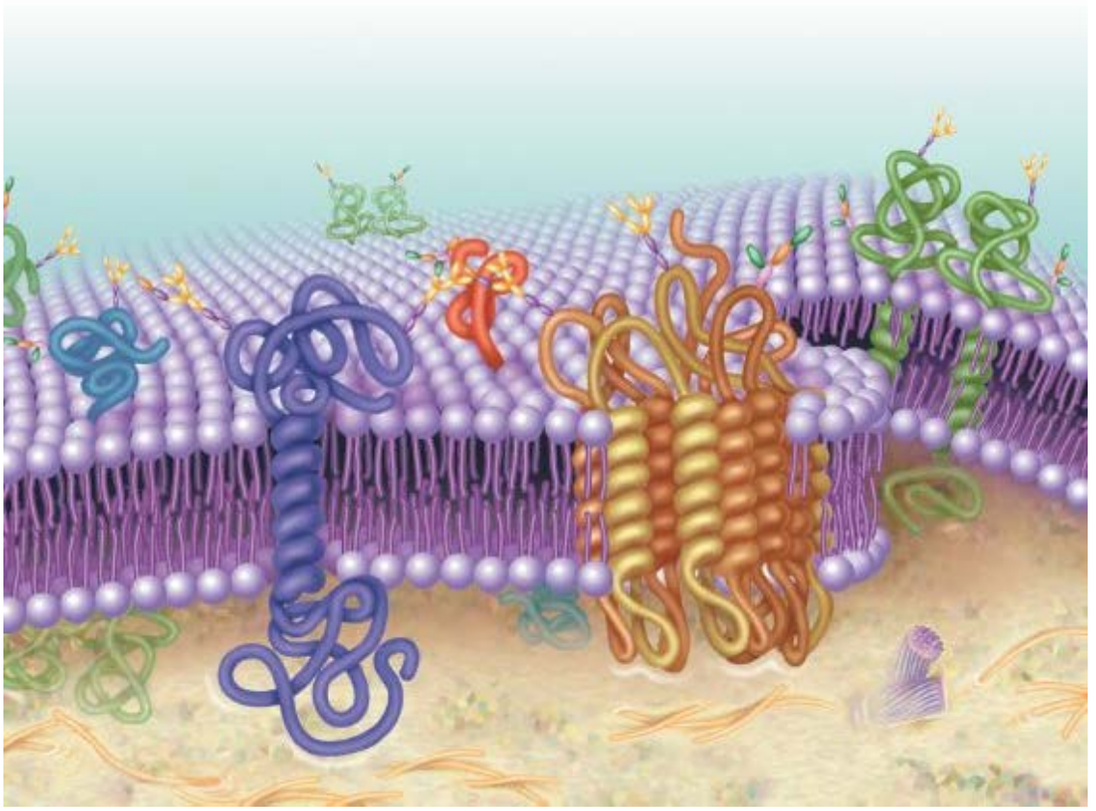
65. Which of the following dehydrogenases delivers electrons to directly to UQ in the electron transport pathway from the intermembrane space side of the inner mitochondrial membrane?

    a. succinate dehydrogenase
    b. acyl CoA dehydrogenase
    c. creatine kinase dehydrogenase
    d. glycerol-3-phosphate dehydrogenase

66. The aspartate–malate shuttle transfers electrons from cytoplasmic NADH to the electron transport system. Which of the following molecules is not directly involved in this pathway?

    a. malate
    b. oxaloacetate
    c. citrate
    d. α-ketoglutarate

CHAPTER 11 Lipids and Membranes



Biological Membrane A biological membrane is a dynamic compartmental barrier composed of a lipid bilayer noncovalently complexed with proteins, glycoproteins, glycolipids, and cholesterol.

OUTLINE

THE LOW-FAT DIET

11.1 LIPID CLASSES

Fatty Acids
The Eicosanoids
Triacylglycerols
Wax Esters
Phospholipids
Phospholipases
Sphingolipids
Sphingolipid Storage Diseases
Isoprenoids

Lipoproteins

11.2 MEMBRANES

Membrane Structure

Membrane Function

Biochemistry in Perspective

Alzheimer's Disease and Apolipoprotein E4

AVAILABLE ONLINE

Biochemistry in Perspective

Botulism and Membrane Fusion

The Low-Fat Diet

Obesity is a worldwide phenomenon. The World Health Organization estimates that there are currently more than 1.9 billion clinically obese adults, with 700 million more described as overweight. These are remarkable statistics considering that obesity was relatively rare before the twentieth century. The obesity rate in the United States was stable at about 12–14% until the late 1970s, when the rate began to rise. By the late 1980s, approximately 25% of American adults were obese or significantly overweight. This number is now over 75%. The current epidemic, and the chronic diseases associated with it (e.g., diabetes mellitus, cardiovascular disease, and certain forms of cancer), has been linked to the ready availability of cheap, energy-dense, and nutrient-deficient processed food and a shift toward physically less demanding lifestyles. Throughout this period there has been an overwhelming amount of dietary advice, most notably the promotion, beginning in the late 1970s, of low-fat diets. Unfortunately, because of the unpalatable nature of low-fat food, carbohydrates (usually sucrose or high-fructose corn syrup) soon replaced the fat content. Apparently, many consumers believed that low-fat foods could be eaten in excess without considering that the caloric content of these foods equals or exceeds that of the fat-containing foods they replaced.

Although low-fat diets are still popular, they have proven to be ineffective for achieving sustained weight loss. Health problems associated with low-fat diets, however, do not end with weight gain in obesity-prone individuals. The original nutritional recommendation that diets should be low in saturated fats was often interpreted as low in all fats. Consequently, without realizing the diverse functions of dietary fats, many weight-conscious individuals reduced or eliminated almost all fats from their diets. Among the consequences of an extreme low-fat diet (15% or less) are deficiencies in the fat-soluble vitamins (A, D, E, and K) and the essential fatty acids linoleic and linolenic acids. The fat-soluble vitamins are important for numerous physiological processes (e.g., growth, immunity, cell repair, and blood clotting). The essential fatty acids (EFAs), so called because the body's cells cannot synthesize them, perform a wide variety of functions. In addition to being important constituents of cell membranes, the EFAs are also converted into biologically active derivatives with roles in immunity, inflammatory responses, and nervous system function. EFA deficiencies in adults have been associated with dry scaly skin, brittle hair, fatigue, high blood pressure, atherosclerosis, depressed immunity, poor wound healing, depression, and cravings for fatty foods. In addition to inadequate growth, EFA deficiencies in children have been linked to impaired brain development, with symptoms that include hyperactivity, attention deficit, and aggressive behavior.

Ironically, recent research indicates that modest amounts of healthy dietary fat not only provide the body with required nutrients but also promote satiety (the perception of fullness that inhibits further eating behavior). Among the most intriguing satiety-inducing mechanisms is the dietary fat-induced release of endorphins by the central nervous system. Endorphins are endogenous opiate-like proteins that are partly responsible for the pleasure that results from the consumption of delicious food.

The substitution of sugar for fat in processed foods, the other unforeseen result of the low-fat recommendation, has had even more serious consequences for human health. The metabolic mechanisms whereby dietary sugar contributes to obesity, hypertension, and atherosclerosis are outlined in Chapter 16.

Overview

LIPIDS ARE NATURALLY OCCURRING WATER-INSOLUBLE SUBSTANCES THAT PERFORM A STUNNING ARRAY OF FUNCTIONS IN LIVING ORGANISMS. Some lipids are vital energy reserves. Others are the primary structural components of biological membranes. Still other lipid molecules act as hormones, antioxidants, pigments, or vital growth factors and vitamins. This chapter describes the structures and properties of the major lipid classes found in living organisms as well as the structural and functional properties of biomembranes.

Lipids are a diverse group of biomolecules. Because of this diversity, the term **lipid** has an operational rather than a structural definition. Lipids are defined as substances from living organisms that dissolve in nonpolar solvents such as ether, chloroform, and acetone but not appreciably in water. The functions of lipids are also diverse. Several types of lipid molecules are important structural components in cell membranes. Another type, the fats and oils, store energy efficiently. Other lipid molecules are chemical signals, vitamins, or pigments. Finally, some lipid molecules that occur in the outer surfaces of various organisms have protective or waterproofing functions.

Chapter 11 describes the structure and function of each major type of lipid and discusses the lipoproteins, complexes of protein and lipid that transport lipids in animals. The chapter ends with an overview of membrane structure and function. Chapter 12 describes the metabolism of several major lipids.

11.1 LIPID CLASSES

Lipids may be classified in many different ways. For this discussion, lipids can be subdivided into the following classes:

1. Fatty acids
2. Triacylglycerols
3. Wax esters
4. Phospholipids (phosphoglycerides and sphingomyelin)
5. Sphingolipids (molecules other than sphingomyelin that contain the amino alcohol sphingosine)
6. Isoprenoids (molecules made up of repeating isoprene units, a branched five-carbon hydrocarbon)

Fatty Acids

Fatty acids are monocarboxylic acids that typically contain hydrocarbon chains of variable lengths (between 12 and 20 or more carbons) (**Figure 11.1**). Fatty acids are numbered from the carboxylate end. Greek letters are used to designate certain carbon atoms. The α-carbon in a fatty acid is adjacent to the carboxylate group, the β-carbon is two atoms removed from the carboxylate group, and so forth. The terminal methyl carbon atom is designated the omega (ω) carbon. **Table 11.1** gives the structures, names, and standard abbreviations of several common fatty acids. Fatty acids occur primarily in triacylglycerols and several types of membrane-bound lipid molecules.

Most naturally occurring fatty acids have an even number of carbon atoms that form an unbranched chain. (Unusual fatty acids with branched or ring-containing chains are found in some species.) Fatty acid chains that contain only carbon–carbon single bonds are referred to as *saturated*; those molecules that contain one or more double bonds are said to be *unsaturated*. Double bonds are rigid structures, so molecules that contain them can occur in two isomeric forms: *cis* and *trans*. In *cis* isomers, similar or identical groups are on the same side of a double bond (**Figure 11.2a**). When such groups are on opposite sides of a double bond, the molecule is said to be a *trans*-isomer (**Figure 11.2b**). The double bonds in most naturally occurring fatty acids are in a cis configuration. The presence of a cis double bond causes an inflexible "kink" in a fatty acid chain (**Figure 11.3**). Because of this structural feature, unsaturated fatty acids do not pack as closely together as saturated fatty acids. Less energy is required to disrupt the intermolecular forces between unsaturated fatty acids. Therefore, they have lower melting points and are liquids at room temperature. For example, a sample of palmitic acid (16:0), a saturated fatty acid, melts at 63°C, whereas palmitoleic acid ($16:1^{\Delta 9}$) melts at 0°C.

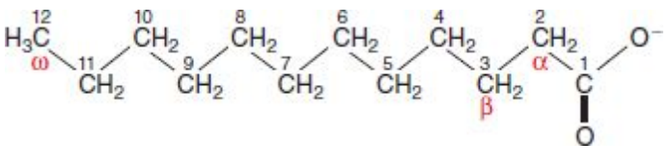


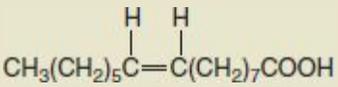
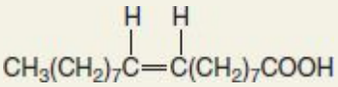
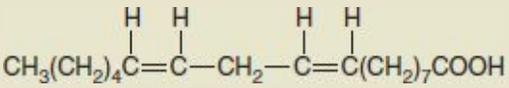
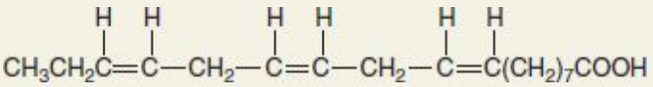
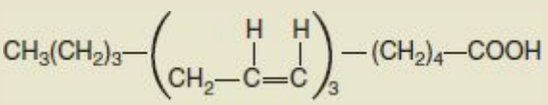
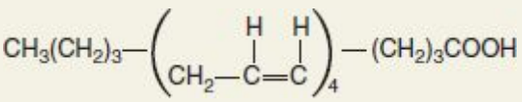
FIGURE 11.1

Fatty Acid Structure

Fatty acids consist of a long-chain hydrocarbon covalently bonded to a carboxylate group. The lipid shown is dodecanoic acid (common name, lauric acid), a 12-carbon saturated fatty acid (12:0).

TABLE 11.1 Examples of Fatty Acids

Common Name	Structure	Abbreviation*
Saturated Fatty Acids		
Myristic acid	$CH_3(CH_2)_{12}COOH$	14:0
Palmitic acid	$CH_3(CH_2)_{12}CH_2CH_2COOH$	16:0
Stearic acid	$CH_3(CH_2)_{12}CH_2CH_2CH_2CH_2COOH$	18:0
Arachidic acid	$CH_3(CH_2)_{12}CH_2CH_2CH_2CH_2CH_2CH_2COOH$	20:0
Lignoceric acid	$CH_3(CH_2)_{12}CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2COOH$	24:0

Cerotic acid	$CH_3(CH_2)_{12}CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2COOH$	26:0
Unsaturated Fatty Acids		
Palmitoleic acid		16:1Δ9
Oleic acid		18:1Δ9
Linoleic acid		18:2Δ9,12
α-Linolenic acid		18:3Δ9,12,15
γ-Linolenic acid		18:3Δ,9,12
Arachidonic acid		20:4Δ5,8,11,14

* Note that in the abbreviations for specific fatty acids, the number to the left of the colon is the total number of carbon atoms, and the number to the right is the number of double bonds. A superscript denotes the placement of a double bond. For example, Δ9 signifies that there are eight carbons between the carboxyl group and the double bond; that is, the double bond occurs between carbons 9 and 10.

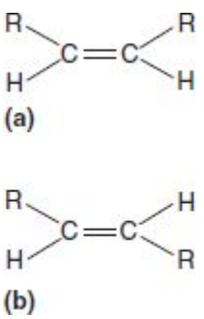


FIGURE 11.2

Isomeric Forms of Unsaturated Molecules

In *cis*-isomers, (a) both R groups are on the same side of the carbon–carbon double bond. In *trans*-isomers (b), the R groups are on different sides.

Unsaturated fatty acids are also classified according to the location of the first double bond relative to the terminal methyl (omega, ω) end of the molecule. For example, linoleic acid and α-linolenic acid can be designated as 18:2ω-6 (equivalent to $18:2^{\Delta 9,12}$) and 18:3ω-3 (equivalent to $18:3^{\Delta 9,12,15}$), respectively. (The number to the right of ω designates the carbon at which the first double bond occurs, counting from the methyl end of the fatty acid. Sequential double bonds are always three carbons apart.) It is noteworthy that fatty acids with trans double bonds have three-dimensional structures similar to those of saturated fatty acids. Trans fatty acids are side products of margarine production in which plant oils are hydrogenated. In addition, the presence of one or more double bonds in a fatty acid makes it susceptible to oxidation (**Figure 10.23**). The

consequences include the effects of oxidative stress on cell membranes and the tendency of oils to become rancid (i.e., they contain unpleasant smelling or tasting short-chain organic acids).

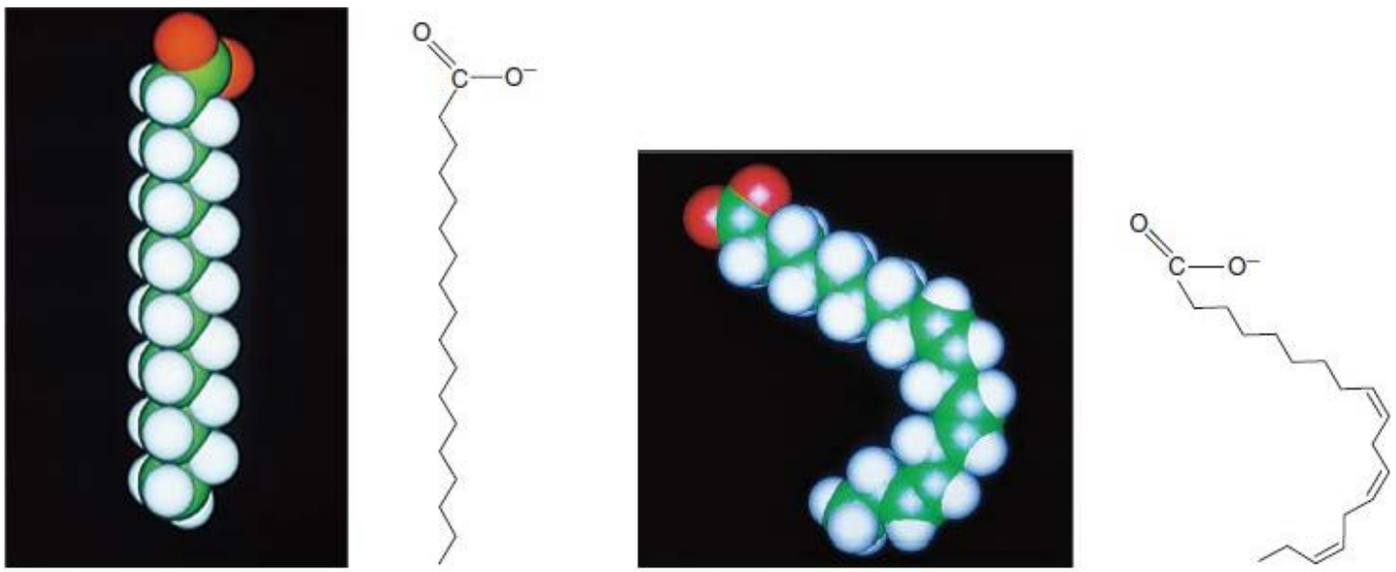


FIGURE 11.3

Space-Filling and Conformational Models

(a) A saturated fatty acid (stearic acid) and (b) an unsaturated fatty acid (α-linolenic acid). (Green spheres = carbon atoms; white spheres = hydrogen atoms; red spheres = oxygen atoms.)

Fatty acids with one double bond are referred to as **monounsaturated** molecules. When two or more double bonds occur in fatty acids, usually separated by methylene groups (—CH_2—), they are referred to as **polyunsaturated**. The monounsaturated fatty acid oleic acid ($18:1^{\Delta 9}$) and the polyunsaturated linoleic acid ($18:2^{\Delta 9,12}$) are among the most abundant fatty acids in living organisms.

Organisms such as plants and bacteria can synthesize all the fatty acids they require from acetyl-CoA (Chapter 12). Mammals obtain most of their fatty acids from dietary sources, but they can synthesize saturated fatty acids and some monounsaturated fatty acids. Mammals can also modify some dietary fatty acids by adding two-carbon units and introducing some double bonds. Fatty acids that can be synthesized are called **nonessential fatty acids**. Because mammals do not possess the enzymes required to synthesize linoleic ($18:2^{\Delta 9,12}$) and α-linolenic ($18:2^{\Delta 9,12,15}$) acids, these **essential fatty acids** must be obtained from the diet.

Linoleic acid ($18:2^{\Delta 9,12}$ or 18:2ω-6) is the precursor for numerous derivatives, formed by elongation and/or desaturation reactions. Prominent examples include γ-linolenic acid ($18:3^{\Delta 6,9,12}$ or 18:3ω-6), arachidonic acid ($20:4^{\Delta 5,8,11,14}$ or 20:4ω-6), and docosapentaenoic acid ($22:5^{\Delta 4,7,10,13,16}$ or 22:5ω-6) (DPA). Together, linoleic acid and its derivatives are referred to as the **omega-6 fatty acids**. Food sources include various vegetable oils (e.g., sunflower and soybean oils), eggs, and poultry. α-Linolenic acid ($18:3^{\Delta 9,12,13}$ or 18:3ω-3) and its derivatives, such as eicosapentaenoic acid ($20:5^{\Delta 5,8,11,14,17}$ or 20:5ω-3) (EPA) and docosahexaenoic acid ($22:6^{\Delta 4,7,10,13,16,19}$ or 22:6ω-3) (DHA), are referred to as the **omega-3 fatty acids**. Sources of α-linolenic acid include flaxseed and soybean oils, and walnuts. EPA and DHA, also found in fish and fish oils (e.g., salmon, tuna, and sardines), are now believed to promote cardiovascular health.

Effects associated with diets with adequate amounts of these two fatty acids include lower blood levels of triacylglycerols (triglycerides), lower blood pressure, and decreased platelet aggregation. Essential fatty acids are used as structural components (e.g., phospholipids in membranes) and as precursors for several important metabolites. Prominent examples of the latter include the eicosanoids and anandamine.

Omega Fatty Acids

The eicosanoids are hormone-like molecules derived from omega-6 or omega-3 fatty acids. In general, omega-6-derived eicosanoids promote inflammation, whereas omega-3 derivatives are less inflammatory. The ratio of omega-6 to omega-3 fatty acids in the diet influences the relative amounts of inflammatory and anti-inflammatory eicosanoids that are synthesized. It is currently believed that 1:1 to 1:4 ratios are healthy. In many developed countries, typical diets provide 10:1 to 30:1 ratios. These ratios favor a net increase in unfavorable inflammatory reactions in the body, an undesirable condition that increases the risk of chronic disease.

Anandamine (*N*-arachidonyl ethanolamine), a derivative of arachidonic acid, is a Δ9-*endocannabinoid*, a substance produced in the body that binds to the same receptor as tetrahydrocannabinol, a psychoactive drug. Anandamine acts as a neurotransmitter in the central and peripheral nervous systems, where it affects feeding and sleep behavior, short-term memory, and pain relief.

REACTIONS OF FATTY ACIDS Fatty acids have several important chemical properties. The reactions that they undergo are typical of short-chain carboxylic acids. For example, fatty acids react with alcohols to form esters:

(1)

This reaction is reversible; that is, under appropriate conditions a fatty acid ester can react with water to produce a fatty acid and an alcohol. Unsaturated fatty acids with double bonds can undergo hydrogenation reactions to form saturated fatty acids. Finally, unsaturated fatty acids are susceptible to oxidation (as described in **Figure 10.23**).

Certain fatty acids are covalently attached to a wide variety of eukaryotic proteins; such proteins are referred to as *acylated* proteins. Fatty acid groups (called **acyl groups**) clearly facilitate the interactions between membrane proteins and their hydrophobic environment. Myristoylation and palmitoylation, the most common forms of protein acylation, are now known to influence a variety of structural and functional properties of proteins. Promotion of protein binding to membranes is a prominent example. In addition, hydrophobic fatty acid molecules are transported from fat cells to body cells for energy-yielding redox reactions by means of the acylation of water-soluble serum proteins.

KEY CONCEPTS

Fatty acids are monocarboxylic acids, most of which are found in triacylglycerol molecules, several types of

membrane-bound lipid molecules, or acylated membrane proteins.

The Eicosanoids

The **eicosanoids**, produced in most mammalian tissues, include the **prostaglandins**, **thromboxanes**, and **leukotrienes**. Together the eicosanoids mediate a wide variety of physiological processes, including smooth muscle contraction, inflammation, pain perception, and blood-flow regulation. Eicosanoids are also implicated in several diseases such as myocardial infarction and rheumatoid arthritis. Because they are generally active within the cell in which they are produced, the eicosanoids are called **autocrine** regulators.

Myocardial Infarction and Arthritis

Eicosanoids (**Figure 11.4**) are usually designated by their abbreviations. They are named according to the following system. The first two letters indicate the type of eicosanoid (PG = prostaglandin, TX = thromboxane, LT = leukotriene). The third letter identifies the type of modification made to the parent compound of the eicosanoid (e.g., A = hydroxyl group and an ether ring, B = two hydroxyl groups). The number in an eicosanoid name indicates the number of double bonds in the molecule. Eicosanoids are extremely difficult to study because they are active for short periods (often measured in seconds or minutes). In addition, they are produced only in small amounts.

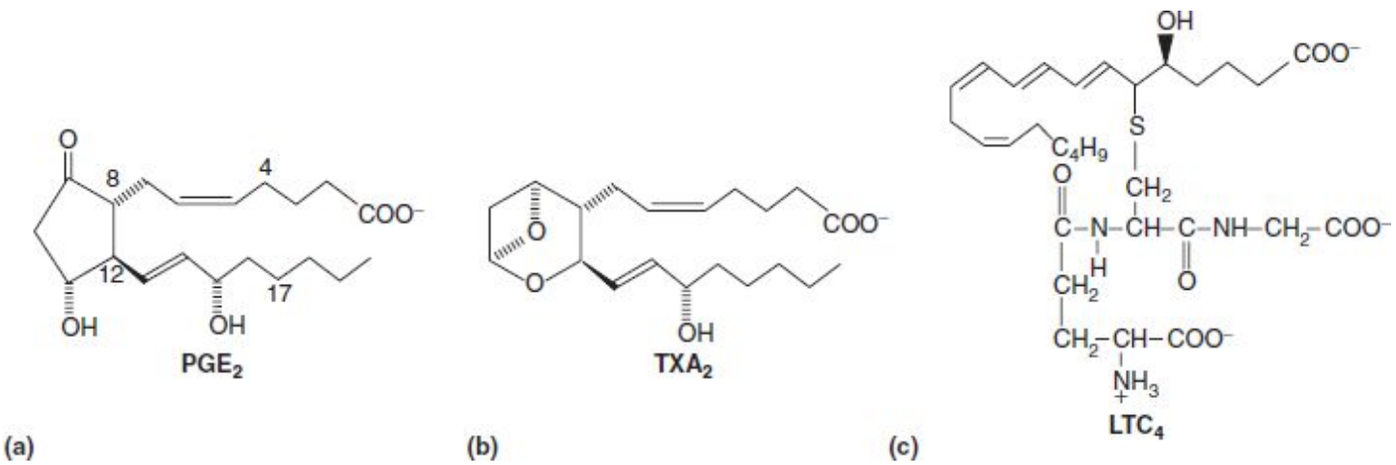


FIGURE 11.4

Eicosanoids

(a) Prostaglandin E_2. (b) Thromboxane A_2. (c) Leukotriene C_4. Note that LTC_4 has a glutathione substituent.

Eicosanoids are derived from either arachidonic acid or EPA. Production of eicosanoids begins after arachidonic acid or EPA is released from membrane phospholipid molecules by the enzyme phospholipase A_2. Individual classes of eicosanoids are described in the online essay The Eicosanoids.

QUESTION 11.1

Rheumatoid arthritis is an autoimmune disease in which the joints are chronically inflamed. In **autoimmune diseases,** the immune system fails to distinguish between self and nonself. For reasons that are not understood, specific lymphocytes are stimulated to produce antibodies, referred to as *autoantibodies*. These molecules bind to surface antigens on the patient's own cells as if they were foreign. In rheumatoid arthritis, the binding of an autoantibody called rheumatoid factor (RF) to the F_c portion of IgG (p. 163) promotes inflammation because it stimulates the infiltration of joint tissue by several types of white blood cells. The leakage of lysosomal enzymes from actively phagocytosing cells (neutrophils and macrophages) leads to further tissue damage. The inflammatory response is perpetuated by the release of several eicosanoids. For example, macrophages are known to produce eicosanoids such as PGE_4, TXA_2, and several leukotrienes.

Autoimmune Diseases

Currently, the treatment of rheumatoid arthritis consists of suppressing pain and inflammation. Despite treatment, however, the disease continues to progress. Aspirin inhibits cyclooxygenase, the enzyme that catalyzes the first reaction in prostaglandin synthesis: the conversion of arachidonic acid to PGE_2. It plays an important role in the treatment of rheumatoid arthritis and other types of inflammation because of its low cost and relative safety. Certain steroids, which inhibit phospholipase A_2, are more potent than aspirin in reducing inflammation; that is, they immediately and dramatically reduce painful symptoms. However, steroids have serious side effects. For example, prednisone may depress the immune system, cause fat redistribution to the neck ("buffalo hump"), and cause serious behavioral changes. For these and other reasons, prednisone is used to treat rheumatoid arthritis only when a patient does not respond to aspirin or similar drugs.

Compare the effects of aspirin and steroids on eicosanoid metabolism and suggest a reason why this information is relevant to the treatment of rheumatoid arthritis. Does it explain the difference between the effectiveness of aspirin and steroids in treating inflammation?

Triacylglycerols

Triacylglycerols are esters of glycerol with three fatty acid molecules (**Figure 11.5**). Glycerides with one or two fatty acid groups, called monoacylglycerols and diacylglycerols, respectively, are metabolic intermediates. They are normally present in small amounts. Because triacylglycerols have no charge (i.e., the carboxyl group of each fatty acid is joined to glycerol through a covalent bond), they are sometimes referred to as **neutral fats**. Most triacylglycerol molecules contain fatty acids of varying lengths; the acids themselves may be unsaturated, saturated, or a combination (**Figure 11.6**). Depending on their fatty acid compositions, triacylglycerol mixtures are referred to as fats or oils. *Fats*, which are solid at room temperature, contain a large proportion of saturated fatty acids. *Oils* are liquid at room temperature because of their relatively high unsaturated fatty acid content. (Recall that unsaturated fatty acids do not pack together as closely as do saturated fatty acids.)

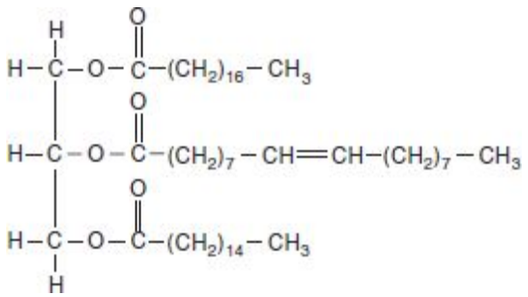


FIGURE 11.5

Triacylglycerol

Each triacylglycerol molecule is composed of glycerol esterified to three (usually different) fatty acids.

3D animation of triacylglycerol

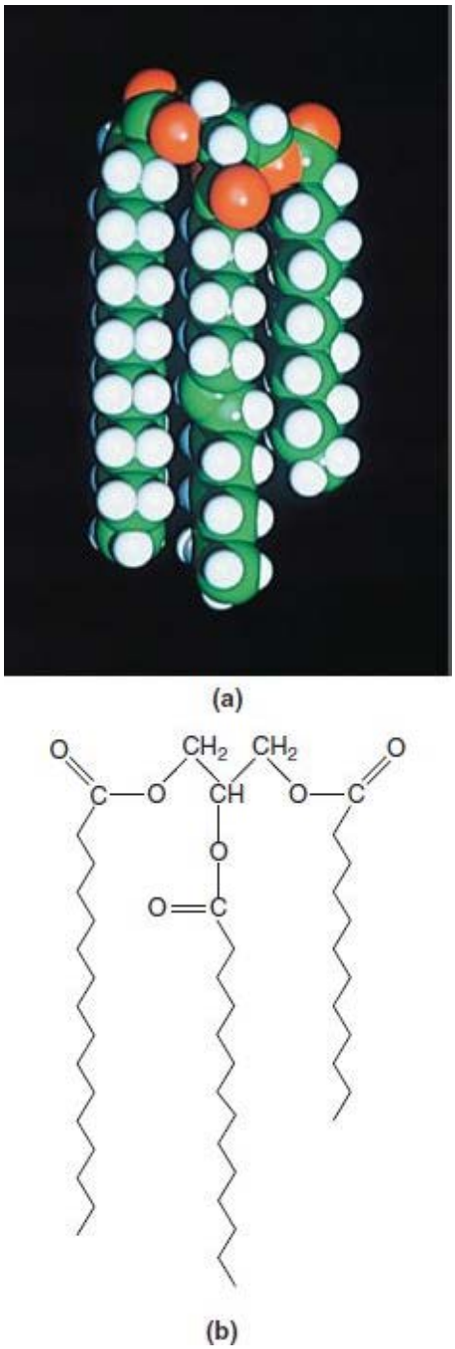


FIGURE 11.6

Space-Filling (a) and Conformational (b) Models of a Triacylglycerol

Triacylglycerols are molecules that serve as a rich source of chemical bond energy.

In animals, triacylglycerols (usually referred to as *fat*) have several roles. First, they are the major storage and transport form of fatty acids. Triacylglycerol molecules store energy more efficiently than glycogen for several reasons:

1. Triacylglycerols are hydrophobic, and therefore they coalesce into compact, anhydrous droplets within cells. A specialized type of cell called the *adipocyte*, found in adipose tissue, stores triacylglycerols. The anhydrous triacylglycerols store an equivalent amount of energy in about one-eighth of the volume of glycogen (the other major energy storage molecule), which binds a substantial amount of water.
2. Triacylglycerols are more reduced and can thus release more electrons when oxidized than

an equivalent amount of carbohydrate. Therefore, triacylglycerols release more energy (38.9 kJ/g of fat compared with 17.2 kJ/g of carbohydrate) when they are degraded.

A second important function of fat is to provide insulation in low temperatures. Fat, a poor conductor of heat prevents heat loss. Adipose tissue, with its high triacylglycerol content, is found throughout the body (especially underneath the skin). Finally, in some animals fat molecules secreted by specialized glands make fur or feathers water-repellent.

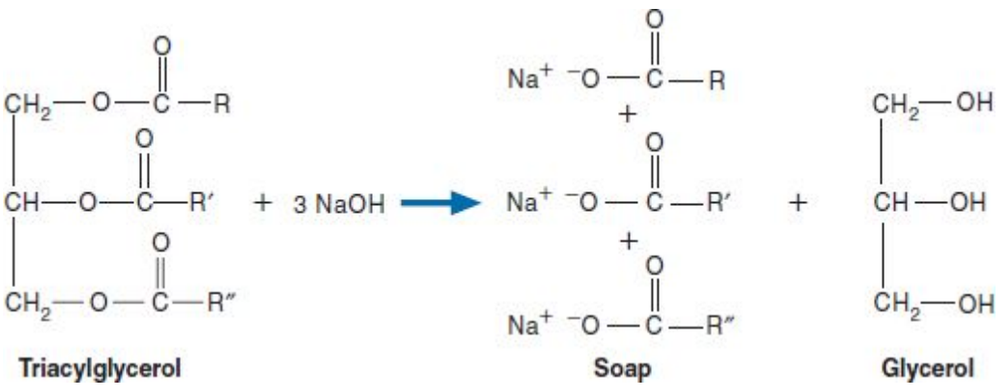
In plants, triacylglycerols constitute an important energy reserve in fruits and seeds. Because these molecules contain relatively large amounts of unsaturated fatty acids (e.g., oleic and linoleic), they are referred to as plant oils. Seeds rich in oil include peanut, corn, palm, safflower, soybean, and flax. Avocados and olives are fruits with a high oil content.

QUESTION 11.2

Oils can be converted to fats in a commercial nickel-catalyzed process referred to as *partial hydrogenation*. Under relatively mild conditions (180°C and pressures of about 1013 torr or 1.33 atm), enough double bonds are hydrogenated for liquid oils to solidify. This solid material, oleomargarine, has a consistency like butter. However, oils are not completely hydrogenated during commercial hydrogenation processes. Propose a practical reason for this.

QUESTION 11.3

Soap making is an ancient process. The Phoenicians, a seafaring people who dominated trade in the Mediterranean area about 3000 years ago, are believed to have been the first to manufacture soap. Traditionally, soap has been made by heating animal fat with potash. [Potash is a mixture of potassium hydroxide (KOH) and potassium carbonate (K_2CO_3) obtained by mixing wood ash with water.] Currently, soap is made by heating beef tallow or coconut oil with sodium or potassium hydroxide. During this reaction, which is a *saponification*, triacylglycerol molecules are hydrolyzed to give glycerol and the sodium or potassium salts of fatty acids:



Fatty acid salts (soaps) are amphipathic molecules (i.e., they possess polar and nonpolar domains) that spontaneously form into micelles (Figure 3.14). Soap micelles have negatively charged surfaces that repel each other. Soap is used to remove dirt mixed with grease because it is an *emulsifying agent*; that is, it promotes the dispersion of one substance in another. The mixing of soap and grease results in an emulsion—specifically, a system in which the oil molecules are dispersed in the soap micelles. Complete the diagram in **Figure 11.7** and explain how this process occurs.

KEY CONCEPTS

- Triacylglycerols are molecules consisting of glycerol esterified to three fatty acids.
- They are a rich energy source in both animals and plants.

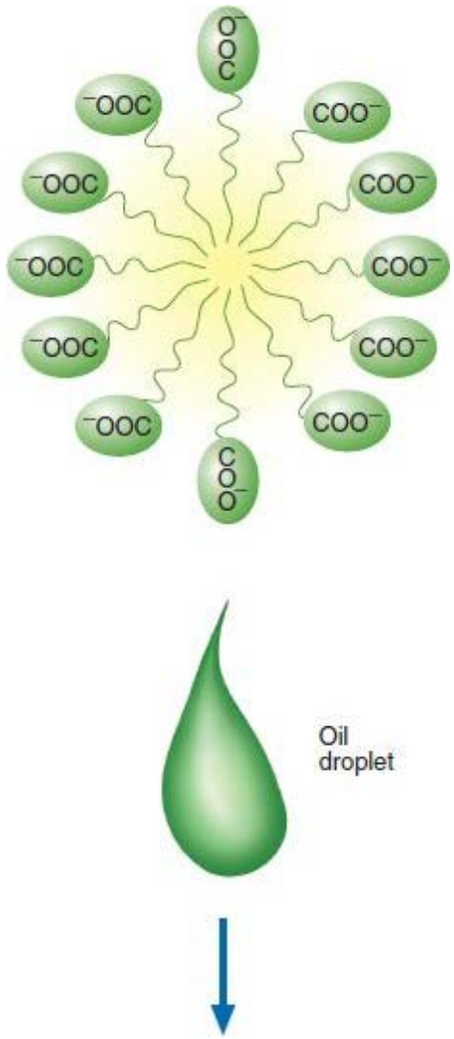


FIGURE 11.7

Soap: An Emulsifying Agent

How does soap interact with an oil droplet? [*Hint*: Recall that "like dissolves like."]

Wax Esters

Waxes are complex mixtures of nonpolar lipids. They are protective coatings on the leaves, stems, and fruits of plants and on the skin and fur of animals. **Wax esters** composed of long-chain fatty acids and long-chain alcohols are prominent constituents of most waxes. Well-known waxes include carnauba wax, produced by the leaves of the Brazilian wax palm, and beeswax. The predominant constituent of carnauba wax is the wax ester melissyl cerotate (**Figure 11.8**). Triacontyl hexadecanoate is one of several important wax esters in beeswax. Waxes also contain hydrocarbons, long-chain alcohols and aldehydes, fatty acids, and sterols (steroid alcohols).

Phospholipids

Phospholipids have several roles in living organisms. They are first and foremost structural components of membranes. In addition, several phospholipids are emulsifying agents and **surface active agents**. Phospholipids are suited to these roles because, like fatty acid salts, they are amphipathic molecules. The hydrophobic domain of a phospholipid is composed largely of the hydrocarbon chains of fatty acids; the hydrophilic domain, called a **polar head group**, contains phosphate and other charged or polar groups.

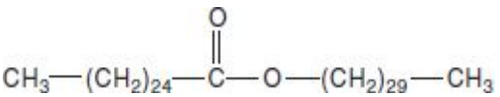


FIGURE 11.8

The Wax Ester Melissyl Cerotate

Found in carnauba wax, melissyl cerotate is an ester formed from melissyl alcohol and cerotic acid.

When phospholipids are suspended in water, they spontaneously rearrange into ordered structures (**Figure 11.9**). As these structures form, phospholipid hydrophobic groups are buried in the interior to exclude water. Simultaneously, hydrophilic polar head groups are oriented so that they are exposed to water. When phospholipid molecules are present in sufficient concentration, they form bimolecular layers. This property of phospholipids (and other amphipathic lipid molecules) is the basis of membrane structure (see pp. 422–35).

There are two types of phospholipid: phosphoglycerides and sphingomyelins. **Phosphoglycerides** are molecules that contain glycerol, fatty acids, phosphate, and an alcohol (e.g., choline). **Sphingomyelins** differ from phosphoglycerides in that they contain sphingosine instead of glycerol. Because sphingomyelins are also classified as sphingolipids, their structures and properties are discussed separately.

Phosphoglycerides are the most numerous phospholipid molecules found in cell membranes. The simplest phosphoglyceride, phosphatidic acid, is the precursor for all other phosphoglyceride molecules. Phosphatidic acid is composed of glycerol-3-phosphate that is esterified with two fatty acids. Phosphoglyceride molecules are classified according to which alcohol becomes esterified to the phosphate group. For example, if the alcohol is choline, the molecule is called phosphatidylcholine (PC) (also referred to as *lecithin*). Other types of phosphoglyceride include phosphatidylethanolamine (PE), phosphatidylserine (PS), diphosphatidylglycerol (dPG) (also called cardiolipin), and phosphatidylinositol (PI). (Refer to **Table 11.2** for the structures of the common types of phosphoglyceride.) The most common fatty acids in the phosphoglycerides have between 16 and 20 carbons. Saturated fatty acids usually occur at C-1 of glycerol. The fatty acid substituent at C-2 is usually unsaturated.

Both phosphatidylethanolamine and cardiolipin have relatively small polar heads. Usually found in the inner leaflet of membranes, PE (25% of all phospholipids in humans) stabilizes membrane curvature and has a role in membrane fusion. Cardiolipin is found almost entirely in membranes whose function is to generate the electrochemical potential that drives ATP synthesis (i.e., in bacterial plasma membrane and mitochondrial inner membrane, where it constitutes 20% of total lipid). Composed of a lipid dimer with two phosphatidyl groups and a total of four hydrocarbon chains, cardiolipin stabilizes ETC supercomplexes (p. 373).

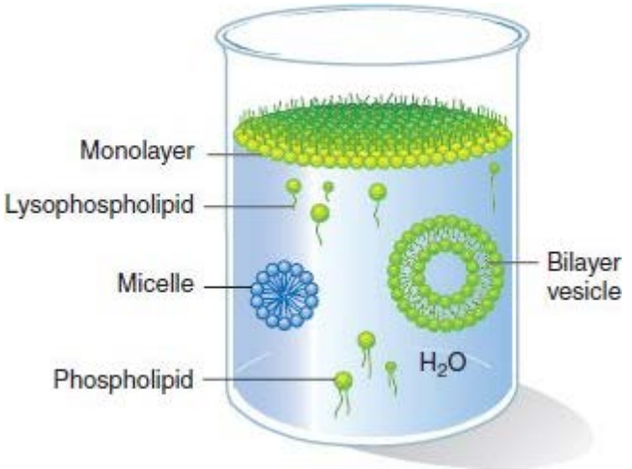
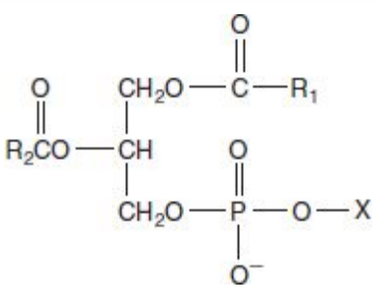


FIGURE 11.9

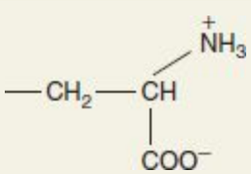
Phospholipid Molecules in Aqueous Solution

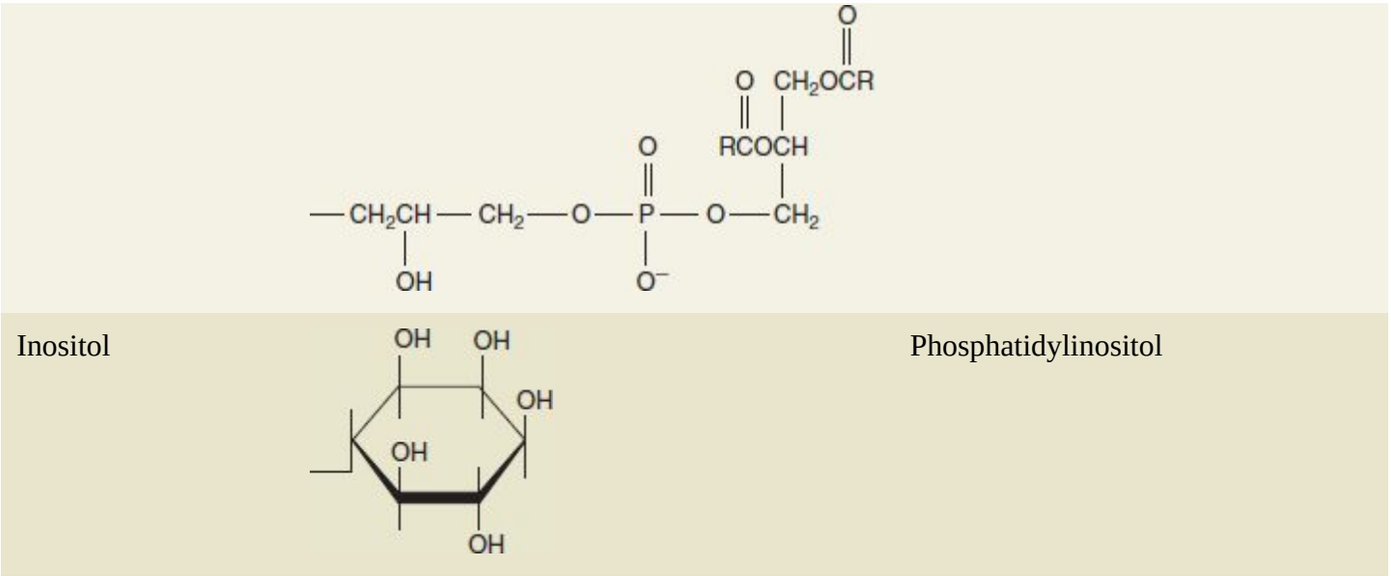
Each molecule is represented as a polar head group attached to one or two fatty acyl chains. (Lysophospholipid molecules possess only one fatty acyl chain.) The monolayer on the surface of the water forms first. As the phospholipid concentration increases, bilayer vesicles begin to form. Because of their wedge shape (compared to the cylindrical shape of phospholipids that contain two fatty acid chains), lysophospholipid molecules form micelles. The micelle and bilayer vesicle in this illustration are shown in cross-section.

TABLE 11.2 Major Classes of Phosphoglycerides



X Substituent

Name of X-OH	Formula of X	Name of Phospholipid
Water	—H	Phosphatidic acid
Choline	$-CH_2CH_2\overset{+}{N}(CH_3)_3$	Phosphatidylcholine (lecithin)
Ethanolamine	$-CH_2CH_2\overset{+}{N}H_3$	Phosphatidylethanolamine (cephalin)
Serine		Phosphatidylserine
Glycerol	$-CH_2CHCH_2OH$ (OH)	Phosphatidylglycerol
Phosphatidylglycerol		Diphosphatidylglycerol (cardiolipin)



A derivative of phosphatidylinositol, namely, phosphatidylinositol-4,5-bisphosphate (PIP_2), is found in only small amounts in plasma membranes. PIP_2 is now recognized as an important component of intracellular signal transduction. The *phosphatidylinositol cycle*, initiated when certain hormones bind to membrane receptors, is described in Section 16.2.

Phosphatidylinositol is also a prominent structural component of GP I anchors. **GPI anchors** (**Figure 11.10**), which are also composed of a trimannosylglucosamine group and phosphoethanolamine, attach certain proteins to the external surface of the plasma membrane. Proteins are attached to the anchor molecule via an amide linkage between the carboxyl terminal of the protein and the amino nitrogen of ethanolamine. The two fatty acids of the phosphatidylinositol component are embedded in the plasma membrane. GPI-linked proteins are believed to be components of lipid rafts (p. 427).

KEY CONCEPTS

- Phospholipids are amphipathic molecules that play important roles in living organisms as membrane components, emulsifying agents, and surface active agents.
- There are two types of phospholipid: phosphoglycerides and sphingomyelins.

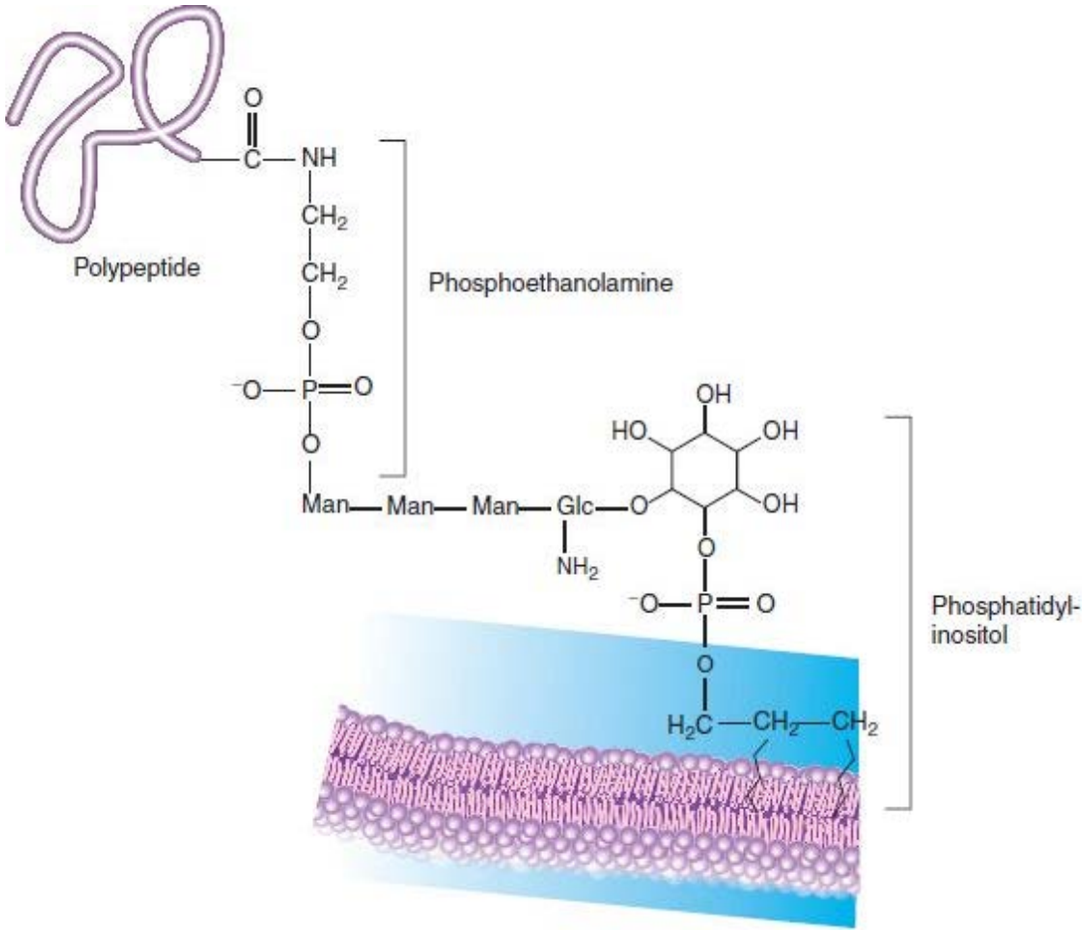


FIGURE 11.10

GPI Anchor

GPI-anchored proteins are attached to the external surface of the membrane through a linker element, phosphoethanolamine-Man_3-$GlcNH_2$, connecting the polypeptide at its carboxy terminus via an amide bond to a membrane phosphatidylinositol via an ether bond. Note that there are variations of this structure. For example, $GlcNH_2$ can be acetylated, and the phosphate of phosphatidic acid can be linked to either C-2 or C-3 of inositol.

QUESTION 11.4

Dipal mitoylphosphatidylcholine is the major component of *surfactant*, or surface active agent (an amphipathic molecule), that is secreted into lung alveoli to reduce the surface tension of the primarily aqueous extracellular fluid of the alveolar epithelia. Alveoli, also referred to as alveolar sacs, are the functional units of respiration. Oxygen and carbon dioxide diffuse across the walls of alveolar sacs, which are one cell thick. The water on alveolar surfaces has a high surface tension because of the attractive forces between the molecules. If the water's surface tension is not reduced, the alveolar sac tends to collapse, making breathing extremely difficult. If premature infants lack sufficient surfactant, they are likely to die of suffocation. This condition is called *respiratory distress syndrome*. Draw the structure of dipalmitoylphosphatidylcholine. Considering the general structural features of phospholipids, propose a reason why surfactant is effective in reducing surface tension.

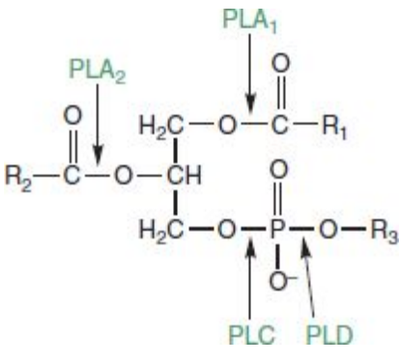


FIGURE 11.11

Phospholipases

Phospholipases hydrolyze ester bonds in phospholipids. PLB (not shown) has both PLA_1 and PLA_2 activities.

Phospholipases

Phospholipases hydrolyze ester bonds in glycerophospholipid molecules (**Figure 11.11**). They are classified according to the specific bond that they cleave. Phospholipases A_1 (PLA_1) and A_2 (PLA_2) hydrolyze the ester bonds at C_1 and C_2 of glycerol, respectively. The products of PLA_1 and PLA_2 are a fatty acid and a *lysophosphatide* (a glycerophospholipid from which one fatty acid has been removed). Phospholipase B (PLB) can hydrolyze both the C-1 and the C-2 ester bonds. Phospholipases C (PLC) and D (PLD) are phosphodiesterases that yield diacylglycerol and phosphatidic acid, respectively. Phospholipases have three major functions: membrane remodeling, signal transduction, and digestion. They are also used by some organisms as biological weapons.

MEMBRANE REMODELING Cells use phospholipases to alter the flexibility of membranes (p. 423) by adjusting the ratio of saturated and unsaturated fatty acids or to replace a damaged fatty acid. Fatty acid removal from a phospholipid is followed by a reacylation reaction catalyzed by an acyltransferase.

SIGNAL TRANSDUCTION Numerous hormones initiate signal transduction mechanisms that involve phospholipid hydrolysis (pp. 609–10). For example, the PLC-catalyzed cleavage of PIP_2, a phosphorylated derivative of phosphatidylinositol, yields the signal molecules inositol-1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG). Synthesis of eicosanoids (p. 407) is initiated with the PLA_2-catalyzed release of arachidonic acid.

DIGESTION In mammals, fat digestion occurs in the small intestine where bile salts (p. 445) convert large fat globules into smaller droplets that can be acted on by enzymes. Pancreatic phospholipases, delivered to the small intestine along with other digestive enzymes, degrade dietary phospholipids. Lysosomal phospholipases degrade the phospholipid components of cellular membranes.

TOXIC PHOSPHOLIPASES Various organisms use membrane-degrading phospholipases as a means of inflicting damage on other species. PLA_2 is a component of snake and bee venoms. PLA_2

in snake venom not only digests cell membranes at the site of a bite, but also causes diverse forms of systemic damage (e.g., necrosis of skeletal and heart muscle, neurotoxicity, and red blood cell lysis). *Clostridium perfringins* is an anaerobic, Gram-positive bacterium that causes *gas gangrene* (tissue death accompanied by gas formation). It secretes a phospholipase called α-toxin that facilitates the organism's penetration into the tissue surrounding a wound.

Toxic Phospholipases

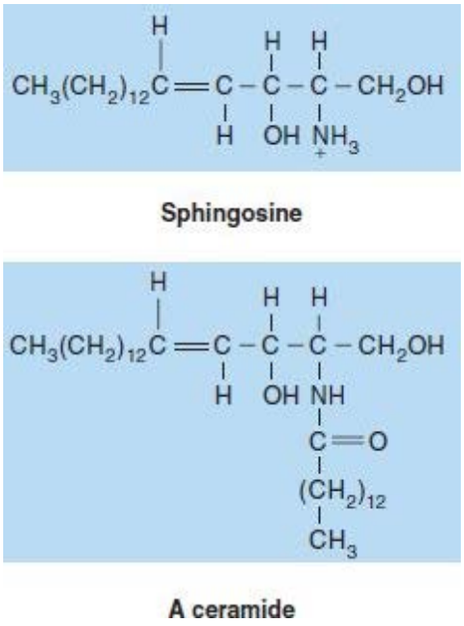


FIGURE 11.12

Sphingolipid Components

Note that the *trans*-isomer of sphingosine occurs in sphingolipids. Ceramide is a fatty acid amide derivative of sphingosine.

Sphingolipids

Sphingolipids are important components of animal and plant membranes. All sphingolipid molecules contain a long-chain amino alcohol. In animals, this alcohol is primarily sphingosine (**Figure 11.12**). The core of each type of sphingolipid is *ceramide*, a fatty acid amide derivative of sphingosine. In *sphingomyelin*, the 1-hydroxyl group of ceramide is esterified to the phosphate group of phosphorylcholine or phosphorylethanolamine (**Figure 11.13**). Sphingomyelin is found in most animal cell membranes. However, as its name suggests, sphingomyelin is found in greatest abundance in the myelin sheath of nerve cells. The myelin sheath is formed by successive wrappings of the cell membrane of a specialized myelinating cell around a nerve cell axon. Its insulating properties facilitate the rapid transmission of nerve impulses.

The **glycolipids** are lipid molecules with carbohydrate groups attached. They include the *glycosphingolipids* (**Figure 11.14**) and GPI anchors (p. 411). The glycosphingolipids contain ceramide, but they differ from sphingomyelin in that they contain no phosphate. The most important glycolipid classes are the cerebrosides, the sulfatides, and the gangliosides.

Cerebrosides are sphingolipids in which the head group is a monosaccharide. (These molecules,

unlike phospholipids, are nonionic.) Galactocerebrosides, the most common example of this class, are found almost entirely in the cell membranes of the brain. If a cerebroside is sulfated, it is referred to as a *sulfatide*. Sulfatides are negatively charged at physiological pH.

Sphingolipids that possess oligos accharide groups with one or more sialic acid residues are called *gangliosides*. Although gangliosides were first isolated from nerve tissue, they also occur in most other animal tissues. The names of gangliosides include subscript letter and numbers. The letters M, D, and T indicate whether the molecule contains one, two, or three sialic acid residues (see **Figure 7.25d**), respectively. The numbers designate the sequence of sugars that are attached to ceramide. The Tay–Sachs ganglioside G_{M2} is illustrated in **Figure 11.14**.

KEY CONCEPTS

- Sphingolipids, important membrane components of animals, contain a complex long-chain amino alcohol (sphingosine).
- The core of each sphingolipid is ceramide, a fatty acid amide derivative of the alcohol molecule. Glycolipids are derivatives of ceramide, which possess a carbohydrate component.

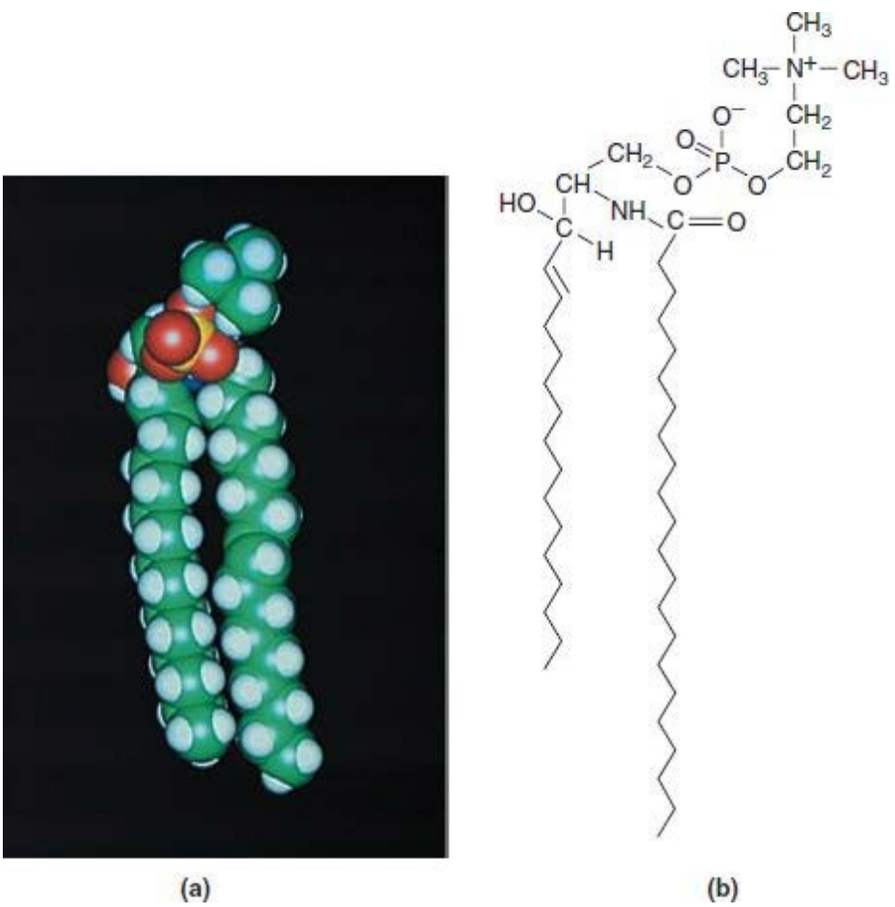


FIGURE 11.13

(a) Space-Filling and Conformational (b) Models of Sphingomyelin

The fatty acid component of sphingomyelins can be saturated or monounsaturated and 16 to 24 carbons in length, depending on the species and tissue of origin. The sphingosine base can be replaced by sphinganine (with no double bond) and other C-20 homologues, although sphingosine is by far the most abundant.

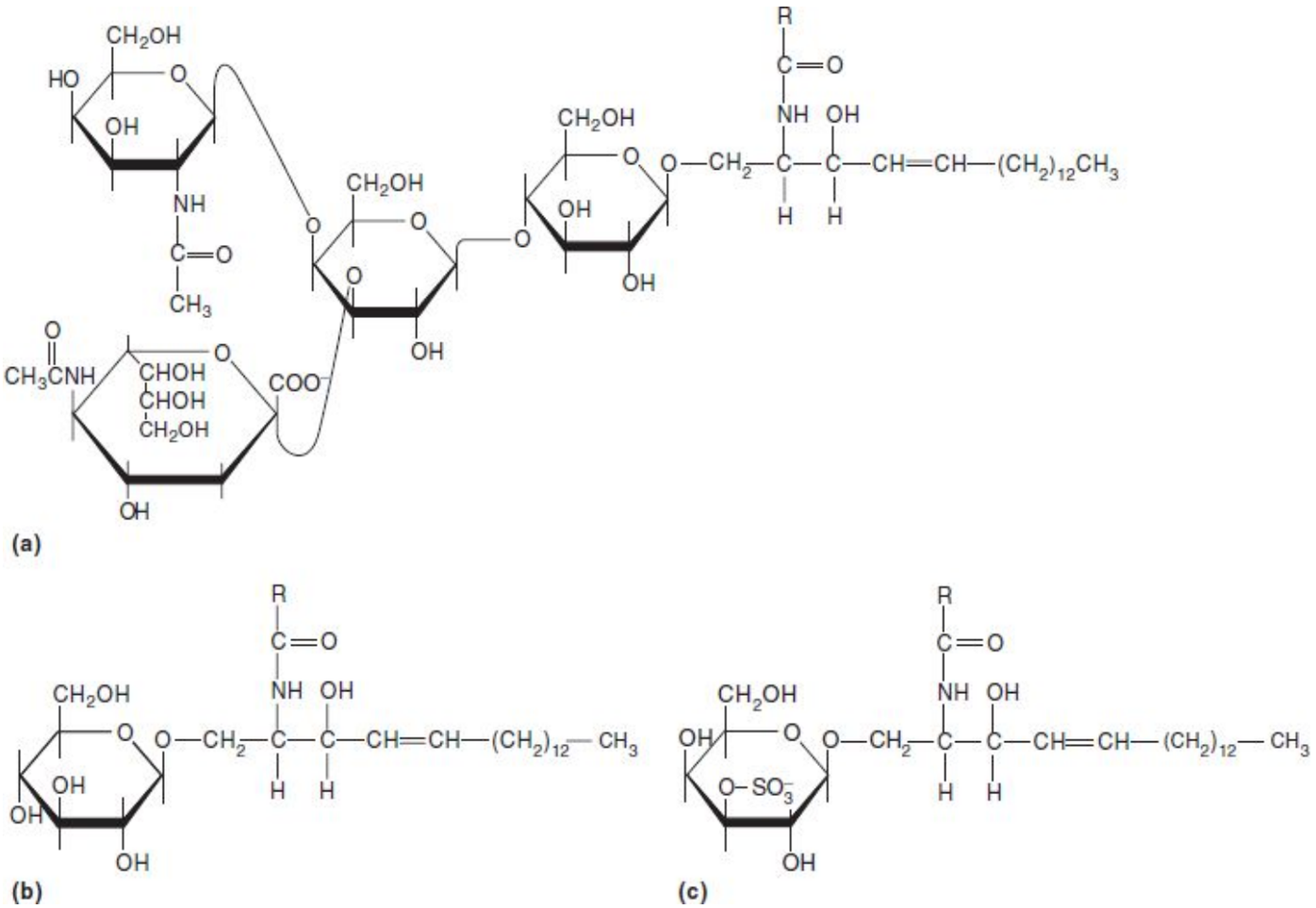


FIGURE 11.14

Structures of Selected Glycolipids

(a) Tay–Sachs ganglioside (G_{M2}), (b) glucocerebroside, and (c) galactocerebroside sulfate (a sulfatide).

The normal role of glycolipids is still unclear. Certain glycolipid molecules may bind bacterial toxins, as well as bacterial cells, to animal cell membranes. For example, the toxins that cause cholera, tetanus, and botulism bind to glycolipid cell membrane receptors. Bacteria that have been shown to bind to glycolipid receptors include *E. coli*, *Streptococcus pneumoniae*, and *Neisseria gonorrhoeae*, the causative agents of urinary tract infections, pneumonia, and gonorrhea, respectively.

Sphingolipidoses

Sphingolipid Storage Diseases

Each lysosomal storage disease (see p. 54) is caused by the hereditary deficiency of an enzyme required for degradation of a specific metabolite. Several lysosomal storage diseases are associated with sphingolipid metabolism. Most of these diseases, also referred to as the *sphingolipidoses*, are fatal. The most common sphingolipid storage disease, Tay–Sachs disease, is caused by a deficiency of *β*-hexosaminidase A, the enzyme that degrades the ganglioside G_{M2}. As cells accumulate this

molecule, they swell and eventually die. Tay–Sachs symptoms (blindness, muscle weakness, seizures, and mental retardation) usually appear several months after birth. Because there is currently no therapy for Tay–Sachs disease or for any other of the sphingolipidoses, the condition is always fatal (usually by age 3). Examples of the sphingolipidoses are summarized in **Table 11.3**.

Isoprenoids

The **isoprenoids** are a vast array of biomolecules that contain repeating five-carbon structural units known as *isoprene units* (**Figure 11.15**). Isoprenoids are not synthesized from isoprene (methylbutadiene). Instead, their biosynthetic pathways all begin with the formation of isopentenyl pyrophosphate from acetyl-CoA (Chapter 12).

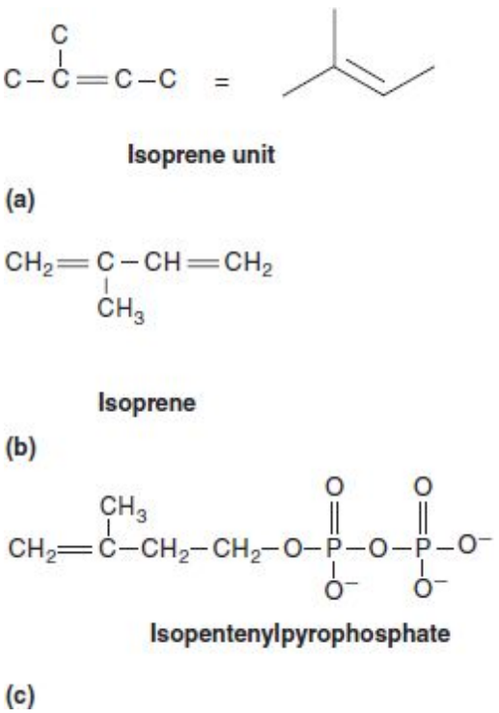


FIGURE 11.15

Isoprene

(a) Basic isoprene structure. (b) The organic molecule isoprene. (c) Isopentenylpyrophosphate.

3D animation of isoprene

The isoprenoids consist of terpenes and steroids. **Terpenes** are an enormous group of molecules that are found largely in the essential oils of plants. Steroids are derivatives of the hydrocarbon ring system of cholesterol.

TABLE 11.3 Selected Sphingolipid Storage Diseases*

Disease	Symptom	Accumulating Sphingolipid	Enzyme Deficiency

Tay–Sachs disease	Blindness, muscle weakness, seizures, mental retardation	Ganglioside G_{M2}	β-Hexosaminidase A
Gaucher's disease	Mental retardation, liver and spleen enlargement, erosion of long bones	Glucocerebroside	β-Glucosidase
Krabbe's disease	Demyelination, mental retardation	Galactocerebroside	β-Galactosidase
Niemann–Pick disease	Mental retardation	Sphingomyelin	Sphingomyelinase

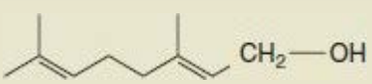
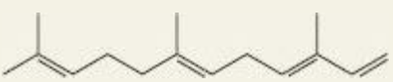
* Many diseases are named for the physicians who first described them. Tay–Sachs disease was reported by Warren Tay (1843–1927), a British ophthalmologist, and Bernard Sachs (1858–1944), a New York neurologist. Phillipe Gaucher (1854–1918), a French physician, and Knud Krabbe (1885–1961), a Danish neurologist, first described Gaucher's disease and Krabbe's disease, respectively. Niemann–Pick disease was first characterized by the German physicians Albert Niemann (1880–1921) and Ludwig Pick (1868–1944).

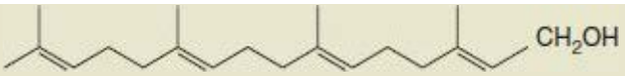
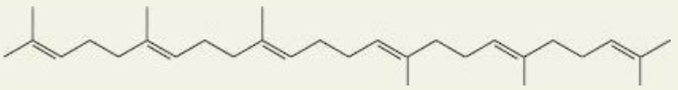
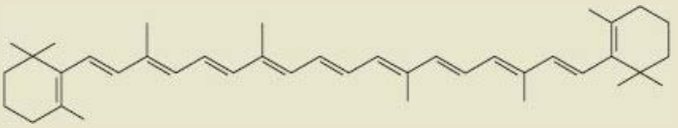
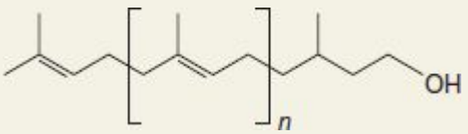
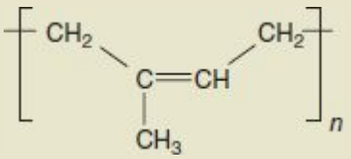
TERPENES The terpenes are classified according to the number of isoprene units they contain (**Table 11.4**). *Monoterpenes* are composed of two isoprene units (10 carbon atoms). Geraniol is a monoterpene found in the *essential oils*, volatile hydrophobic liquid mixtures extracted from plants, fruits, or flowers (e.g., roses, lemon, and geranium). Each essential oil has a characteristic odor, and some are used to make perfumes.

Terpenes that contain three isoprenes (15 carbons) are referred to as *sesquiterpenes*. Farnesene, an important constituent of oil of citronella, which is used in soap and perfumes, is a sesquiterpene. Phytol (p. 459), a plant alcohol, is a *diterpene*, a molecule composed of four isoprene units. Squalene is a prominent example of the *triterpenes*; this intermediate in the synthesis of the steroids is found in large quantities in shark liver oil, olive oil, and *yeast*. **Carotenoids**, the orange pigments found in most plants, are the only *tetraterpenes* (molecules composed of eight isoprene units). The *carotenes* are hydrocarbon members of this group. The *xanthophylls* are oxygenated derivatives of the carotenes. *Polyterpenes* are high-molecular-weight molecules composed of up to thousands of isoprene units. Natural rubber is a polyterpene composed of between 3000 and 6000 isoprene units. *Dolichols* are polyisoprenoid alcohols (16–19 isoprene units) that function as sugar carriers in glycoprotein synthesis (pp. 776–77).

Several important biomolecules are composed of nonterpene components attached to isoprenoid groups (often referred to as *prenyl* or *isoprenyl* groups). Examples of these biomolecules, referred to as **mixed terpenoids**, include vitamin E (α-tocopherol) (**Figure 10.27a**), ubiquinone (coenzyme Q) (**Figure 10.2**, vitamin K (**Figure 11.16**), and plastoquinone (p. 497).

TABLE 11.4 Examples of Terpenes

		Example	
Type	**Number of Isoprene Units**	**Name**	**Structure**
Monoterpene	2	Geraniol	
Sesquiterpene	3	Farnesene	

Diterpene	4	Phytol	
Triterpene	6	Squalene	
Tetraterpene	8	β - Carotene	
Polyterpene	9–24	Dolichol	
	Thousands	Rubber	

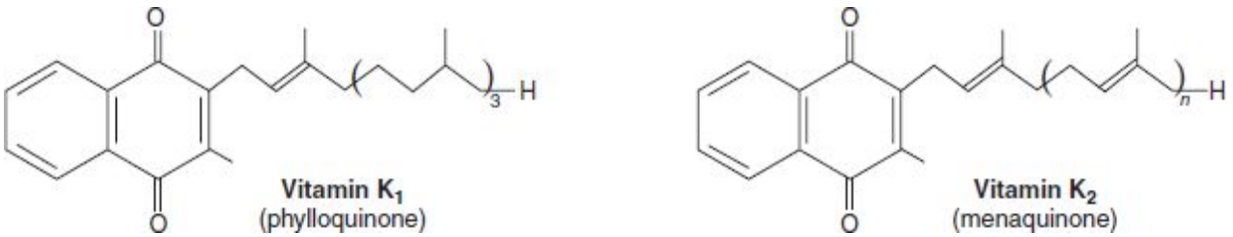


FIGURE 11.16

Vitamin K, a Mixed Terpenoid

Vitamin K_1 (phylloquinone) is found in plants, where it acts as an electron carrier in photosynthesis. Vitamin K_2 (menaquinone) is synthesized by intestinal bacteria and plays an important role in blood coagulation.

3D animation of Vitamin K1

A variety of proteins in eukaryotic cells are covalently attached to prenyl groups after their biosynthesis on ribosomes. The prenyl groups most often involved in this process, referred to as **prenylation**, are farnesyl and geranylgeranyl groups (**Figure 11.17**). Farnesyl and geranylgeranyl groups are intermediates in the cholesterol biosynthesis pathway (**Figure 12.29**). Apparently, the function of protein prenylation is the tethering of peripheral proteins to cell membranes. There is some evidence that it plays a role in the control of cell growth. For example, *Ras proteins*, a group of cell growth regulators, are activated by prenylation reactions.

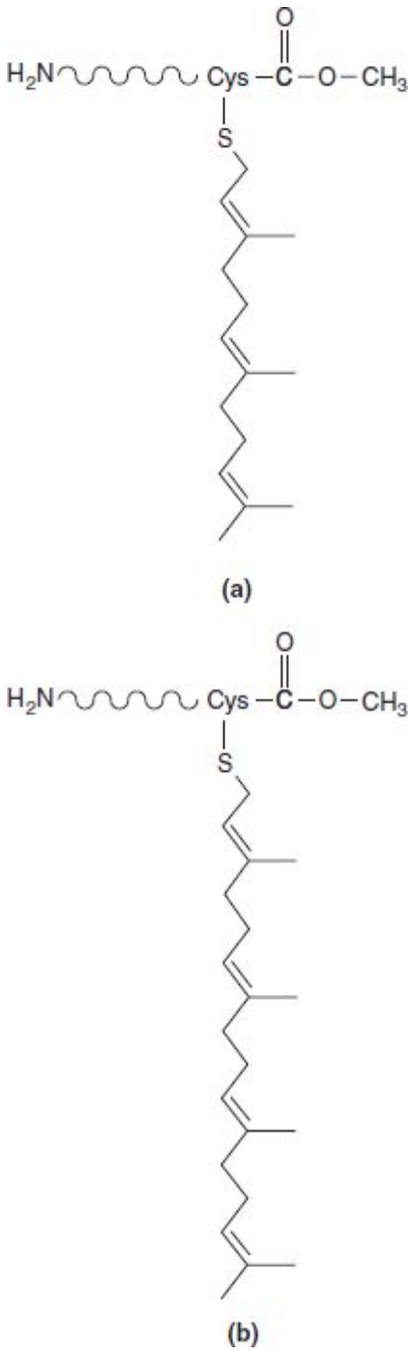


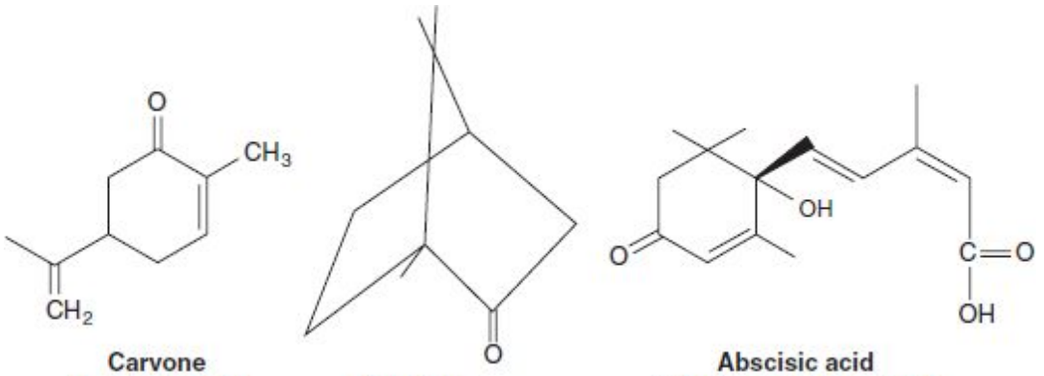
FIGURE 11.17

Prenylated Proteins

Prenyl groups are covalently attached at the SH group of C-terminal cysteine residues. Many prenylated proteins are also methylated at this residue. (a) Farnesylated protein. (b) Geranylgeranylated protein.

QUESTION 11.5

The majority of terpenes contain one or more ring structures. Consider the following examples. Determine which terpene class they belong to, and outline the positions of the isoprene units.



Carvone (spearmint oil) **Camphor** **Abscisic acid (plant growth regulator)**

3D animation of carvone

3D animation of camphor

3D animation of abscisic acid

STEROIDS **Steroids** are derivatives of triterpenes with four fused rings. They are found in all eukaryotes and a small number of bacteria. Steroids are distinguished from each other by the placement of carbon–carbon double bonds and various substituents (e.g., hydroxyl, carbonyl, and alkyl groups).

Cholesterol, an important molecule in animals, is an example of a steroid (**Figure 11.18**). In addition to being an essential component in animal cell membranes, cholesterol is a precursor in the biosynthesis of all steroid hormones, vitamin D, and bile salts (**Figure 11.19**). Cholesterol (C-27) is formed from the linear triterpene squalene (C-30) by intramolecular ring closure, oxidation, and cleavage. The only double bond retained migrates to the Δ5 position, and C-3 is oxidized to a hydroxyl group, which justifies its classification as a *sterol*. (Although the term *steroid* is most properly used to designate molecules that contain one or more carbonyl or carboxyl groups, it is often used to describe all derivatives of the steroid ring structure.) Cholesterol is usually stored within cells as a fatty acid ester. The esterification reaction is catalyzed by the enzyme *acyl-CoA:cholesterol acyltransferase* (ACAT), which is also referred to as sterol O-acetyltransferase (SOAT) to distinguish it from acetyl-CoA acetyltransferase (p. 457). ACAT is located on the cytoplasmic face of the ER. In liver cells, cholesterol esterification is involved in the synthesis of ApoB-containing lipoproteins.

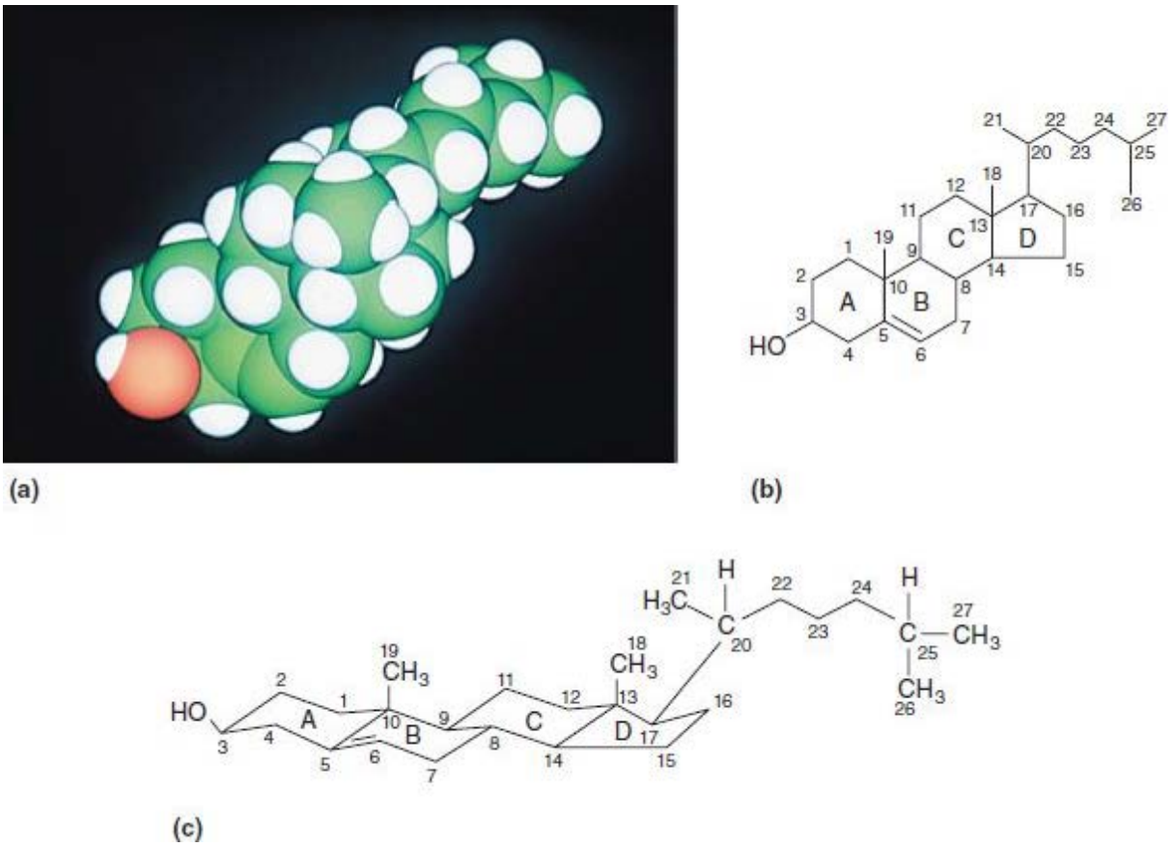


FIGURE 11.18

Structure of Cholesterol

(a) Space-filling model, (b) conventional view, and (c) conformational model. Space-filling models and conformational models represent molecular structure more accurately than the conventional view.

3D animation of cholesterol

Digitalis and Heart Failure

KEY CONCEPTS

- Isoprenoids are a large group of biomolecules with repeating units derived from isopentenyl pyrophosphate.
- There are two types of isoprenoids: terpenes and steroids.

Cardiac glycosides, molecules that increase the force of cardiac muscle contraction, are among the most interesting steroid derivatives. Glycosides are carbohydrate-containing acetals (see p. 255). Although several cardiac glycosides are extremely toxic (e.g., *ouabain*, obtained from the seeds of the plant *Strophanthus gratus* or climbing oleander), others have valuable medicinal properties (**Figure 11.20**). For example, *digitalis*, an extract of the dried leaves of *Digitalis purpurea* (the foxglove plant), is a time-honored stimulator of cardiac muscle contraction.

Digitoxin, the major "cardiotonic" glycoside in digitalis, is used to treat congestive heart failure, an illness in which the heart is so damaged by disease processes (e.g., myocardial infarcts) that pumping is impaired. In higher than therapeutic doses, digitoxin is extremely toxic. Both ouabain and digitoxin inhibit Na^+–K^+ ATPase (see p. 429).

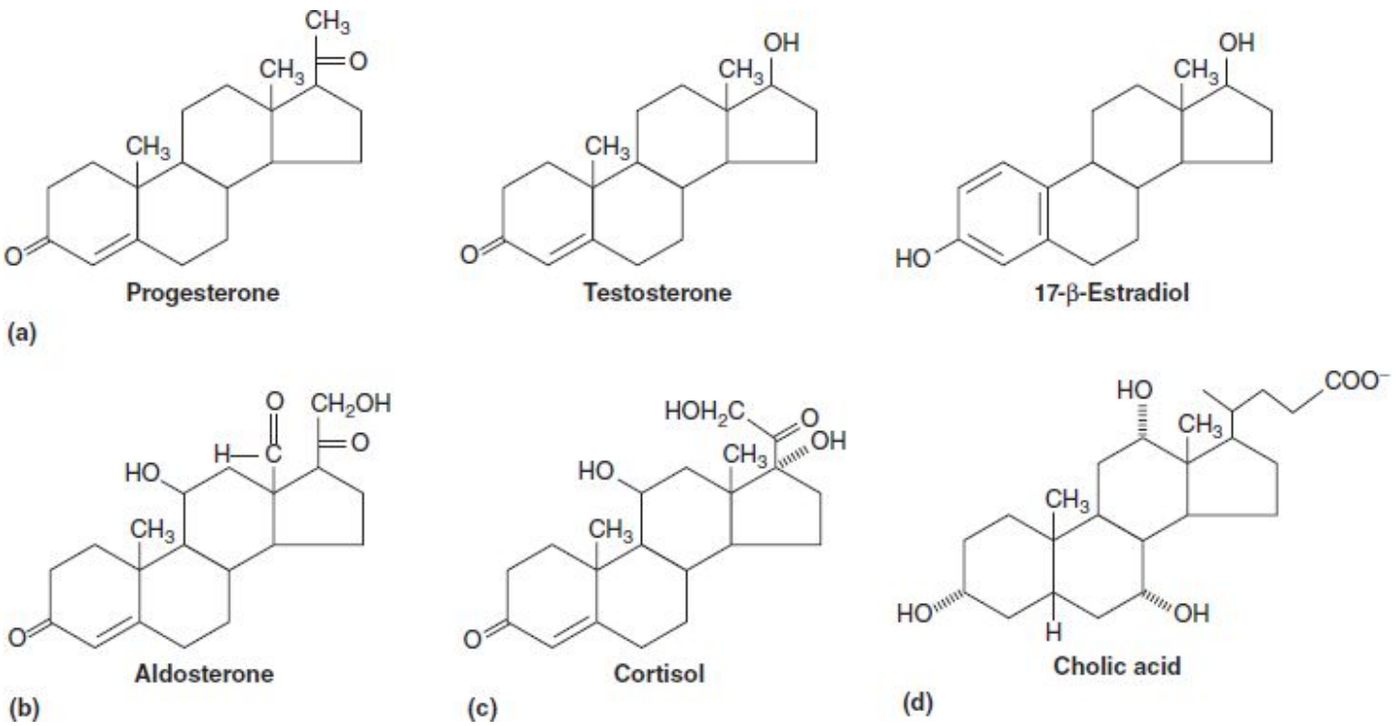


FIGURE 11.19

Animal Steroids

(a) Sex hormones (progesterone, testosterone and 17-*β*-estradiol) are molecules that regulate the development of primary and secondary sex characteristics and various reproductive behaviors). (b) Aldosterone is a mineralocorticoid (a molecule produced in the adrenal cortex) that regulates plasma concentrations of several ions, especially sodium. (c) Cortisol is a glucocorticoid (also an adrenal cortex molecule) that regulates the metabolism of carbohydrates, fats, and proteins. (d) Cholic acid is a bile acid that is converted into a bile salt in the liver. (Bile salts aid the absorption of dietary fats and fat-soluble vitamins in the intestine.)

3D animation of progesterone

3D animation of aldosterone

3D animation of cortisol

3D animation of cholic acid

QUESTION 11.6

Bile salts are emulsifying agents; that is, they promote the formation of mixtures of hydrophobic substances and water. Produced in the liver, bile salts assist in the digestion of fats in the small intestine. They are formed by linking bile acids to hydrophilic substances such as the amino acid glycine. After reviewing the structure of cholic acid in **Figure 11.19**, suggest how the structural features of bile salts contribute to their function.

Lipoproteins

Although the term *lipoprotein* can describe any protein that is covalently linked to lipid groups (e.g., fatty acids or prenyl groups), it is most often applied to a group of molecular complexes found in the blood plasma of mammals (especially humans). Plasma lipoproteins transport lipid molecules (triacylglycerols, phospholipids, and cholesterol) through the bloodstream from one organ to another. Lipoproteins also contain several types of lipid-soluble antioxidant molecules (e.g., α-tocopherol and several carotenoids). (The function of *antioxidants*, substances that protect biomolecules from free radicals, is described in Chapter 10.) The protein components of lipoproteins, called *apolipoproteins* or *apoproteins*, are synthesized in the liver, intestine, and brain. (See Biochemistry in Perspective: Alzheimer's Disease and Apolipoprotein E4 on pp. 436–37.) There are five major classes of apolipoproteins: A, B, C, D, and E. A generalized lipoprotein is shown in **Figure 11.21**. The relative amounts of lipid and protein components of the major types of lipoprotein are summarized in **Figure 11.22**.

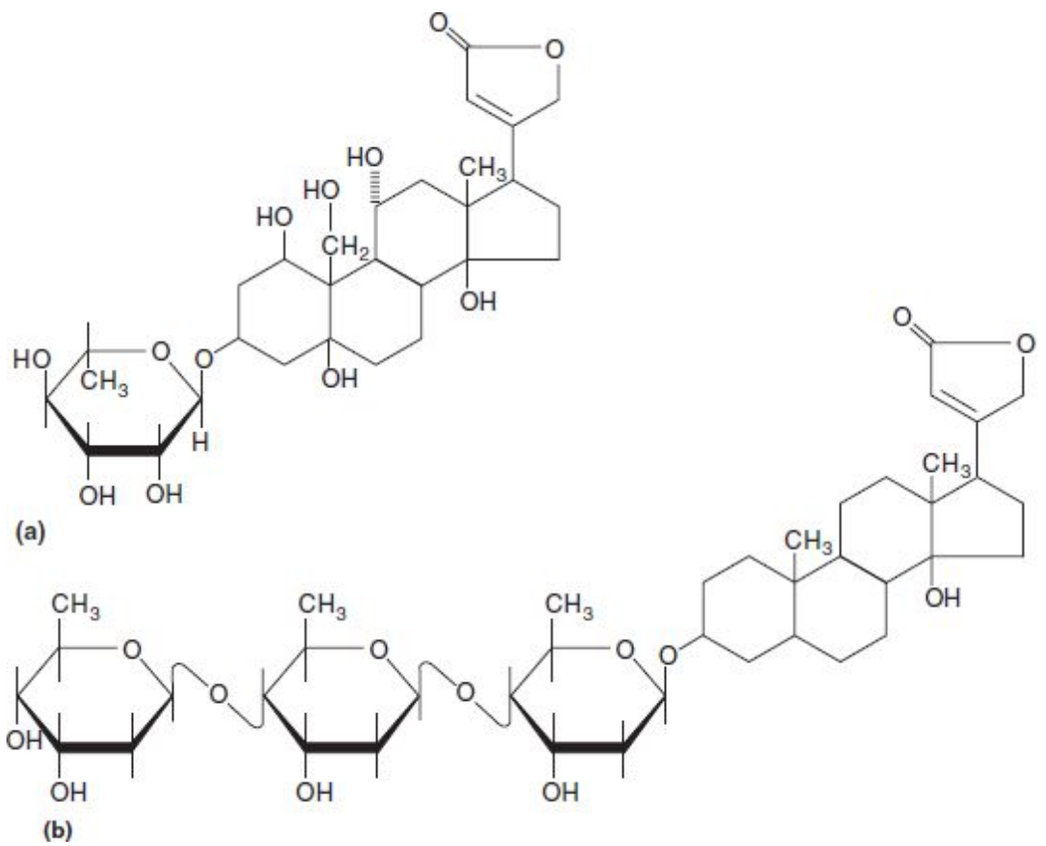


FIGURE 11.20

Cardiac Glycosides

Each cardiac glycoside possesses a glycone (carbohydrate) and an aglycone component. (a) In ouabain, the glycone is one rhamnose residue. The steroid aglycone of ouabain is called ouabagenin. (b) The glycone of digitoxin is composed of three digitoxose residues. The aglycone of digitoxin is called digitoxigenin. Cardiac glycosides also have a lactone ring attached at the 17-position of the steroid ring.

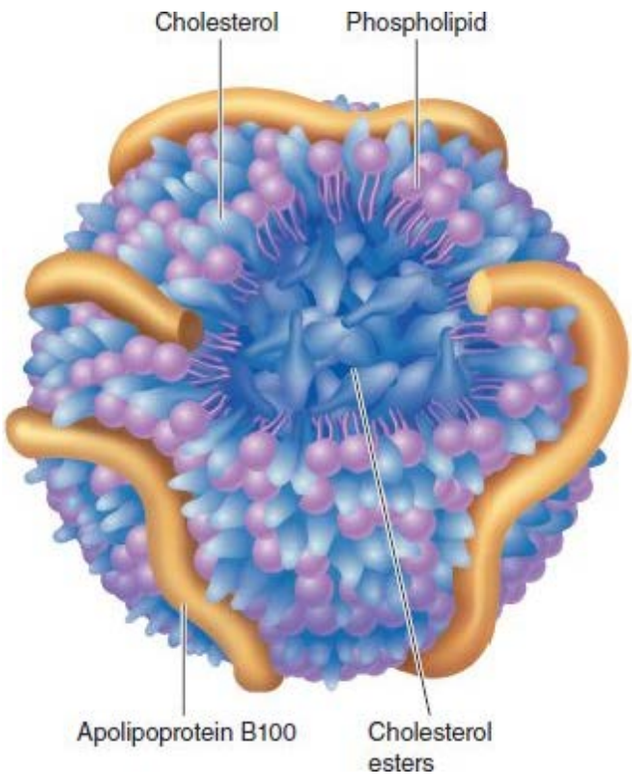


FIGURE 11.21

Plasma Lipoproteins

Lipoproteins vary in diameter from 5 to 1000 nm. Each type of lipoprotein contains a neutral lipid core composed of cholesteryl esters and/or triacylglycerols. This core is surrounded by a layer of phospholipid, cholesterol, and protein. Charged and polar residues on the surface of a lipoprotein enable it to dissolve in blood. In a low-density lipoprotein (LDL), as illustrated in this figure, each particle is composed of a core of cholesteryl esters and triglycerides, surrounded by a monolayer that consists of hundreds of cholesterol and phospholipid molecules and several apolipoproteins, including apolipoprotein B100, the ligand for the LDL receptor (p. 434).

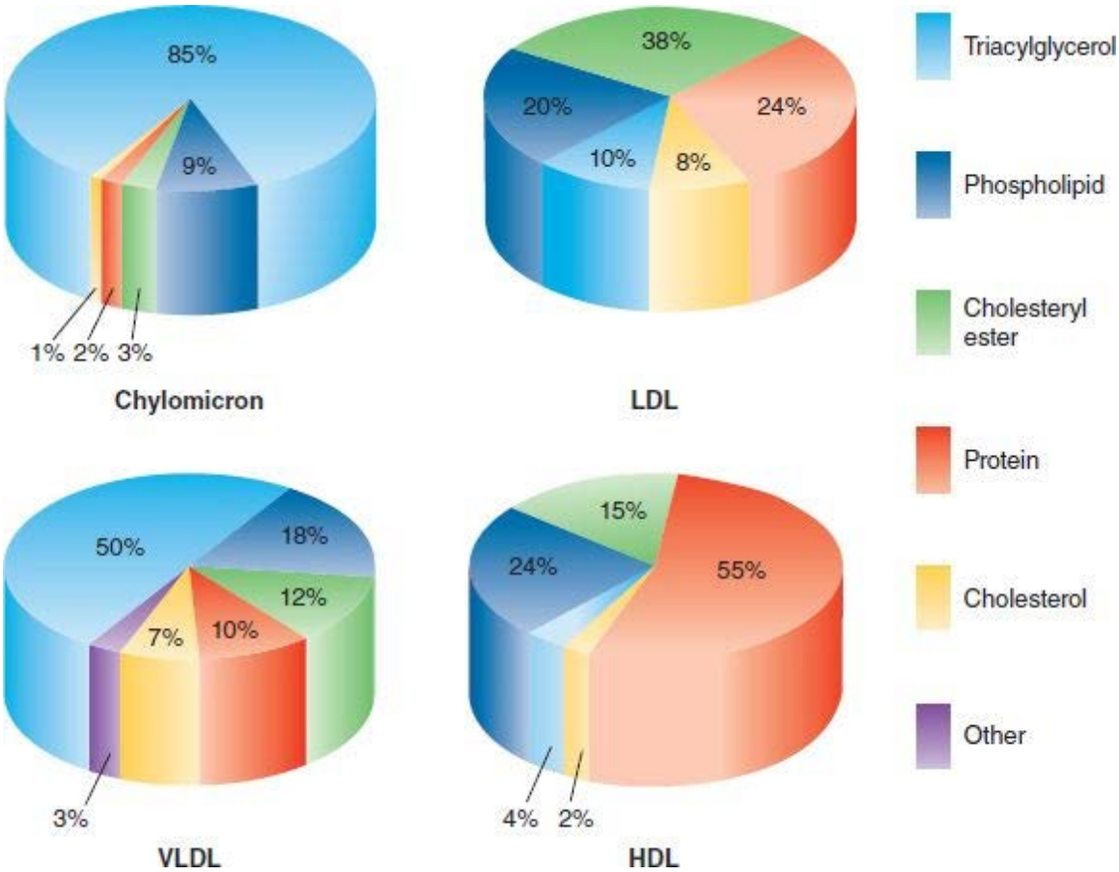


FIGURE 11.22

Proportional (Relative) Mass of Cholesterol, Cholesteryl Ester, Phospholipid, and Protein Molecules in Four Major Classes of Plasma Lipoproteins

Chylomicrons are the largest but least dense of the plasma lipoproteins because of their high content of triacylglycerols. In contrast, high-density lipoprotein (HDL) is a smaller dense particle that contains a high percentage by mass of protein and low triacylglycerol content.

Lipoproteins are classified according to their density. **Chylomicrons**, which are large lipoproteins (diameter 1000 nm) of extremely low density (<0.95 g/cm^3), transport dietary triacylglycerols and cholesteryl esters from the intestine to muscle and adipose tissues. Chylomicron remnants are then taken up by the liver via endocytosis. **Very-low-density lipoproteins** (VLDLs) (0.98 g/cm^3, D = 30–90 nm), synthesized in the liver, transport lipids to tissues. As VLDLs are depleted of triacylglycerol and some apolipoprotein and phospholipids, they shrink in size, become more dense, and are referred to as **intermediate-density lipoproteins** (IDLs) or VLDL remnants (1 g/cm^3, D = 40 nm). IDLs may continue to lose triacylglycerol to form a higher density lipoprotein called **low-density lipoprotein** (LDL), or they may be removed from the bloodstream by the liver.

LDL or IDL remnants (1.04 g/cm^3, D = 20 nm), are the principal transporters of cholesterol and cholesteryl esters to tissues. In an intricate process (p. 435) elucidated by Michael Brown and Joseph Goldstein (recipients of the 1985 Nobel Prize in Physiology or Medicine), LDL particles bind to LDL receptors and then are engulfed by cells via endocytosis. LDLs are also classified according to their diameters. LDLs with diameters less than 25 nm are referred to as small dense LDLs (sdLDLs). Those with diameters greater than 25 nm are called large buoyant LDL. sdLDLs are more atherogenic (i.e., prone to promotion of fatty plaques in arteries) than buoyant LDLs because they easily enter artery walls, where they are susceptible to oxidation. Risk factors for high

blood sdLDL levels include genetic predisposition, high-carbohydrate diet, physical inactivity, and insulin resistance (p. 616).

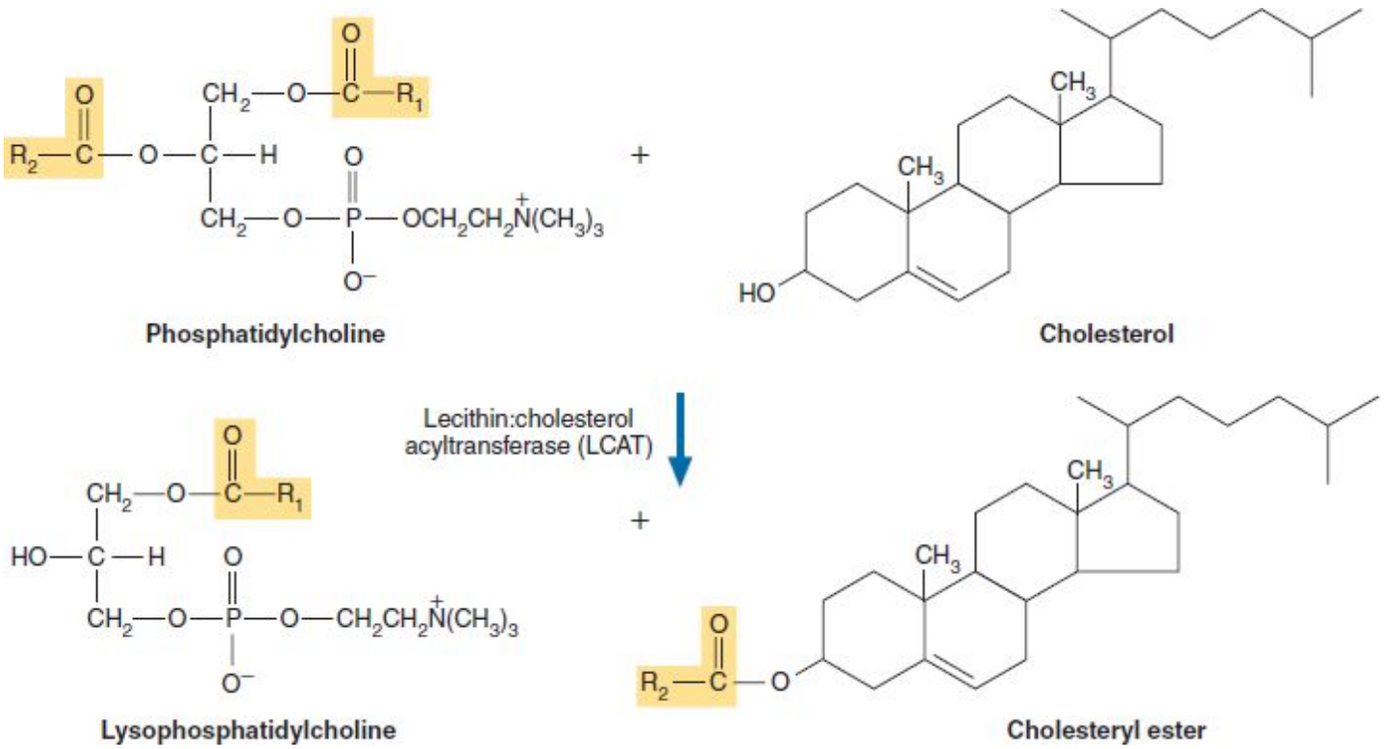


FIGURE 11.23

Reaction Catalyzed by Lecithin:Cholesterol Acyltransferase (LCAT)

Cholesteryl ester transfer protein, a protein associated with the LCAT–high- density lipoprotein (HDL) complex, transfers cholesteryl esters from HDL to very-low-density lipoprotein and LDL. The acyl groups are highlighted in color.

The role of **high-density lipoproteins** (HDLs) (1.2 g/cm^3, D = 9 nm), protein-rich particles produced in the liver and the intestine, appears to be the scavenging of excess cholesterol from cell membranes and cholesteryl esters from VLDL and LDL. Cholesteryl esters are formed when the plasma enzyme lecithin:cholesterol acyltransferase (LCAT) transfers a fatty acid residue from lecithin to cholesterol (**Figure 11.23**). HDL transports cholesteryl esters to the liver and steroid-producing organs such as the adrenal glands, ovary, and testes. The liver, the only organ that can dispose of excess cholesterol, converts most of it to bile acids (p. 480). The role of lipoproteins in atherosclerosis, a chronic disease of the cardiovascular system, is discussed in Chapter 12.

In addition to functioning as structural components, apolipoproteins have roles in the redistribution of lipids among various cells and tissues by acting as ligands for plasma membrane receptors for specific lipoproteins and as cofactors for certain enzymes. For example, apolipoprotein C-II (ApoC-II), a component of chylomicrons and VLDL, activates lipoprotein lipase, the enzyme in capillaries that hydrolyzes triacylglycerols to yield fatty acids for use by nearby cells. ApoA1, a component of HDL, is a cofactor for LCAT.

KEY CONCEPTS

- Plasma lipoproteins transport lipids through the bloodstream.

- On the basis of density, lipoproteins are classified into five major classes: chylomicrons, VLDL, IDL, LDL, and HDL.

11.2 MEMBRANES

Most of the properties attributed to living organisms (e.g., movement, growth, reproduction, and metabolism) depend, either directly or indirectly, on membranes. All biological membranes have the same general structure. As previously mentioned (Chapter 2), membranes contain lipid and protein molecules. In the currently accepted concept of membranes, referred to as the **fluid mosaic model**, or the Singer–Nicholson model, a membrane is a noncovalent **lipid bilayer** and associated proteins. The nature of these molecules determines each membrane's biological functions and mechanical properties. Because of the importance of membranes in biochemical processes, the remainder of this chapter is devoted to a discussion of their structure and functions.

TABLE 11.5 Chemical Composition of Some Cell Membranes

Membrane	Protein (%)	Lipid (%)	Carbohydrate (%)
Human erythrocyte plasma membrane	49	43	8
Mouse liver cell plasma membrane	46	54	2–4
Amoeba plasma membrane	54	42	4
Mitochondrial inner membrane	76	24	1–2
Spinach chloroplast lamellar membrane	70	30	6
Halobacterium purple membrane	75	25	0

Source: Guidotti (1972), Membrane proteins, Annu Rev Biochem 41:731.

Membrane Structure

Each type of living cell has a unique set of functions. Consequently, the membranes of each cell type have unique features. Not surprisingly, the proportions of lipid and protein vary considerably among cell types and among organelles within each cell (**Table 11.5**). The types of lipid and protein found in each membrane also vary.

MEMBRANE LIPIDS When amphipathic molecules are suspended in water, they spontaneously rearrange into ordered structures (**Figure 11.9**). As these structures form, hydrophobic groups become buried in the water-depleted interior. Simultaneously, hydrophilic groups become oriented so that they are exposed to water. Phospholipids form into bimolecular layers at relatively low concentration. This property of phospholipids (and other amphipathic lipid molecules) is the basis of membrane structure. Membrane lipids are largely responsible for several other important features of biological membranes as well.

MEMBRANE FLUIDITY. Membrane-dependent functions such as cell signaling and endocytosis appear to be regulated by plasma membrane fluidity. The term *fluidity* refers to the viscosity of the lipid bilayer (i.e., the degree of resistance of membrane components to movement). Membrane

fluidity is largely determined by the percentage of unsaturated fatty acids in its phospholipid molecules. (Recall that the hydrocarbon chains of unsaturated fatty acids pack less densely than saturated chains; see p. 404). A membrane's fluidity increases as its percentage of unsaturated fatty acids increases. Cholesterol contributes stability to animal cell membranes because of its rigid ring system and ability to form van der Waals interactions with contiguous hydrocarbon chains. However, fluidity remains high because of cholesterol's incomplete penetration into the membrane and the flexibility of its hydrocarbon tail (**Figure 11.24**).

Membrane fluidity is an important feature of biological membranes because rapid lateral movement (**Figure 11.25**) of lipid molecules is apparently responsible for the proper functioning of many membrane proteins. The movement of lipid molecules from one side of a lipid bilayer to the other occurs only during membrane synthesis or under conditions of lipid imbalance and requires the function of ATP-requiring mediator proteins, in the process called facilitated diffusion. *Flippase* transfers phospholipids from the outer to the inner membrane leaflet, whereas *floppase* transfers them in the opposite direction. *Scramblase* is a nonspecific, energy-independent redistributor of phospholipids across membranes. One measure of membrane fluidity, the ability of membrane components to diffuse laterally, can be demonstrated when two different cell types are fused to form a *heterokaryon*. (Certain viruses or chemicals are used to promote cell–cell fusion.) The plasma membrane proteins of each cell type can be tracked when they are labeled with different fluorescent markers. Initially, the proteins are confined to their own side of the heterokaryon membrane. As time passes, the two fluorescent markers intermix, indicating that the proteins observed in this type of experiment move freely in the lipid bilayer.

Selective Permeability. Hydrophobic hydrocarbon chains in lipid bilayers provide a virtually impenetrable barrier to the transport of ionic and polar substances. Specific membrane proteins regulate the movement of such substances into and out of cells. To cross a lipid bilayer, a polar substance must shed some or all of its hydration sphere and bind to a carrier protein for membrane translocation or pass through an aqueous protein channel. Both methods shield the hydrophilic molecule from the hydrophobic core of the membrane. Water crosses the membrane through protein channels called *aquaporins* (p. 430), which exhibit cell-type variation in their permeability to both water and accompanying ions. Nonpolar substances simply diffuse through the lipid bilayer down their concentration gradients. Each membrane exhibits its own transport capability or selectivity based on its protein component.

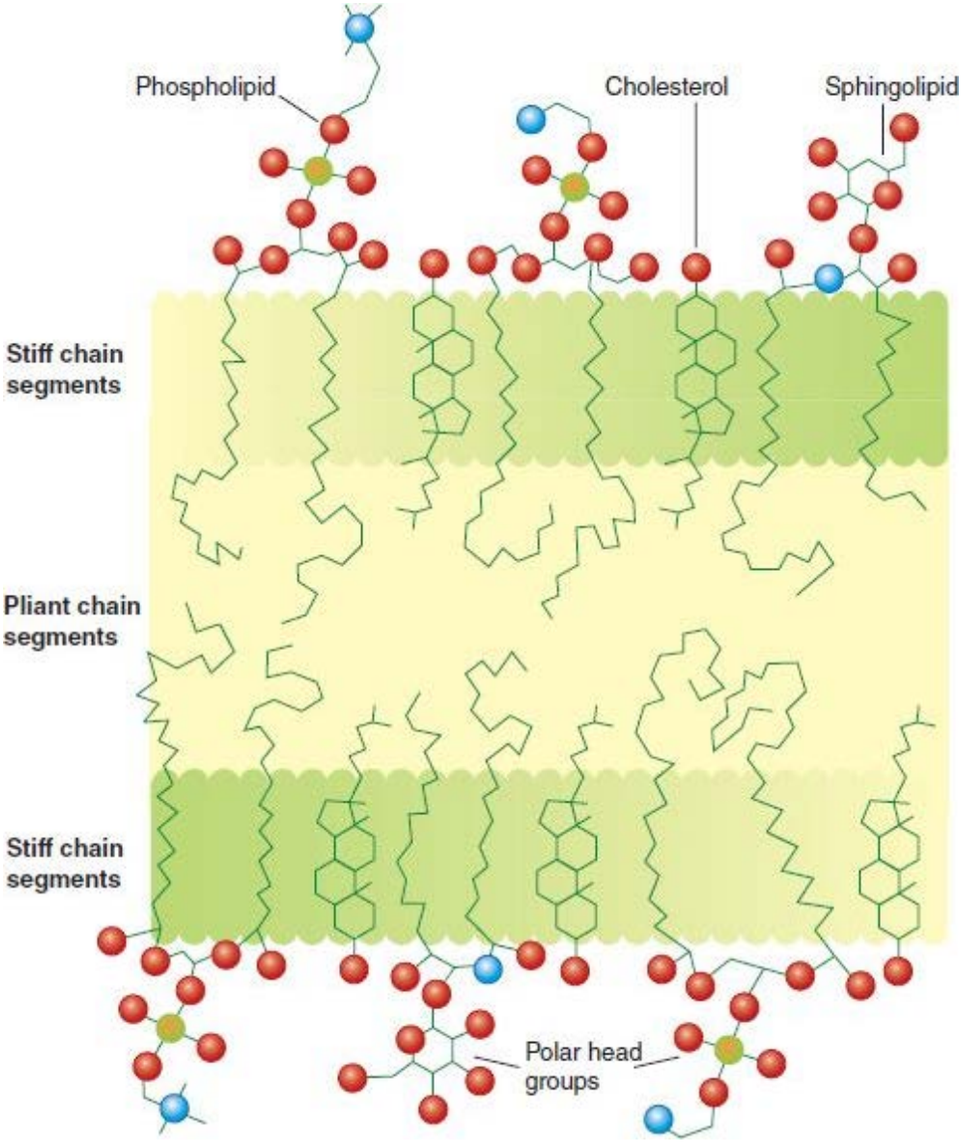


FIGURE 11.24

Diagrammatic View of a Lipid Bilayer

The flexible hydrocarbon chains in the hydrophobic core (lightly shaded area in the middle) make the membrane fluid. The phospholipids in the membrane have different levels of unsaturation and vary in the nature of the polar head group. Cholesterol's compact and rigid ring system creates structural stability in the outer region of each leaflet. Red blood cells and other cells that are subjected to mechanical stress have a high content of cholesterol and cardiolipin (two phospholipids linked by a glycerol). Cell membranes are about 7 to 9 nm thick. (Nitrogen atoms are blue, oxygen atoms are red, and phosphorus atoms are orange.)

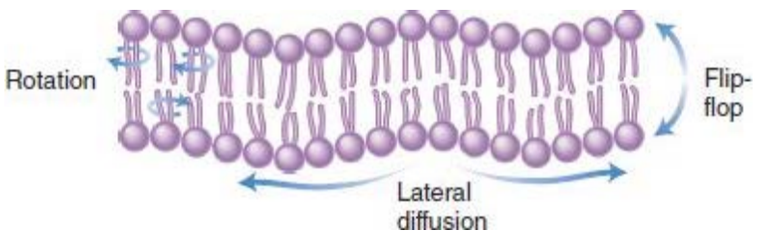


FIGURE 11.25

Lateral Diffusion in Biological Membranes

Lateral movement of phospholipid molecules is usually relatively rapid. "Flip-flop," the transfer of a lipid molecule from one side of a lipid bilayer to the other, occurs during new-membrane synthesis and membrane remodeling. Rotation of phospholipids within cell membranes is rapid.

Self-Sealing. When lipid bilayers are disrupted, they soon reseal. Small plasma membrane breaks spontaneously seal via the lateral flow of lipid molecules. Repair of larger lesions caused by mechanical stress, however, is an energy-requiring, Ca^{2+}-dependent process. A tear in the plasma membrane results in the inward flow of Ca^{2+} down its concentration gradient. Calcium ions trigger the movement of nearby endomembrane-derived vesicles to the lesion site. In an exocytosis-like process that involves cytoskeletal rearrangements, motor proteins such as dynein and kinesin, and membrane fusion proteins, the vesicles fuse with the plasma membrane to form a membrane patch. The repair process is rapid, usually occurring within a few seconds of the traumatic event.

Asymmetry. Biological membranes are asymmetric; that is, the lipid composition of each half of a bilayer is different. New membrane is synthesized by insertion of additional phospholipid molecules from the cytoplasmic face of existing membranes. Lipid molecules are transferred by mediator proteins to the opposite leaflet until membrane stability has been attained. Because the two faces of the resulting membrane are not chemically equivalent, the resulting leaflets are not identical in lipid composition. For example, the human red blood cell membrane possesses substantially more phosphatidylcholine and sphingomyelin on its outside surface. Most of the membrane's phosphatidylserine and phosphatidylethanolamine are on the inner side. The protein components of membranes (discussed next) also exhibit considerable asymmetry, with distinctly different functional domains within the membrane and on the cytoplasmic and extracellular faces of the membrane.

MEMBRANE PROTEINS Most of the functions associated with biological membranes require protein molecules. Membrane proteins are often classified by the function they perform: structure, transport, catalysis, signal transduction, or immunological identity. Membrane proteins are also classified according to their structural relationship to membrane. Proteins that are embedded in and/or extend through a membrane are referred to as **integral proteins** (**Figure 11.26**). Such molecules can be extracted only by disrupting the membrane with organic solvents or detergents. **Peripheral proteins** are bound to membrane primarily through noncovalent interactions with integral membrane proteins or covalent bonds to myristic, palmitic, or prenyl groups. GPI anchors link a wide variety of cell-surface proteins (e.g., lipoprotein lipase, folate receptor, alkaline phosphatase, and the core proteins of glypicans) to plasma membranes. Some peripheral proteins interact directly with the lipid bilayer.

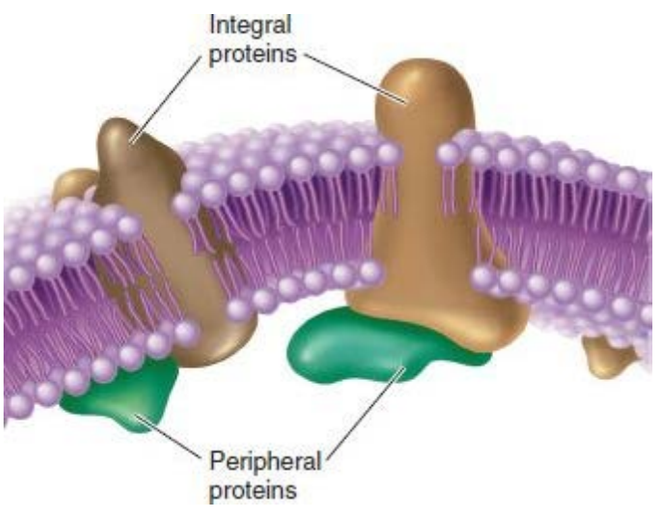


FIGURE 11.26

Integral and Peripheral Membrane Proteins

Integral membrane proteins are released only if the membrane is disrupted with detergents. Many peripheral membrane proteins can be removed with mild reagents, such as high salt concentration.

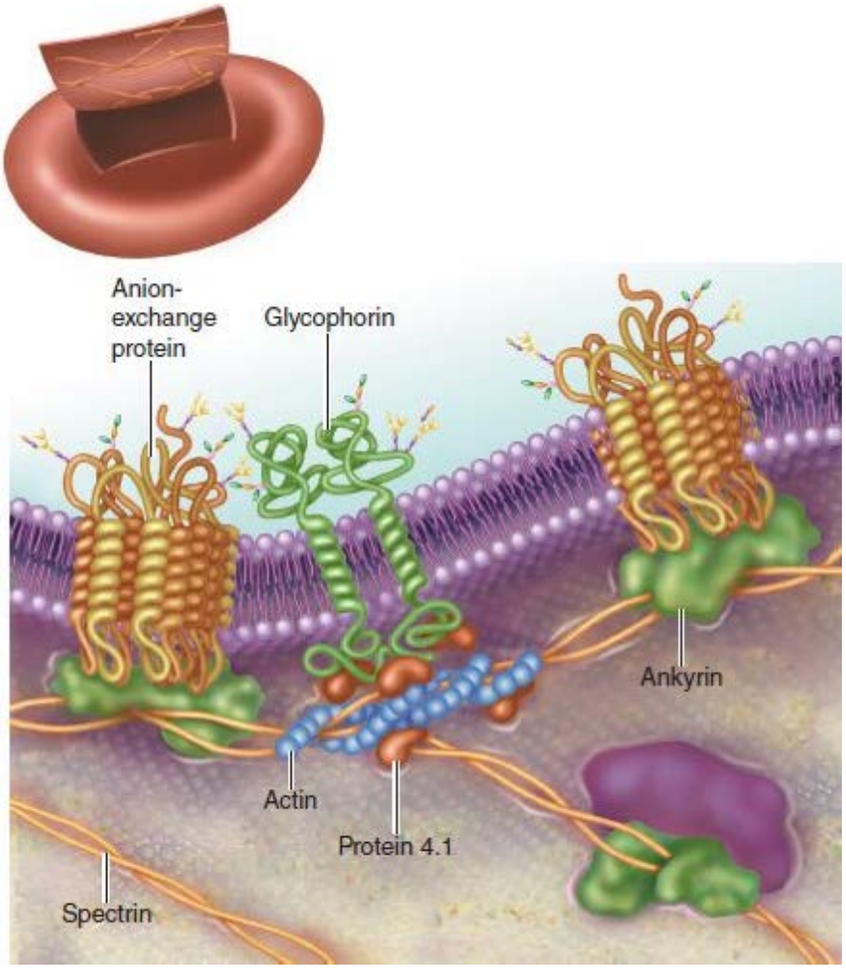


FIGURE 11.27

Red Blood Cell Integral Membrane Proteins

The integral membrane proteins glycophorin and anion exchange protein are components in a network of linkages that connect the plasma membrane to structural elements of the cytoskeleton (e.g., actin, spectrin, protein 4.1, and ankyrin). Note that the oligosaccharides on glycophorin are the ABO and MN blood group antigens.

KEY CONCEPTS

- The basic structural feature of membrane structure is the lipid bilayer, which consists of phospholipids and other amphipathic lipid molecules.
- Membrane proteins, embedded in or associated with the phospholipid bilayer, contribute specialized functions to the membrane, depending on its cell type and its role in biological processes.

The band 3 anion exchanger protein (AE1) (**Figure 11.27**), found in red blood cell membrane, is a well-researched example of integral membrane protein. AE1 is composed of two identical subunits, each consisting of 929 amino acids. With more than 1 million copies per cell, this protein channel plays an important role in CO_2 transport in blood. The HCO_3^- ion formed from CO_2 with the aid of carbonic anhydrase diffuses into and out of the red blood cell through the anion channel

in exchange for chloride ion Cl^-. (The exchange of Cl^- for HCO_3^-, called the *chloride shift*, preserves the electrical potential of the red blood cell membrane.)

Red blood cell membrane peripheral proteins are composed largely of spectrin, ankyrin, and protein 4.1. They are primarily involved in preserving the cell's unique biconcave shape, which maximizes the surface area-to-volume ratio and exposure of diffusing O_2 to intracellular hemoglobin. *Spectrin* is a tetramer, composed of two $\alpha\beta$ dimers, that binds to ankyrin and band 4.1. *Ankyrin* is a large globular polypeptide (215 kDa) that links spectrin to the anion channel protein. (This is a connecting link between the red blood cell's cytoskeleton and its plasma membrane.) *Protein 4.1* binds to both spectrin and *actin filaments* (a cytoskeletal component). Because protein 4.1 also binds to the membrane spanning protein glycophorin, it too links the cytoskeleton and the membrane.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on botulism and membrane fusion.

THE BAND 3 ANION EXCHANGER MACROCOMPLEX AND THE CHLORIDE SHIFT The red blood cell protein AE1, also found in the acid-secreting cells of the kidney, contains three domains: a 12-segment membrane-spanning domain that carries out anion exchange; an extended cytoplasmic N-terminal domain; and a short cytoplasmic C-terminal domain. In its tetramer form, AE1 is a macrocomplex with numerous integral and peripheral membrane proteins. The transmembrane domains bind to glycophorin and AQP1 (a water channel; see p. 430). The C-terminal domains bind carbonic anhydrase (CA), and the N-terminal domains bind several glycolytic enzymes (e.g., phosphofructokinase, aldolase, and glyceraldehyde-3-phosphate dehydrogenase), deoxyhemoglobin, and ankyrin, among others.

As freshly oxygenated blood reaches the tissues, carbon dioxide molecules released by respiring cells enter red blood cells, where CA rapidly converts them into HCO_3^- and H^+. Bicarbonate ions then leave the cell via the AE1 channel in exchange for chloride ions. The water molecules required for the CA-catalyzed reaction are supplied by AQP1. Excess protons cause the intracellular pH to drop, which decreases the affinity of hemoglobin for oxygen (the Bohr effect; see p. 179). As deoxyHb forms, a few molecules displace the glycolytic enzymes from the N-terminal domains of AE1 in a process that increases their catalytic activity. Two products of glycolysis, 2,3-bisphosphoglycerate (BPG; p. 179) and ATP, have unique roles in erythrocytes. As described previously, BPG binds to and stabilizes deoxyHb, thereby promoting oxygen release to the tissues. As red blood cells squeeze through narrow capillaries, ATP drives the membrane pumps that restore the ion concentrations disrupted by leakage caused by mechanical stress.

When blood flows through the lungs, its pH and oxygen level rise, and the process reverses. The T-to-R shift in hemoglobin conformation (p. 179) occurs as oxygen binds to hemoglobin and deoxyHb is destabilized by the release of BPG, protons, and other allosteric effectors. The resulting shift in the chemical equilibrium of the CA-catalyzed reaction is in favor of the conversion of HCO_3^- to CO_2. Bicarbonate enters the cell via AE1 in exchange for chloride ions, and CO_2 flows out of the blood down its concentration gradient and into the lung's alveolar cells. The release of newly oxygenated hemoglobin from the N-terminal domain of AE1 facilitates the rebinding of the glycolytic enzymes, depressing their activity. As a result, more glucose molecules are diverted into the pentose phosphate pathway, causing the synthesis of higher levels of NADPH. NADPH is required in oxygenated red blood cells for reduction of the ferrous iron in methemoglobin (p. 180)

and for protection of the red blood cell membrane from oxidative stress.

MEMBRANE MICRODOMAINS The lipids and proteins in membranes are not distributed uniformly. A notable example is provided by "lipid rafts," specialized microdomains in the external leaflet of eukaryotic plasma membranes (**Figure 11.28**). The components of lipid rafts are primarily cholesterol and sphingolipids and certain membrane proteins. The rigid fused rings of cholesterol pack tightly alongside the more saturated acyl chains of sphingolipid molecules. Consequently, the lipid molecules in these microdomains are more ordered (i.e., less fluid) than those in nonraft regions, where unsaturated acyl chains (in phospholipids) are more common. As their name suggests, lipid rafts seemingly float in a sea of more loosely packed membrane lipids.

Lipid rafts are enriched in some classes of proteins, and they exclude others. Proteins associated with lipid rafts include GPI-anchored proteins, doubly acetylated tyrosine kinases, and certain transmembrane proteins. Some proteins are always present in lipid rafts, whereas others enter lipid rafts only as a result of an activation process. Lipid rafts have been implicated in a variety of cellular processes; examples include exocytosis, endocytosis, and signal transduction. Lipid rafts are believed to function as platforms where the molecules that drive these processes are spatially organized. Caveolae, small invaginations of the plasma membrane of endothelial cells and adipocytes, are a special type of lipid raft. Caveolar membrane curvature occurs as a result of the formation of caveolin oligomers. (Refer to the online Chapter 2 Biochemistry in Perspective essay, Caveolar Endocytosis.)

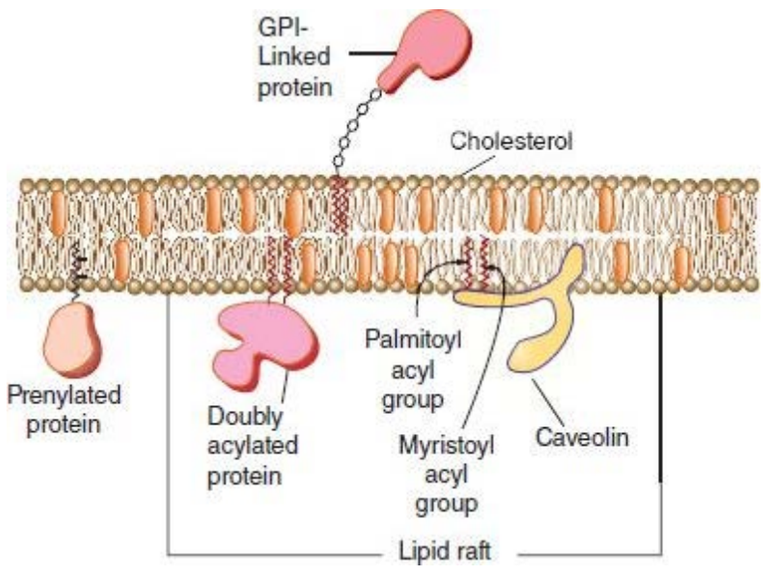


FIGURE 11.28

Lipid Rafts

As a result of stable associations between cholesterol and sphingolipids, slightly thicker membrane microdomains called lipid rafts form. Lipid rafts are also enriched in specific types of membrane protein. (Caveolin is a membrane protein found in caveolae, curved lipid rafts involved in clathrin-independent endocytosis and other processes.)

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on caveolar endocytosis in Chapter 2.

Membrane Function

Among the vast array of membrane functions are the transport of polar and charged substances into and out of cells and organelles and the relay of signals that initiate change in metabolic and developmental aspects of cell function.

MEMBRANE TRANSPORT Membrane transport mechanisms are vital to living organisms. Ions and molecules constantly move across cell plasma membranes and across the membranes of organelles. This flux must be carefully regulated to meet each cell's metabolic needs. Additionally, the plasma membrane regulates intracellular ion concentrations. Because lipid bilayers are generally impenetrable to ions and polar substances, specific transport components must be inserted into cellular membranes. Several examples of these structures, referred to as transport proteins, are discussed.

Biological transport mechanisms are classified according to whether they require energy. Major types of biological transport are illustrated in **Figure 11.29**. In **passive transport** (simple and facilitated diffusion), substances are moved across the membrane down their concentration gradients (i.e., from an area of high concentration to an area of low concentration) and require no direct input of energy. In contrast, **active transport** requires energy to transport molecules against a concentration gradient.

In **simple diffusion**, each solute, propelled by random molecular motion, moves down its concentration gradient. In this spontaneous process, there is a net movement of solute until an equilibrium is reached. A system reaching equilibrium becomes more disordered: that is, entropy increases. Because there is no input of energy, transport occurs with a negative change in free energy. In general, the higher the concentration gradient, the faster the rate of solute diffusion. The diffusion of gases such as O_2 and CO_2 across membranes is proportional to their concentration gradients. The diffusion of nonpolar organic molecules (such as steroid hormones) also depends on molecular weight and lipid solubility.

In **facilitated diffusion**, certain large or charged molecules are moved through special channels or carriers. *Channels* are tunnel-like transmembrane proteins. Each type usually transports a specific solute. Many channels are regulated. Chemically regulated channels open or close in response to a specific chemical signal. For example, a *chemically gated Na^+ channel* in the nicotinic acetylcholine receptor complex (found in muscle cell plasma membranes) opens when the neurotransmitter molecule acetylcholine binds. Then Na^+ rushes into the cell, and the membrane potential falls. Because membrane potential is an electrical gradient across the membrane (see p. 92), a decrease in membrane potential is membrane *depolarization*. Local depolarization caused by acetylcholine leads to the opening of nearby *voltage-gated Na^+ channels*. *Repolarization*, the reestablishment of the membrane potential, begins with the diffusion of K^+ ions out of the cell through *voltage-gated K^+ channels*. (The diffusion of K^+ ions out of the cell makes the inside less positive, that is, more negative.)

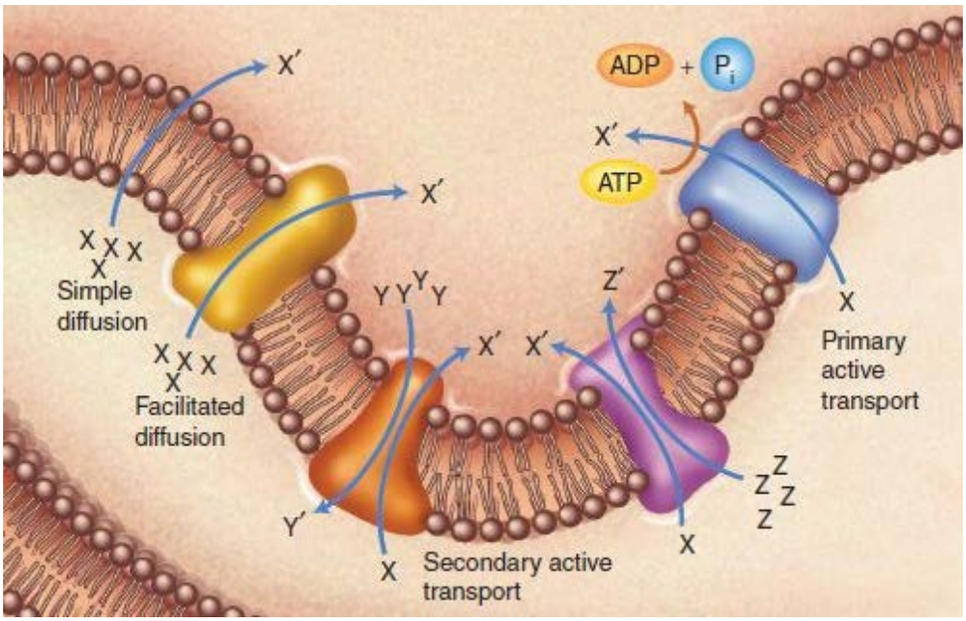


FIGURE 11.29

Transport across Membranes

The major transport processes are simple and facilitated diffusion and primary and secondary active transport. In simple diffusion, the spontaneous transport of a specific solute is driven by its concentration gradient. Facilitated diffusion, the movement of a solute down its concentration gradient across a membrane, occurs through protein channels or carriers. Both primary and secondary active transport require energy to move solutes across a membrane against their concentration gradients. In primary active transport, this energy is usually provided directly by ATP hydrolysis. In secondary active transport, solutes (X) are moved across a membrane by energy stored in a concentration gradient of a second substance (Y or Z) that has been created by ATP hydrolysis or another energy-requiring mechanism.

Another form of facilitated diffusion involves membrane proteins called *carriers* (sometimes referred to as *passive transporters*). In carrier-mediated transport, a specific solute binds to the carrier on one side of a membrane and causes a conformational change in the carrier. The solute is then translocated across the membrane and released. The red blood cell *glucose transporter* is the best-characterized example of passive transporters. It allows D-glucose to diffuse across the red blood cell membrane down its concentration gradient for use in glycolysis and the pentose phosphate pathway.

The two forms of active transport are primary and secondary. In **primary active transport**, energy is provided by ATP. Transmembrane ATP-hydrolyzing enzymes use the energy derived from ATP to drive the transport of ions or molecules. The Na^+–K^+ pump (also referred to as the Na^+–K^+ ATPase) is a prominent example of a primary transporter. (The Na^+ and K^+ gradients are required for maintaining normal cell volume and membrane potential. Refer to the Biochemistry in Perspective essay entitled Cell Volume Regulation and Metabolism in Chapter 3). In **secondary active transport**, concentration gradients generated by primary active transport are harnessed to move substances across membranes. For example, the Na^+ gradient created by the Na^+–K^+ ATPase pump is used in kidney tubule cells and intestinal cells to transport D-glucose against its concentration gradient (**Figure 11.30**). Two examples of membrane transport proteins, the aquaporins and cystic fibrosis-related chloride channel, are described next.

KEY CONCEPTS

- Membrane transport mechanisms are classified as passive or active according to whether they require energy.
- In passive transport, solutes moving across membranes move down their concentration gradient.
- In active transport, energy derived directly or indirectly from ATP hydrolysis or another energy source is required to move an ion or molecule against its concentration gradient.

QUESTION 11.7

The H^+–K^+ ATPase is a transporter in the membrane of several types of body cells, most notably in the luminal side of the stomach's parietal cells. ATP hydrolysis drives the transport of protons across the cell membrane in exchange for K^+. Proton secretion into gastric juice, composed of water, electrolytes, and enzymes, lowers its pH to about 1. Chloride ions move into the lumen down its concentration gradient through a membrane channel to yield HCl, which is required for protein digestion.

Alligators are predators that require exceptionally large amounts of HCl to digest their relatively large meals. They accomplish this feat in part with a vascular shunt that diverts venous blood to the stomach before returning it to the lungs for oxygenation. Why does this mechanism allow these animals to generate large amounts of HCl? [*Hint:* Review the discussion of carbonic anhydrase on pp. 101 and 426. Note that chloride ions enter the parietal cell from blood in exchange for bicarbonate.]

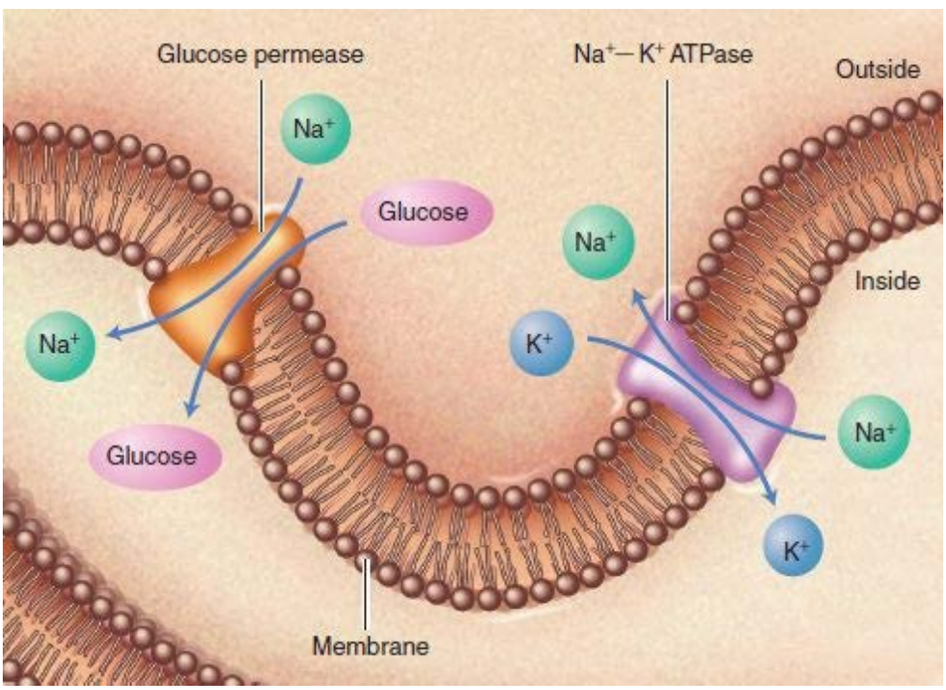


FIGURE 11.30

The Na^+–K^+ ATPase and Glucose Transport

The Na^+–K^+ ATPase preserves the Na^+ gradient essential to maintain membrane potential. In certain cells, glucose transport depends on the Na^+ gradient. Glucose permease in the apical membrane of kidney proximal tubule cells transports both Na^+ and glucose into the cells. Only when both substrates are bound does the protein change its conformation, thus initiating transport. Once they enter the tubular cell the glucose molecules

are transported down their concentration gradient across the basolateral membrane and into the blood by a facilitative glucose transporter (GLUT1 or 2) (not shown).

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on cell volume regulation and metabolism.

THE AQUAPORINS A basic characteristic of living cells is the capacity to rapidly move water across cell membranes in response to changes in osmotic pressure. For years, many researchers assumed that simple diffusion was responsible for most water flow. In a wide variety of cell types such as red blood cells and certain kidney cells, however, water flow is extraordinarily rapid. In the early 1990s, investigators characterized the first of a series of water channel proteins, now called the aquaporins. Found initially in red blood cell membrane and then in kidney tubule cells, **aquaporin** 1 (AQP1) is an intrinsic membrane protein complex that facilitates water flow, about 3×10^9 water molecules/s/channel. Aquaporins have been found in almost all living organisms, with at least 13 different forms in mammals and different water and ion permeability characteristics. Of these, 11 aquaporins have been characterized. AQP0, 1, 2, 4, 6, and 8 are selectively permeable to water; aquaglyceroporins (AQP3, 7, 9, and 10) are permeable to water and glycerol; andAQ 3, 7, and 9 also transport ammonia and urea.

Diabetes Insipidis

Recent experimental evidence suggests that water flow through aquaporin channels is regulated. For example, three mammalian aquaporins appear to be regulated by pH. Others are regulated by phosphorylation reactions or by the binding of specific signal molecules. For example, adipose tissue aquaglyceroporin expression is regulated by catecholamines and insulin. In 1993, the cause of a rare inherited form of **nephrogenic diabetes insipidis** (a disease in which the kidneys of affected individuals cannot produce concentrated urine) was discovered to be a mutation in the gene for AQP2. The mutated AQP2 does not respond to the antidiuretic hormone vasopressin (**Table 5.3**, p. 149).

Of all the aquaporins, AQP1, a homotetramer that has high permeability to water only, is the best characterized. Each subunit is a polypeptide containing 269 amino acid residues that form a water-transporting pore with six α-helical membrane-spanning domains connected by five loops. Although each monomer is an independent water channel, the formation of the tetramer is required for full function. In the functional monomer, the two loops that both possess an Asn-Pro-Ala (NPA) sequence meet in the middle to form the water-binding site. The pore, which has been measured at 3 Å, is only slightly larger than a water molecule (2.8 Å). The movement through the channel of water molecules only, and not smaller species such as H^+, is believed to be made possible by the formation of hydrogen bonds between water molecules and the Asn residues of the two NPA sequences (**Figure 11.31**). The hydrophobic environment created by the amino acid residues on the other helices that comprise the pore causes the hydrogen bonds between water molecules to break as they move in single file toward the narrowest portion of the pore. It also forces the oxygen atom of each water molecule to orient itself toward the Asn residues. When the water molecule approaches the 3 Å constriction in the pore, its oxygen atom sequentially forms and

breaks hydrogen bonds with the side chains of the two Asn residues. The absence of other hydrogen-bonding partners prevents the ionization of H_2O and the generation of protons. Peter Agre received the Nobel Prize in Chemistry in 2003 for his discovery and investigation of aquaporins.

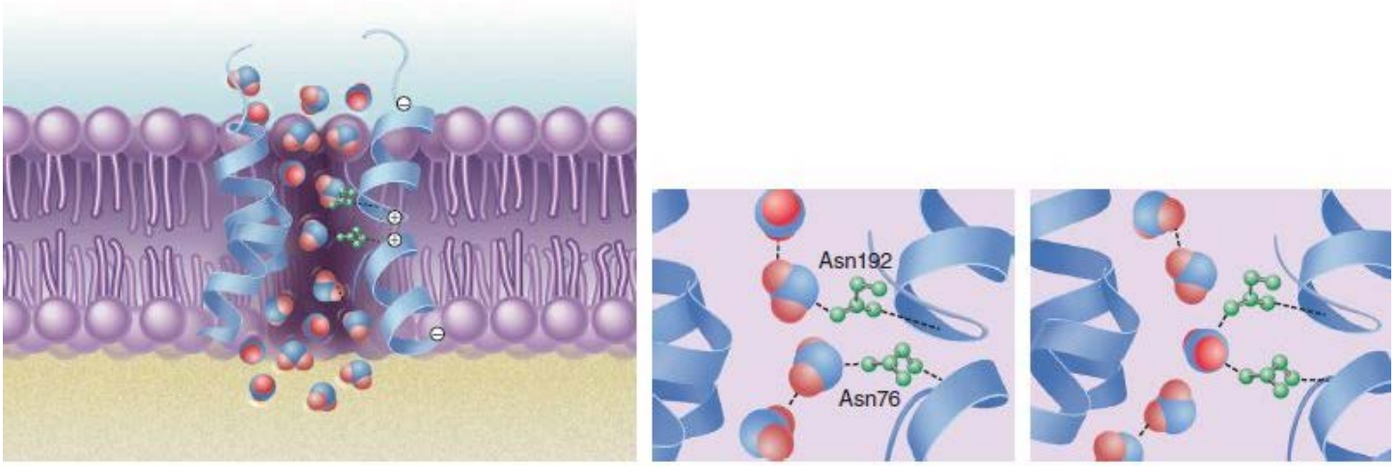


FIGURE 11.31

Water Transport through the AQP1 Monomer

Water molecules move through the pore in single file. As each molecule approaches the constriction in the pore, it is forced to orient its oxygen atom so that it can then form and break hydrogen bonds with the side chains of the two Asn residues. (a) Within the pore of the aquaporin monomer, there is a positive electrostatic environment in which the oxygen atom of each water molecule is oriented toward the two Asn residues. (b) and (c) The sequential formation and breakage of hydrogen bonds between the oxygen of water molecules and the two Asn residue side chains mediate the movement of water through the pore.

THE CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR Impaired membrane transport mechanisms can have serious consequences. One of the best-understood examples of dysfunctional transport occurs in cystic fibrosis. **Cystic fibrosis** (CF), a fatal autosomal recessive disease, is caused by a missing or defective plasma membrane glycoprotein called **cystic fibrosis transmembrane conductance regulator** (**CFTR**). CFTR (**Figure 11.32**), which functions as a chloride channel in epithelial cells, is a member of a family of proteins referred to as ABC transporters. (ABC transporters are so named because each contains a polypeptide segment called an ATP-binding cassette.) Unlike other ABC proteins, which are ATP-dependent transporters, CFTR is an ATP-gated chloride channel. CFTR is a large protein with 1480 amino acid residues that are organized into five domains. Two domains (MSD1 and MSD2), each containing six membrane-spanning helices, form the Cl^- channel pore. Chloride transport through the pore is controlled by the other three domains (all of which occur on the cytoplasmic side of the plasma membrane). Two are nucleotide-binding domains (NBD1 and NBD2) that bind and hydrolyze ATP and use the released energy to drive conformational changes in the pore. The regulatory (R) domain contains several amino acid residues that must be phosphorylated by cAMP-dependent protein kinase (PKA) for chloride transport to occur.

The chloride channel is vital for proper absorption of salt (NaCl) and water across the apical membrane surface of epithelial cells that line ducts and tubes in tissues such as lungs, liver, small intestine, and sweat glands. Chloride channel opening occurs in response to a signal molecule, cAMP. The cAMP-dependent kinase PKA then phosphorylates specific residues in the R domain, causing a change in its conformation that triggers the binding of ATP molecules to NBD1 and

NBD2. The two nucleotide-binding domains then form a head-to-tail heterodimer-like structure with the ATP-binding sites on the inside surfaces.

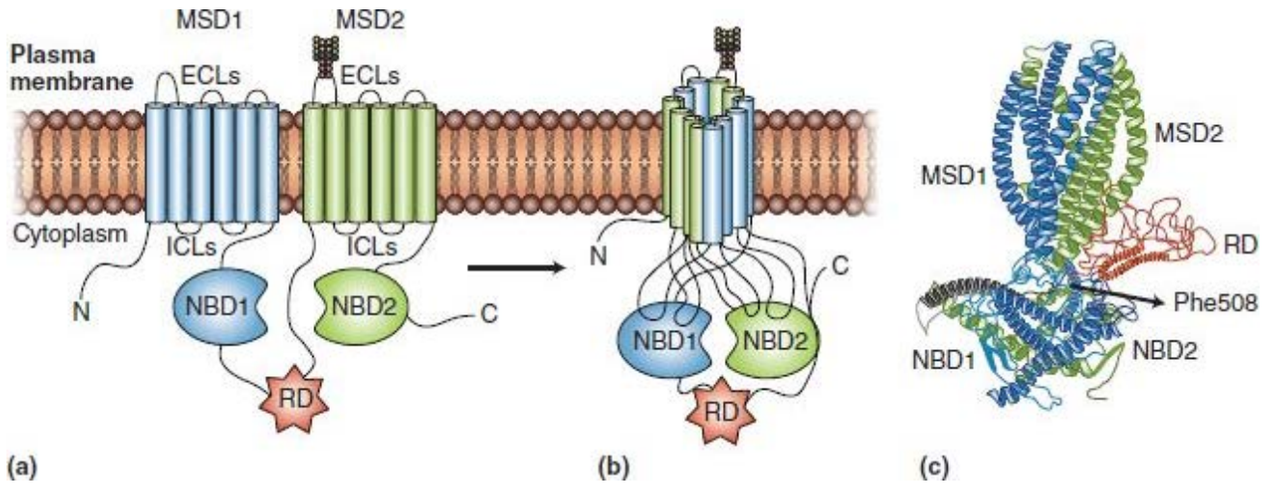


FIGURE 11.32

The Cystic Fibrosis Transmembrane Conductance Regulator (CFTR)

CFTR is a chloride channel composed of five domains: two transmembrane domains (MSD1 and MSD2), two nucleotide-binding domains (NBD1 and NBD2), and a regulatory domain (RD). The structure shown in (a) illustrates the domain sequence of the 1480 amino acid-containing polypeptides, including the extracellular loops (ECLs) and intracellular loops (ICLs) that link the transmembrane segments. The packing of the domains to form the functional protein is illustrated in (b). Note that activation of the channel requires dimerization of NBD1 and NBD2, which is triggered when ATP binds at the interface between these domains (not shown). A ribbon model of CFTR is illustrated in (c). Transport of the Cl^- or other small anions through the pore, driven by ATP hydrolysis, occurs when specific amino acid side chains in the regulatory domain are phosphorylated. Phosphorylation by PKA (a cAMP-dependent kinase) causes conformation changes in the intrinsically disordered R domain that are required for pore opening. The position of Phe508, the deletion of which is the most common CF mutation, is illustrated in (c).

As a result of these intramolecular rearrangements, the chloride channel gate opens and chloride ions flow down their concentration gradient. Hydrolysis of one of the NBD-bound ATP molecules causes dimer disruption that results in channel closing. The NBD dimer acts as a timing device in that the rate of ATP hydrolysis determines the length of time that the channel is open.

Cystic Fibrosis

In cystic fibrosis, the failure of CFTR channels results in the retention of $C1^-$ within the cells. Thick mucus or other secretion forms because osmotic pressure causes the excessive uptake of water from the mucus. The most obvious features of CF are lung disease (obstructed air flow and chronic bacterial infections) and pancreatic insufficiency (impaired production of digestive enzymes that can result in severe nutritional deficits). In the majority of CF patients, CFTR is defective because of a deletion mutation at Phe508, which causes protein misfolding, thereby preventing the processing and insertion of the mutant protein into the plasma membrane. Less common causes of CF (more than 100 have been described) include defective formation of CFTR mRNA molecules, mutations in the nucleotide-binding domains that result in ineffective binding or

hydrolysis of ATP, and mutations in the pore-forming domains that cause reduced chloride transport.

Before the development of modern therapies, CF patients rarely survived childhood. It is only because of antibiotics (used principally in the treatment of lung infections) and commercially available digestive enzymes (replacing the enzymes normally produced in the pancreas) that many CF patients can now expect to live into their thirties. Yet as with the sickle-cell gene (see p. 155), defective *CF* genes are not rare. With an approximate incidence of 1 in every 2500 Caucasians, CF is the most common fatal genetic disorder in this population. Recent experiments with "knockout" mice indicate that carriers of the mutant gene are protected from diseases that kill because of diarrhea. (Knockout animals are genetically modified so that a functional gene is replaced by a defective copy.) The experimental animals used in CF research lose significantly less body fluid because they have a reduced number of functional chloride channels. It is suspected that CF carriers (individuals having only one copy of a defective *CF* gene) are also less susceptible to fatal diarrhea (e.g., in cholera) for the same reason. The *CF* gene did not spread beyond western Europe (the incidence among East Asians is approximately 1 in 100,000) because CF carriers secrete slightly more salt in their sweat than noncarriers, and the epithelial cells that line sweat gland ducts cannot reabsorb chloride efficiently. In warmer climates, where sweating is a common feature of daily life, chronic excessive salt loss is far more dangerous than intermittent exposure to diarrhea-causing microorganisms.

In recent years, researchers have attempted to solve mutation-caused CFTR dysfunction with small molecules called correctors (molecules that improve folding in the ER and trafficking to the cell surface) and potentiators (molecules that enhance the function of plasma membrane-bound CFTR). Several such molecules have completed clinical trials. For example, the corrector molecule VX809 partially restores CFTR function in individuals with the Phe508 deletion mutation. VX809 acts as a molecular chaperone by stabilizing the folded MSD1 in a process that results in some CFTR molecules reaching the cell surface. VX770 (also called Kalydeco) is a potentiator that improves chloride transport in individuals with a G551D mutation (glycine is replaced by aspartate). The G551D mutation results in a CFTR protein that reaches the cell surface but cannot transport Cl^- through the ion channel.

QUESTION 11.8

Suggest the mechanism(s) by which each of the following substances is transported across cell membranes:

a. CO_2
b. Glucose
c. $C1^-$
d. K^+
e. Fat molecules
f. α-tocopherol

QUESTION 11.9

Describe the types of noncovalent interaction that promote the stability and functional properties of biological membranes.

QUESTION 11.10

Transport mechanisms are often categorized according to the number of transported solutes and the direction of solute transport.

1. *Uniporters* transport one solute.
2. *Symporters* transport two different solutes simultaneously in the same direction.
3. *Antiporters* transport two different solutes simultaneously in opposite directions.

After examining the examples of transport discussed in this chapter, determine the category to which each belongs.

MEMBRANE RECEPTORS Membrane receptors provide mechanisms by which cells monitor and respond to changes in their environment. In multicellular organisms, the binding to membrane receptors of chemical signals, such as the hormones and neurotransmitters of animals, is a vital link in intercellular communication. Other receptors are engaged in cell–cell recognition or adhesion. For example, lymphocytes perform a critical function of the immune system by transiently binding to the body's cells. When they bind to the cell surface of virus-infected cells, this binding event triggers the lymphocyte-induced death of the infected cell. Similarly, the capacity of cells to recognize and adhere to other appropriate cells in a tissue is of crucial importance in organismal processes, such as embryonic and fetal development.

The binding of a ligand to a membrane receptor results in a conformational change, which then causes a specific programmed response. Sometimes, receptor responses appear to be relatively straightforward. For example, the binding of acetylcholine to an acetylcholine receptor opens a cation channel. However, most responses are complex. The most intensively researched example of membrane receptor function is LDL receptor-mediated endocytosis, which is discussed next.

LDL RECEPTORS The low-density lipoprotein receptor is responsible for the uptake of cholesterol-containing lipoproteins into cells. The LDL receptor is a glycoprotein found on the surface of many cells, which binds to the ApoB-100 on the surface of LDL and ApoE found in chylomicrons and VLDL in peripheral blood. ApoE-containing lipoproteins formed in the central nervous system also bind to LDL receptors. When cells need cholesterol for the synthesis of membrane or steroid hormones, they produce LDL receptors and insert them into discrete regions of plasma membrane. (These membrane regions usually constitute about 2% of a cell's surface). The protein *clathrin* forms a lattice-like polymer on the cytosolic side of the membrane during the initial stages of endocytosis. The number of receptors per cell varies from 15,000 to 70,000, depending on cell type and cholesterol requirements.

The process of LDL receptor-mediated endocytosis occurs in several steps (**Figure 2.17**). It begins within several minutes after LDLs have bound to LDL receptors clustered in *coated* pits. Coated pits are concave regions of the membrane with a clathrin cage on the intracellular face that have large numbers of LDL receptor proteins. The LDL-occupied coated pit pinches off inside the cell to become a coated vesicle. Subsequently, *uncoated vesicles* are formed as clathrin depolymerizes. Before uncoated vesicles fuse with early endosomes (p. 52), LDLs are uncoupled from LDL receptors as the pH changes from 7 to 5. (This change is created by ATP-driven proton pumps in the vesicle membrane.) LDL receptors are recycled to the plasma membrane, and LDL-containing late endosomes fuse with lysosomes. Proteins associated with the LDL particles are degraded to amino acids, and cholesteryl esters are hydrolyzed to cholesterol and fatty acids.

Under normal circumstances, LDL receptor-mediated endocytosis is a highly regulated process. In liver cells, transcription factors called SREBPs (sterol regulatory element binding proteins; see p. 471) have been characterized. SREBP precursors are membrane-bound ER proteins. When liver cell cholesterol levels are low, the precursor proteins are transported to the Golgi complex, where they are cleaved to form the active transcription factors. The SREBPs then migrate into the nucleus and bind to sterol regulatory elements (SREs); afterward, they together activate up to 30 genes involved in lipid metabolism, including the gene for the LDL receptor. Cellular cholesterol levels then rise in response to a combination of LDL uptake and increased endogenous cholesterol synthesis. High-fat diets block LDL receptor synthesis because the accumulation of ingested cholesterol in ER membranes prevents SREBP processing reactions.

FAMILIAL HYPERCHOLESTEROLEMIA The LDL receptor and its role in the uptake of LDL into fibroblasts were discovered by Michael Brown and Joseph Goldstein during an investigation of an inherited disease, *familial hypercholesterolemia* (FH). The biochemical defects that cause FH were identified as mutations in the LDL receptor gene. Depending on the types of mutation, their consequences include failures of receptor synthesis, intracellular transport to the plasma membrane, LDL receptor–LDL binding, or internalization of the LDL receptor–LDL complex.

LDL Receptors and FH

Patients with FH have elevated levels of plasma cholesterol because they have missing or defective LDL receptors. Heterozygous individuals (also referred to as *heterozygotes*) inherit one defective LDL receptor gene. Consequently, they possess half the number of functional LDL receptors. With blood cholesterol values of 300 to 600 mg/100 mL, heterozygotes have heart attacks as early as the age of 40. They also develop disfiguring *xanthomas* (cholesterol deposits in the skin) in their 30s. With a population frequency of 1 in 500, heterozygous FH is one of the most common human genetic anomalies. In contrast, *homozygotes* (individuals who have inherited a defective LDL receptor gene from both parents) are rare (approximately 1 in 1 million). These patients have plasma cholesterol values of 650 to 1200 mg/100 mL. Both xanthomas and heart attacks occur during childhood or early adolescence. Death usually occurs before the age of 20. The genetic defects that cause FH prevent affected cells from obtaining sufficient cholesterol from LDLs.

Biochemistry IN PERSPECTIVE

Alzheimer's Disease and Apolipoprotein E4

What structural property of apolipoprotein E4 is responsible for making it a high-risk factor for Alzheimer's disease? **Alzheimer's disease** (AD), the most common form of dementia, is a progressive brain disease in which there are severe and irreversible cognitive

deficits and personality changes. Pathological features of AD include extracellular amyloid β (Aβ) deposits (senile plaques composed of aggregated protein fragments), intracellular neurofibrillary tangles (aggregates formed by tau, a neuronal microtubule-associated protein, as a result of hyperphosphorylation), and neuroinflammation, a pathological process caused by the release of inflammatory proteins by several types of brain cells.

There are two major forms of AD: early onset (before age 65) and late onset (after age 65). Early-onset AD (EOAD), which represents less than 5% of all cases, is an autosomal dominant disease sometimes referred to as familial AD, or FAD. In EOAD, there is a mutation in one of three genes: APP, PSEN1, or PSEN2. APP (amyloid precursor protein) is a transmembrane protein that in neurons is involved in processes such as cell growth and synapse formation. PSEN1 (presenilin 1) and PSEN2 (presenilin 2) are homologous proteins that function as the catalytic subunit of γ-secretase, the intramembrane enzyme complex that catalyzes the cleavage of APP to yield Aβ peptides of various lengths (from 36 to 43 amino acid residues). Aβ42, the most toxic Aβ peptide, forms both insoluble fibrillar plaques and soluble Aβ oligomers. The soluble oligomers are believed to trigger tau hyperphosphorylation, which in turn results in severe microtubule dysfunction and neuronal death.

Late-Onset AD and Apolipoprotein E4

Although the symptoms of late onset AD (LOAD) resemble those of EOAD, there is no specific gene dysfunction that directly causes the disease. Instead, AD has been linked to a combination of genetic, metabolic, and environmental risk factors. After aging itself, the most important risk factor is a specific variant of the gene for apolipoprotein E (ApoE). Within the brain, ApoE is synthesized primarily by astrocytes (star-shaped cells that support neuron function) and then secreted into the extracellular space. After its secretion, the ApoE protein binds to a lipid transporter protein in the astrocyte plasma membrane. The transporter protein then transfers phospholipids and cholesterol to ApoE to form HDL-like lipoproteins in a process called *lipidation*. The roles of ApoE lipoproteins include brain lipid homeostasis (transport of cholesterol and phospholipids between brain cells for repair and remodeling), cellular uptake of Aβ via endocytosis of ApoE lipoproteins into astrocytes, and inhibition of neuroinflammation by stimulating the release of anti-inflammation proteins primarily from microglia (immune defense cells).

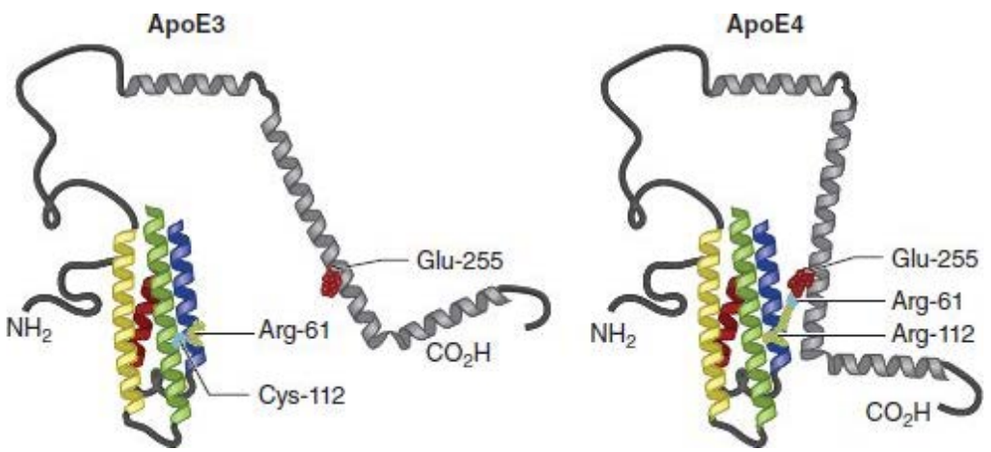


FIGURE 11A

Structures of ApoE3 and ApoE4

ApoE3 (a) is a fully functional protein in which the N-terminal domain binds to LDL receptors to initiate lipoprotein endocytosis and the C-terminal binds to lipid molecules. The domain interaction in ApoE4 (b), caused by the noncovalent bond between arg61 and glu255, alters the protein's structure so as to reduce its stability and functional properties. As a result, the brain produces fewer and smaller ApoE4

lipoproteins, thereby negatively affecting cholesterol homeostasis. Other consequences include increased amyloid plaque formation and tau hyperphosphorylation.

ApoE is a 35kDa protein with two independently folded domains that are separated by a hinge region. The N-terminal sequence (about two-thirds of the molecule) folds into a four-helix domain that binds to the low-density lipoprotein receptor (LDLR), the plasma membrane protein that mediates endocytosis; the hydrophobic C-terminal sequence folds into a lipid-binding domain. There are three human variants that differ at residues 112 and 158: ApoE2 (cys112, cys158), ApoE3 (cys112, arg158), and ApoE4 (arg112, arg158) (**Figure 11A**). Their approximate frequencies among humans and AD risk are

- ApoE2 (5–10%) (protective)
- ApoE3 (65–70%) (neutral risk)
- ApoE4 (15–20%) (high risk).

An individual with a genotype of APOE-ε4/-ε3 has an AD risk four times as high as an individual with the genotype APOE-ε3/-ε3. With a homozygous genotype (APOE-ε4/-ε4), the risk is about 12 times higher.

The high risk for carriers of the APOE-ε4 allele is linked to the protein's unique structure. In ApoE3, arg158 forms a salt bridge with asp154. In ApoE4, the substitution of cys112 with an arginine residue disrupts this salt bridge and causes a change in the relative position of arg61, enabling the latter residue to form a salt bridge with glu255 in the C-terminal domain. These structural changes result in the formation of a structure called a *domain interaction*. This domain interaction is believed to influence several properties of ApoE4 that contribute to LOAD risk. Examples include sensitivity to proteolysis within the ER (a process that yields lower levels of the secreted protein and generates neurotoxic C-terminal peptides, which then stimulate tau phosphorylation and promote mitochondrial dysfunction); lower lipoprotein Aβ binding capacity, leading to increased plaque formation; inadequate ApoE lipidation, resulting in impaired cholesterol transport; and reduced anti-inflammatory activity.

SUMMARY The substitution of a single amino acid in the primary structure of apolipoprotein E4 causes a profound change in its three-dimensional conformation. As a result, ApoE4-containing lipoproteins are substantially less able to perform their vital functions in the brain.

Chapter Summary

1. Lipids are a diverse group of biomolecules that dissolve in nonpolar solvents. They can be separated into the following classes: fatty acids and their derivatives, triacylglycerols, wax esters, phospholipids, lipoproteins, sphingolipids, and the isoprenoids.
2. Fatty acids are monocarboxylic acids that occur primarily in triacylglycerols, phospholipids, and sphingolipids. The eicosanoids are a group of powerful hormone-like molecules derived from long-chain fatty acids. The eicosanoids include the prostaglandins, the thromboxanes, and the leukotrienes.
3. Triacylglycerols are esters of glycerol with three fatty acid molecules. Triacylglycerols that are solid at room temperature (i.e., possess mostly saturated fatty acids) are called fats. Those that are liquid at room temperature (i.e., possess a high unsaturated fatty acid content) are referred to as oils. Triacylglycerols, the major storage and transport form of fatty acids, are an important energy storage form in animals. In

plants, they store energy in fruits and seeds.

4. Phospholipids are structural components of membranes. There are two types of phospholipid: phosphoglycerides and sphingomyelins.
5. Sphingolipids are also important components of animal and plant membranes. Like the sphingomyelins, they contain a ceramide base (*N*-acylsphingosine) but not phosphate. Their polar head group is one or more sugar residues.
6. Isoprenoids are molecules that contain repeating five-carbon isoprene units. The isoprenoids consist of the terpenes and the steroids.
7. Plasma lipoproteins transport lipid molecules through the bloodstream from one organ to another. They are classified according to their density. Chylomicrons are large lipoproteins of extremely low density that transport dietary triacylglycerols and cholesteryl esters from the intestine to adipose tissue and skeletal muscle. VLDLs, which are synthesized in the liver, transport lipids to tissues. As VLDLs unload some of their lipid molecules, they are converted to LDLs. LDLs bind to LDL receptors on the plasma membrane and then are engulfed by cells. HDLs, also produced in the liver, scavenge cholesterol from cell membranes and other lipoprotein particles. LDLs play an important role in the development of atherosclerosis.
8. According to the fluid mosaic model, the basic structure of membranes is a lipid bilayer in which proteins float. Membrane lipids (the majority of which are phospholipids) are primarily responsible for the fluidity, regional partitioning (lipid rafts), and sealing and fusion properties of membranes. Membrane proteins usually define the biological functions of specific membranes. Depending on their location, membrane proteins can be classified as integral or peripheral. Membranes are involved in transport and in the binding of hormones and other extracellular metabolic signals.
9. Membranes exhibit heterogeneity with respect to their lipid, protein, and carbohydrate components. New lipid bilayer is synthesized from the cytoplasmic face, and lipids are distributed throughout the bilayer by specific mediator proteins. The compartments on either side of the membrane are chemically different, and the membrane surfaces reflect that difference.
10. The movement of substances across a cell membrane can be accomplished by non-energy-requiring passive transport with the concentration gradient (the passive diffusion of small nonpolar molecules and the facilitated diffusion of large or polar molecules or ions via carrier or channel proteins), primary active transport (ATP energy used to concentrate a substance on one side of the membrane), secondary active transport (ion gradient generated by primary active transport used to concentrate another substance on one side of the membrane), or receptor-mediated transport (receptor and ligand in coated pits engulfed by endocytosis). Some transport channels are gated; that is, they open only in the presence of a certain gating substance (neurotransmitter, ions, etc.) or membrane condition (voltage, pH).

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on lipids and membranes to help you prepare for exams.

Chapter 11 Review Quiz

Suggested Readings

Aronson JK. 2016. Inhibiting the proton pump: mechanisms, benefits, harms, and questions. BMC Medicine 14:172–5.

Brown D. 2017. The discovery of water channels (aquaporins). Ann Nutr Metab 70(Suppl 1):37–42.

Cheng PX, Nichols BJ. 2016. Caveolae: one function or many? Trends Cell Biol 26(3):177–89.

Chin S, Hung M, Bear CE. 2017. Current insight into the role of PKA phosphorylation in CFTR channel activity and the pharmacological rescue of cystic fibrosis disease−causing mutants. Cell Mol Life Sci 74:57–66.

Laforenza U, Bottino C, Gastaldi G. 2016. Mammalian aquaglyceroporin function in metabolism. Biochim et Biophys Acta—Biomembranes 1858:1–11.

Mahley RW. 2016. Central nervous system lipoproteins. Atheroscler Thromb Vasc Biol 36(7):1305–15.

Saint-Criq V, Gray MA. 2017. Role of CFTR in epithelial physiology. Cell Mol Life Sci 74:93–115.

Underwood E. 2017. How ApoE4 endangers brains. Science 357:1224.

Key Words

active transport, *428*
acyl group, *406*
aquaporin, *430*
autocrine, *406*
autoimmune disease, *407*
carotenoid, *416*
chylomicron, *421*
cystic fibrosis, *431*
cystic fibrosis transmembrane conductance regulator, *431*
eicosanoid, *406*
essential fatty acid, *405*
facilitated diffusion, *428*
fluid mosaic model, *422*
glycolipid, *413*
GPI anchor, *411*
high-density lipoprotein, *422*
integral protein, *425*
intermediate-density lipoprotein, *421*
isoprenoid, *415*
leukotriene, *406*
lipid, *403*
lipid bilayer, *422*
low-density lipoprotein, *421*
mixed terpenoid, *417*
monounsaturated, *405*
nephrogenic diabetes insipidis, *430*
neutral fat, *408*
nonessential fatty acid, *405*
omega-3 fatty acid, *405*
omega-6 fatty acid, *405*
passive transport, *428*

peripheral protein, *425*
phosphoglyceride, *410*
phospholipid, *409*
polar head group, *409*
polyunsaturated, *405*
prenylation, *417*
prostaglandin, *406*
simple diffusion, *428*
sphingolipid, *413*
sphingomyelin, *410*
steroid, *417*
surface active agent, *409*
terpene, *415*
thromboxane, *406*
very-low-density lipoprotein, *421*
wax, *409*
wax ester, *409*

Review Questions

SECTION 11.1

Comprehension Questions

1. Define the following terms:
    a. lipid
    b. *cis*-isomer
    c. *trans*-isomer
    d. fatty acid
    e. anandamine
2. Define the following terms:
    a. monounsaturated fatty acid
    b. polyunsaturated fatty acid
    c. saturated fatty acid
    d. nonessential fatty acid
    e. essential fatty acid
3. Define the following terms:
    a. eicosanoid
    b. omega-6 fatty acid
    c. omega-3 fatty acid
    d. trans fatty acid
    e. acyl group

4. Define the following terms:
    a. autocrine
    b. triacylglycerol
    c. neutral fat
    d. fat
    e. oil
5. Define the following terms:
    a. partial hydrogenation
    b. saponification
    c. emulsifying agent
    d. wax ester
    e. wax
6. Define the following terms:
    a. phospholipids
    b. soap
    c. surface active agent
    d. phosphoglyceride
    e. sphingomyelin
7. Define the following terms:
    a. lecithin
    b. phosphatidylcholine
    c. cardiolipin
    d. GPI anchors
    e. respiratory distress syndrome
8. Define the following terms:
    a. surfactant
    b. membrane remodeling
    c. glycolipid
    d. ceramide
    e. cerebroside
9. Define the following terms:
    a. ganglioside
    b. sphingolipidoses
    c. isoprenoid
    d. terpene
    e. mixed terpenoid
10. Define the following terms:
    a. prenylation
    b. steroid
    c. cardiac glycoside
    d. digitalis
    e. lipoprotein

11. Define the following terms:
    a. apolipoprotein
    b. chylomicron
    c. VLDL
    d. LDL
    e. HDL

Fill in the Blanks

12. A derivative of a triterpene that contains four fused rings is referred to as a _____.
13. _____ is a type of cardiac glycoside that increases the force of cardiac contraction.
14. Bile salts emulsify dietary _____.
15. The protein components of lipoproteins are called _____.
16. In _____ signaling the signal molecule is active within the cell that produces it.

Short-Answer Questions

17. What role do plasma lipoproteins play in the human body?
18. Describe the possible consequences of a low-fat diet.
19. What do the abbreviations ACAT and LCAT stand for? What functions do these molecules serve in lipid metabolism?
20. What is the difference between an autocrine regulator and a hormone?
21. Describe the functions of the triacylglycerols.
22. Review the amphipathic structure of sphingomyelin. Which components are hydrophilic, and which are hydrophobic.
23. Explain why sdLDLs are more atherogenic than buoyant LDLs.
24. How do lipoproteins transport water-insoluble lipid molecules in the bloodstream?
25. Suggest a reason why trans fatty acids have melting points similar to analogous saturated fatty acids.

Critical-Thinking Questions

26. Suggest a reason why elevated LDL levels are a risk factor for coronary artery disease.
27. Plants often produce waxes on the surface of their leaves to prevent dehydration and protect against insects. What structural feature of waxes makes them suitable for this function?
28. Compare the energy contents of triacylglycerols and glycogen. Explain the differences between these two energy sources.
29. Describe the consequences of the substitution of sugar for fat in processed food.
30. Describe the function of dipalmitoylphosphatidylcholine in lung tissue.
31. Describe how a toxic phospholipase causes tissue damage.
32. What are the sphingolipid storage diseases? What is the process that causes these fatal diseases?
33. Describe the roles of several examples of terpenes.
34. What are endorphins, and what is their connection to dietary fat?

SECTION 11.2

Comprehension Questions

35. Define the following terms:
    a. lipid bilayer
    b. fluid mosaic model
    c. membrane fluidity
    d. flippase
    e. floppase
36. Define the following terms:
    a. heterokaryon
    b. aquaporin
    c. integral protein
    d. peripheral protein
    e. AE1
37. Define the following terms:
    a. chloride shift
    b. lipid raft
    c. caveolae
    d. simple diffusion
    e. facilitated diffusion
38. Define the following terms:
    a. depolarization
    b. repolarization
    c. passive transport
    d. active transport
    e. voltage-gated channel
39. Define the following terms:
    a. membrane carrier
    b. primary active transport
    c. secondary active transport
    d. nephrogenic diabetes insipidis
    e. cystic fibrosis
40. Define the following terms:
    a. CFTR
    b. ApoE4
    c. LDL receptor
    d. clathrin
    e. Na^+–K^+ ATPase
41. Define the following terms:
    a. coated pit
    b. familial hypercholesterolemia
    c. xanthoma
    d. homozygote

e. heterozygote

Fill in the Blanks

42. The term _____ describes the viscosity of a membrane.
43. In active transport, ATP provides the energy required to transport a substance _____ its concentration gradient.
44. _____ proteins are embedded within a membrane.
45. Transport of a substance across membranes that requires no direct input of energy is referred to as _____ transport.
46. Three proteins that facilitate the movement of phospholipids across a membrane are _____, _____, and _____ .
47. _____ is a form of membrane transport that requires a channel or carrier but no direct energy input.
48. _____ formation is initiated by AGE formation caused by high-glucose levels.
49. A mutation in the gene for _____ is the cause of nephrogenic diabetes insipidis.

Short-Answer Questions

50. Describe several factors that influence membrane fluidity.
51. Explain the differences in the ease of lateral movement and bilayer translocation movement of phospholipids.
52. How do most water molecules move through hydrophobic cell membranes?
53. Describe the structural and functional properties of lipid rafts.
54. Describe and give an example of a facilitated diffusion process.
55. Detergents are synthetic soap-like substances that biochemists use to disrupt cell membranes and extract membrane proteins. Explain how this process works.
56. Explain why triacylglycerols are not components of lipid bilayers.
57. What changes to the structure of a cell membrane will increase the cell's resistance to mechanical stress?
58. Describe how glucose is transported across membranes in the kidney. What type of transport is involved?
59. Explain how K^+ moves across a nerve cell membrane during depolarization and repolarization.

Critical-Thinking Questions

60. Glycolipids are nonionic lipids that can orient themselves unto bilayers as phospholipids do. They accomplish this feat without an ionic head group like that of the phospholipids. Suggest a reason why this is possible.
61. The myelin sheath insulates the axons of certain neurons in the body. What structural feature of this covering makes it a good insulator?
62. Changes in temperature affect membrane properties. How would you expect the lipid composition of thermophile membranes to differ from those of prokaryotes that live at temperate temperatures?
63. The fluid mosaic model of membrane structure has been very useful in explaining membrane behavior. However, the description of membrane as proteins floating in a phospholipid sea is oversimplified. Describe some components of membrane that are restricted in their lateral motion.

64. Explain why entropy increases when a lipid bilayer forms from phospholipid molecules.
65. Mammals in the Arctic Circle such as reindeer have higher levels of unsaturated fatty acids in the cell membranes in their legs than in other parts of their bodies. Suggest a reason for this phenomenon. Does it have a survival advantage?
66. Animal cells are enclosed in a cell membrane that according to the fluid mosaic model is held together largely by hydrophobic interactions. Considering the shear forces that membranes are subject to, why don't such membranes break every time an animal moves.

MCAT Study Questions

67. Saponification is a reaction in which

    a. triacylglycerols are incorporated into lipoproteins
    b. triacylglycerols bind to the serum protein albumin
    c. fatty acid salts react to form triacylglycerols
    d. triacylglycerols undergo a base-catalyzed ester hydrolysis

68. Which of the following statements concerning glycolipids is true?

    a. All glycolipids contain sphingosine.
    b. All glycolipids contain one sugar residue.
    c. All glycolipids contain phosphatidylglycerol.
    d. Some glycolipids contain ceramide, while others do not.

69. Which of the following statements concerning terpenes is true?

    a. Monoterpenes are composed of one isoprene unit.
    b. Xanthophylls are derivatives of carotenes.
    c. Vitamin E is a precursor of ubiquinone.
    d. Prenylation is a process in which hydrocarbon chains are attached to cholesterol.

70. Digitoxin, the molecule used to treat congestive heart failure, is what type of molecule?

    a. mixed terpenoid
    b. phosphlipid
    c. steroid glycoside
    d. glycolipid

71. Which of the following molecules is not capable of forming lipid bilayers?

    a. glycolipids
    b. wax esters
    c. phospholipids
    d. cerebrosides

CHAPTER 12 Lipid Metabolism

Endurance Exercise Carbohydrate (skeletal muscle and liver glycogen and blood glucose) is the preferred fuel of skeletal muscle for brief bursts of activity. However, carbohydrate storage in the body is limited, while fat reserves are not. Endurance training (aerobic exercise) occurring over weeks or months promotes physiological and metabolic adaptations (e.g., faster release of calorie-dense fatty acids from fat cells and transport to muscle cell mitochondria) that increase fatty acid oxidation and reduce glucose oxidation. In effect, fatty acid oxidation has a carbohydrate-sparing role that increases muscle ATP synthesis and prevents blood glucose depletion, thereby delaying fatigue.

OUTLINE

ABETALIPOPROTEINEMIA

12.1 FATTY ACIDS, TRIACYLGLYCEROLS, AND THE LIPOPROTEIN PATHWAYS

Dietary Fat: Digestion, Absorption, and Transport

Triacylglycerol Metabolism in Adipocytes

Fatty Acid Degradation

The Complete Oxidation of a Fatty Acid

Fatty Acid Oxidation: Double Bonds and Odd Chains
Fatty Acid Biosynthesis
Regulation of Fatty Acid Metabolism in Mammals
Lipoprotein Metabolism: The Endogenous Pathway

12.2 MEMBRANE LIPID METABOLISM
Phospholipid Metabolism

12.3 ISOPRENOID METABOLISM
Cholesterol Metabolism
The Cholesterol Biosynthetic Pathway and Drug Therapy

Biochemistry in Perspective

Atherosclerosis

Biochemistry in Perspective

Biotransformation

AVAILABLE ONLINE

Biochemistry in Perspective

Sphingolipid Metabolism

Abetalipoproteinemia

After months of chronic diarrhea and symptoms described as "failure to thrive," a 10-month-old boy was taken to a gastroenterologist (a physician specializing in disorders of the stomach and intestines) by his increasingly desperate parents. Although he had no feeding difficulties, the child's height and weight were lower than the fifth percentile. Laboratory exams were normal except for a fasting lipid profile, which revealed exceptionally low values for triacylglycerols (less than 10 mg/dl; normal is less than 150 mg/dl) and cholesterol (less than 50 mg/dl; normal is 170 mg/dl). In addition, the values for blood hemoglobin and hematocrit (proportion of blood volume occupied by red blood cells) were also low because of malabsorption of iron and folic acid. Visual examination of blood showed the presence of *acanthocytes*, abnormal red blood cells with spiny projections caused by abnormal lipid content in the red blood cell membrane. The use of an endoscope, an instrument fitted with a camera used to view the inside of the body, showed that much of the duodenum, the first part of the small intestine immediately beyond the stomach, was almost white in color and not a normal pink. Microscopic examination of retrieved enterocytes revealed that the cytoplasm contained large lipid droplets, which, combined with steatorrhea (excessive fats in the feces), indicated an inability to absorb dietary fat.

Diagnosis and Cause

This constellation of symptoms led the child's physician to suspect *abetalipoproteinemia*, an exceptionally rare autosomal recessive metabolic disorder (a frequency of 1 in 1 million humans). The disorder affects the absorption of dietary fat, cholesterol, and the fat-soluble vitamins. The diagnosis was confirmed by a blood test for lipoproteins, which showed that the patient's blood contained virtually no chylomicrons, VLDL, and LDL (see pp. 419–22). In recent years, the molecular defect that causes alipoproteinemia was finally identified. ApoB-containing lipoproteins fail to form as a result of mutations in the gene that codes for a molecular chaperone called microsomal triglyceride transfer protein (MTTP). MTTP facilitates the transfer of lipids onto ApoB during the production of ApoB containing lipoproteins.

Prognosis and Treatment

Clinical research over the past several decades has improved the lives of individuals with abetalipoproteinemia. Early diagnosis and treatment can now help patients manage their symptoms and delay or prevent serious long-term damage. Just a few examples of the damage observed in patients not diagnosed until adulthood include ataxia (progressive neurological damage that causes uncoordinated muscle movement), muscle weakness, and loss of peripheral vision. Treatment of affected children consists of several years of intravenous feeding with a lowfat formula supplemented with medium-chain triacylglycerols and fat-soluble vitamins. This is followed by a rigorous low-fat diet with oral doses of medium-chain fatty acids and fat-soluble vitamins. (Medium-chain fatty acids are more easily absorbed than are long-chain fatty acids, and they are transported by albumin in the bloodstream instead of chylomicrons.)

Overview

LIPIDS PLAY A UNIQUE ROLE IN LIVING ORGANISMS LARGELY BECAUSE OF THEIR HYDROPHOBIC STRUCTURES. LIPIDS SERVE AS (1) HIGHLY EFFICIENT AND compact energy storage molecules (triacylglycerols), (2) essential components of biological membranes (phospholipids, sphingolipids, and cholesterol), and (3) diverse molecules that have signaling (e.g., steroid hormones and prostaglandins) or protective (e.g., α-tocopherol) functions. Chapter 12 focuses on the metabolism of the major classes of lipids, that is, how they are synthesized and degraded and how these processes are regulated. Major emphasis is placed on the central metabolite in lipid metabolism: acetyl-CoA. The metabolism of cholesterol is also discussed because of its prominent role in cardiovascular disease.

The structural and functional diversity of lipids is impressive. All lipids are derived in whole or in part from acetyl-CoA (**Figure 9.9**). For example, acetyl-CoA is the substrate in the synthesis of fatty acids, the terpenes (e.g., β-carotene), and the steroids (e.g., cholesterol). When cells require energy, fatty acids are degraded to yield acetyl-CoA, which is then diverted into the citric acid cycle. Chapter 12 explores the metabolism of the major lipid classes: fatty acids, triacylglycerols, phospholipids, and isoprenoids. In addition, we review the synthesis of the ketone bodies and consider several metabolic control mechanisms.

12.1 FATTY ACIDS, TRIACYLGLYCEROLS, AND THE LIPOPROTEIN PATHWAYS

Triacylglycerols (fat molecules) are an important and efficient energy source in animals. For example, 30–40% of the calories the average American ingests come from fat. Triacylglycerol molecules (TGs) are digested within the lumen of the small intestine (**Figure 12.1**). The absorption of TGs and other lipid nutrients and their distribution to the body's tissues via lipoproteins is referred to as the *exogenous pathway*. (The *endogenous pathway* in which lipoproteins transport lipids produced in the liver to the body's cells is described on p. 472.)

Dietary Fat: Digestion, Absorption, and Transport

This section describes the digestion and absorption of dietary fat, TG metabolism in adipose tissue, energy-yielding fatty acid degradation reactions, and fatty acid biosynthesis. The section ends with

an overview of fatty acid metabolism regulation and a short review of the endogenous lipoprotein pathway, the mechanism whereby the liver packages lipids into lipoproteins for distribution throughout the body.

After dietary fat mixes with **bile salts**, amphipathic molecules with detergent properties (see p. 445) are digested by pancreatic lipase to form fatty acids and monoacylglycerol. These latter molecules are then transported across the plasma membrane of intestinal wall cells (enterocytes). Short-chain (C4 to C6) and medium-chain (C6 to C12) fatty acids are transferred to the bloodstream, where they bind to serum albumin, which carries them to the liver. Long-chain fatty acids are delivered to the enterocyte smooth endoplasmic reticulum (SER), where they are incorporated into TGs. Enterocytes combine triacylglycerols with dietary cholesterol, newly synthesized phospholipids, and ApoB-48 to form *nascent* (newly made) chylomicrons (large, low-density lipoproteins; see p. 421). (ApoB-48, the main lipoprotein component of nascent chylomicrons, is synthesized from an mRNA that is a truncated version of the ApoB-100 mRNA, the protein found on lipoprotein produced in the liver.) After their secretion into the lymph (tissue fluid derived from blood), chylomicrons pass from the lymph into the bloodstream at the thoracic duct. Nascent chylomicrons are converted into mature chylomicrons as they circulate in blood and lymph when HDLs transfer two lipoprotein molecules. ApoC-II activates lipoprotein lipase (LPL), and ApoE binds to a specific receptor on the surface of hepatocytes.

Most of the triacylglycerol content of circulating chylomicrons is removed from blood by muscle and the adipose tissue cells (adipocytes), which comprise the body's primary lipid storage depot. Lipoprotein lipase, synthesized by cardiac and skeletal muscle, lactating mammary gland, and adipose tissue, is transferred as a homodimer to the lumenal surface of capillary endothelial cells. Once it is activated by ApoC-II, LPL converts the triacylglycerol in chylomicrons into fatty acids and glycerol. As many as 40 LPL dimers have been observed to bind simultaneously to a lipoprotein. Insulin and glucagon activate LPL in adipose tissue and in skeletal and cardiac muscle, respectively. LPL deficiency causes elevated blood levels of TGs (hypertriglyceridemia).

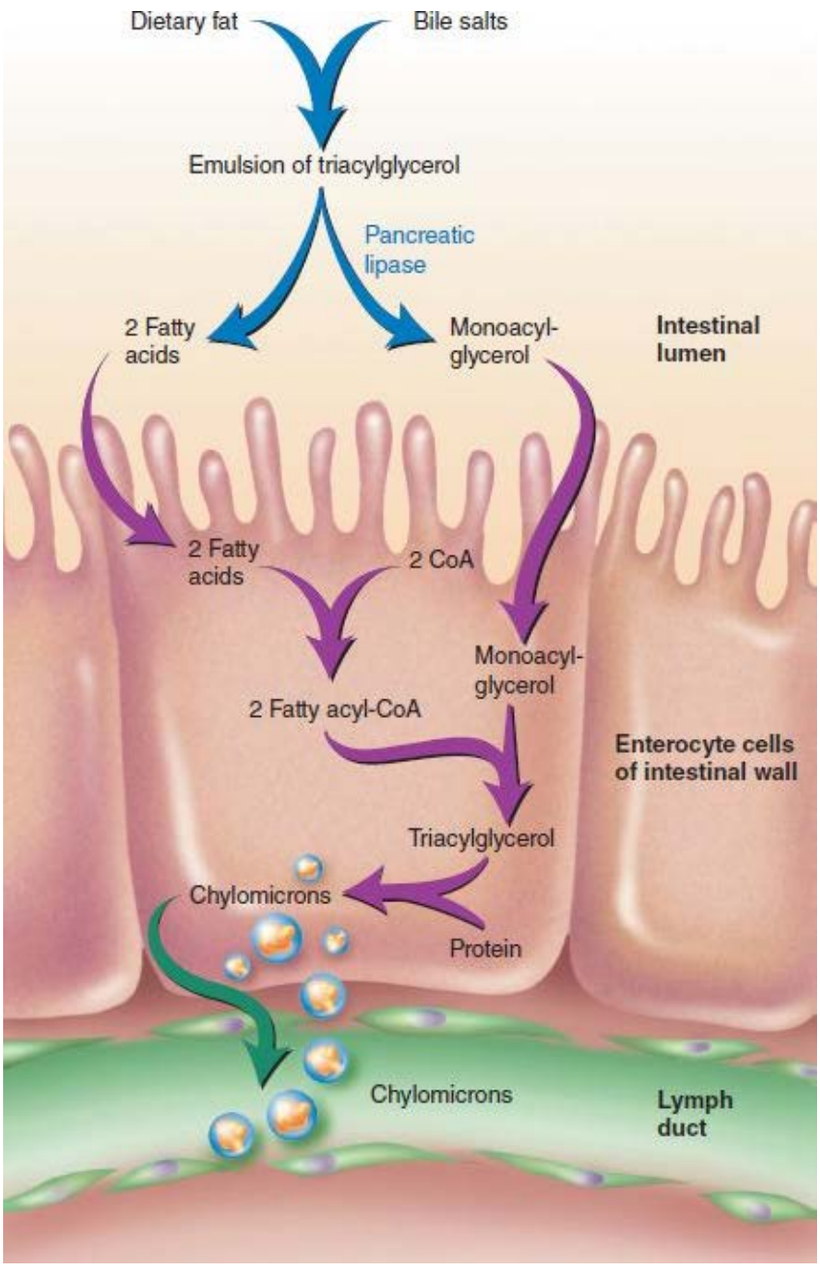


FIGURE 12.1

Digestion and Absorption of Triacylglycerols in the Small Intestine

After triacylglycerols have been emulsified (solubilized) by mixing with bile salts, they are digested by intestinal lipases, the most important of which is pancreatic lipase. The products, fatty acids and monoacylglycerol, are transported into enterocytes and resynthesized to form triacylglycerol. The triacylglycerol molecules, along with newly synthesized phospholipid and protein, are then incorporated into chylomicrons. After the chylomicrons have been transported into lymph, via exocytosis, and then into blood, the triacylglycerols are drawn off by muscle and fat cells. Chylomicron remnants are removed from the blood by the liver.

The fatty acid products of LPL-catalyzed hydrolysis are taken up by cells, whereas glycerol is carried in the blood to the liver, where the enzyme glycerol kinase converts it to glycerol-3-phosphate. Glycerol-3-phosphate is then used in the synthesis of triacylglycerols, phospholipids, or glucose. When LPL has removed about 90% of TGs in chylomicrons, the **chylomicron remnants** are removed from the blood by liver cells via the binding of ApoE to chylomicron remnant receptors. Hydrolysis of the remaining TGs within lysosomes releases fatty acids and glycerol that can be either metabolized by liver cells immediately or stored for later use. Cholesterol molecules released from chylomicron remnants have several fates. Some are esterified with fatty acids and then packaged into nascent lipoproteins, whereas others are either converted into bile acids or secreted directly into bile.

Triacylglycerol Metabolism in Adipocytes

Fatty acids, stored in TGs primarily in adipocytes, are the body's most concentrated energy source. Depending on an animal's current metabolic needs, fatty acids may be released from triacylglycerols to be degraded to generate energy or used in membrane synthesis. Immediately after a meal, for example, insulin is released in response to high blood glucose levels. Insulin promotes triacylglycerol storage by inactivating *hormone-sensitive lipase* (one of the enzymes that hydrolyzes the ester bonds of fat molecules in adipose tissue) and activating triacylglycerol synthesis in adipocytes and muscle cells. Insulin also stimulates the release of VLDL from the liver and activates LPL synthesis and translocation of the enzyme to the surface of the endothelial cells serving adipose and muscle tissue. As a result, the uptake of fatty acids and storage as triacylglycerols is increased. When blood glucose falls after a meal, insulin levels decrease and glucagon levels increase, promoting net TG hydrolysis in fat and muscle cells. Triacylglycerols are degraded to form glycerol and fatty acids.

THE TG CYCLE AND GLYCERONEOGENESIS The **triacylglycerol cycle** (**Figure 12.2**) is a mechanism that regulates the level of fatty acids that are available to the body for energy generation and synthesis of molecules such as phospholipids. TGs are constantly both synthesized and hydrolyzed to fatty acids and glycerol. This seemingly futile cycle occurs at the cellular level (e.g., in adipocytes) and at the whole body level. **Figure 12.2** illustrates TG cycling between adipocytes and the liver. TGs are hydrolyzed in adipocytes with the release of a relatively small fraction of fatty acids into blood. Once in the blood, fatty acids are transported to the body's other tissues. In the liver, a high proportion of the fatty acids are reincorporated into TGs, most of which are packaged into VLDL. The net result of TG cycling is that a flexible system ensures that sufficient fatty acids are available for the body's energy and biosynthetic requirements. Excess fatty acids, which can have toxic effects on cells, are efficiently reesterified into TGs. The pathways by which TGs are synthesized and hydrolyzed are described next.

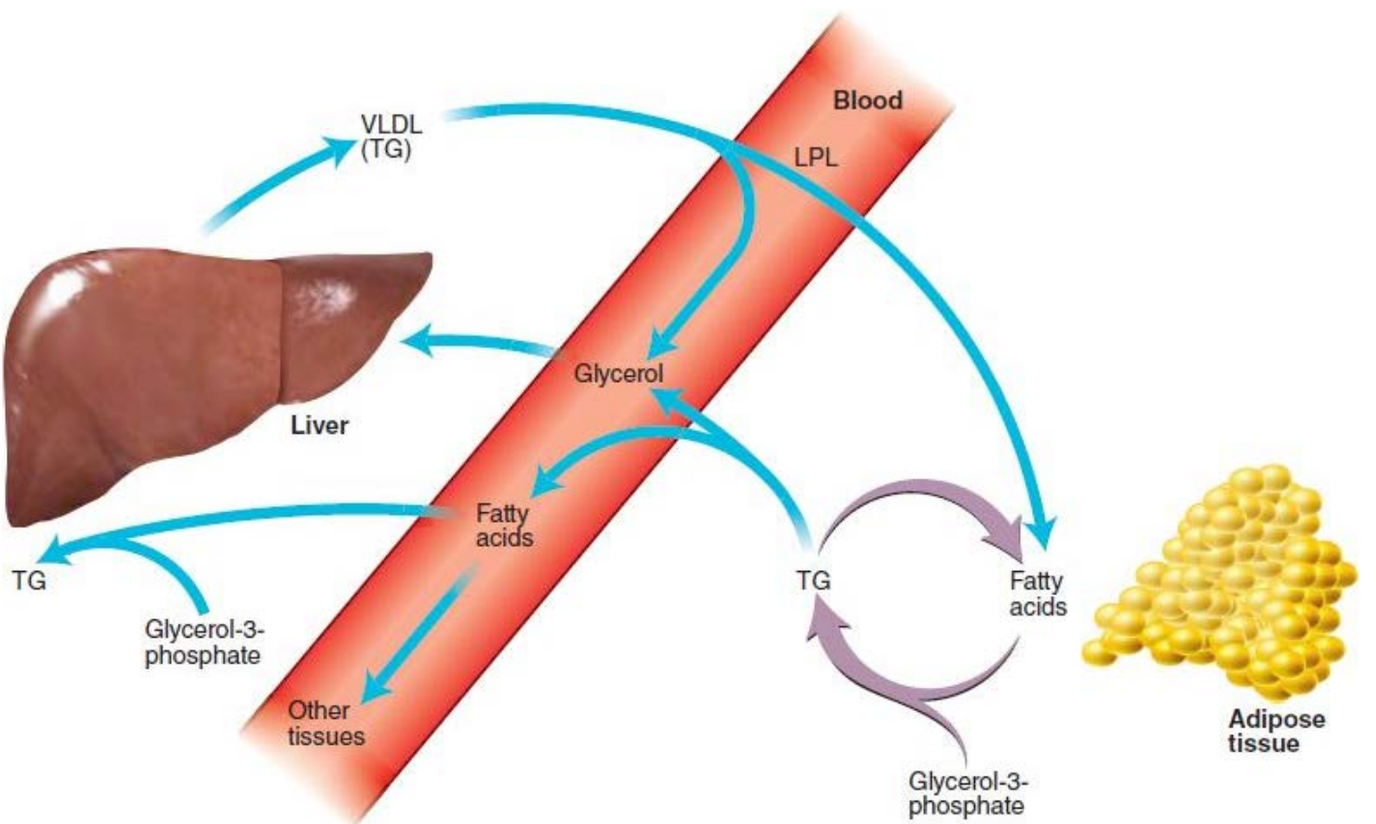


FIGURE 12.2

The Triacylglycerol Cycle

In adipocytes, the body's primary energy storage depot, TGs are synthesized from fatty acids obtained from the blood and glycerol-3-phosphate. The rate at which fatty acids are released into blood to meet current energy needs of other tissues is increased by glucagon and epinephrine and depressed by insulin. Under all metabolic conditions, however, the percentage of adipocyte fatty acids (about 75%) that are reesterified is remarkably constant. In the liver most fatty acids removed from blood are used to synthesize TGs that are incorporated into VLDL. Once VLDLs are secreted into blood, they travel to tissues such as adipose tissue where TGs are hydrolyzed by lipoprotein lipase. Fatty acids are then transported into adipocytes. Glycerol, the other product of TG hydrolysis, is removed from blood by the liver.

TRIACYLGLYCEROL BIOSYNTHESIS Triacylglycerol synthesis (referred to as **lipogenesis**) is illustrated in **Figure 12.3**. Glycerol-3-phosphate or DHAP reacts sequentially with three molecules of acyl-CoA (fatty acid esters of CoASH). Acyl-CoA molecules are produced in the following reaction:

(1)

Note that the reaction is driven to completion by the subsequent hydrolysis of pyrophosphate by pyrophosphatase.

In the synthesis of triacylglycerols, phosphatidic acid is formed by two sequential acylations of glycerol-3-phosphate or by a pathway involving the direct acylation of DHAP. In the latter pathway, acyldihydroxyacetone phosphate is later reduced to form lysophosphatidic acid. Depending on the pathway used, lysophosphatidic acid synthesis utilizes either an NADH or an NADPH cofactor. Phosphatidic acid is produced when lysophosphatidic acid reacts with a second acyl-CoA. Once formed, phosphatidic acid is converted to diacylglycerol by phosphatidic acid phosphatase. A third acylation reaction forms triacylglycerol. Fatty acids derived from both the diet and de novo synthesis are incorporated into triacylglycerols. (The term *de novo* is used by biochemists to indicate new synthesis.) De novo synthesis of fatty acids is discussed on p. 460. Glyceroneogenesis, the principal means of producing glycerol-3-phosphate required in TG synthesis, is described next.

Glyceroneogenesis (**Figure 12.4**) is an abbreviated version of gluconeogenesis in which glycerol-3-phosphate (required for TG synthesis) is synthesized from substrates other than glucose or glycerol. Examples include pyruvate, alanine, glutamine, and citric acid cycle intermediates. The key enzymes for glyceroneogenesis are pyruvate carboxylase (PC) and the cytoplasmic isoform of phosphoenolpyruvate carboxykinase (PEPCK-C). Both enzymes are found in large amounts in lipogenic (TG-producing) tissues, such as adipose tissue and lactating mammary glands, and organs involved in gluconeogenesis (i.e., liver and kidney). Note that glyceroneogenesis is more active in brown adipocytes (p. 384) than in white adipocytes. PC and PEPCK-C are also found in moderate amounts in brain, heart, and adrenal glands.

TRIACYLGLYCEROL HYDROLYSIS When energy reserves are low, the body's fat stores are mobilized in a process referred to as **lipolysis** (**Figure 12.5**). Lipolysis in adipose tissue occurs during fasting or vigorous exercise and in response to stress. Several hormones (e.g., the catecholamines epinephrine and norepinephrine) bind to specific adipocyte plasma membrane receptors and begin a reaction sequence similar to the activation of glycogen phosphorylase. Hormone binding to the receptor elevates cytoplasmic cAMP levels, which, in turn, activates the phosphorylation of a protein called perilipin that is linked to another protein referred to as CGI-58. Perilipin-1 then releases CGI-58. Once perilipin-1 is phosphorylated, its conformational change exposes TGs to lipase-catalyzed hydrolysis. The first and committed step in TG hydrolysis is

catalyzed by adipose triglyceride lipase (ATGL), now bound to its coactivator, CGI-58. The products are diacylglycerol and a fatty acid. The hydrolysis of diacylglycerol to yield monoacylglycerol and a fatty acid is catalyzed by hormone-sensitive lipase (HSL). The final reaction, the conversion of monoacylglycerol to glycerol and a fatty acid, is catalyzed by monoacylglycerol lipase (MGL). The products of lipolysis (i.e., fatty acids and glycerol) are released into the blood. The synthesis of ATGL is promoted by the transcription factor PPARγ (p. 471) and suppressed by insulin. Insulin also suppresses lipolysis by decreasing intracellular cAMP.

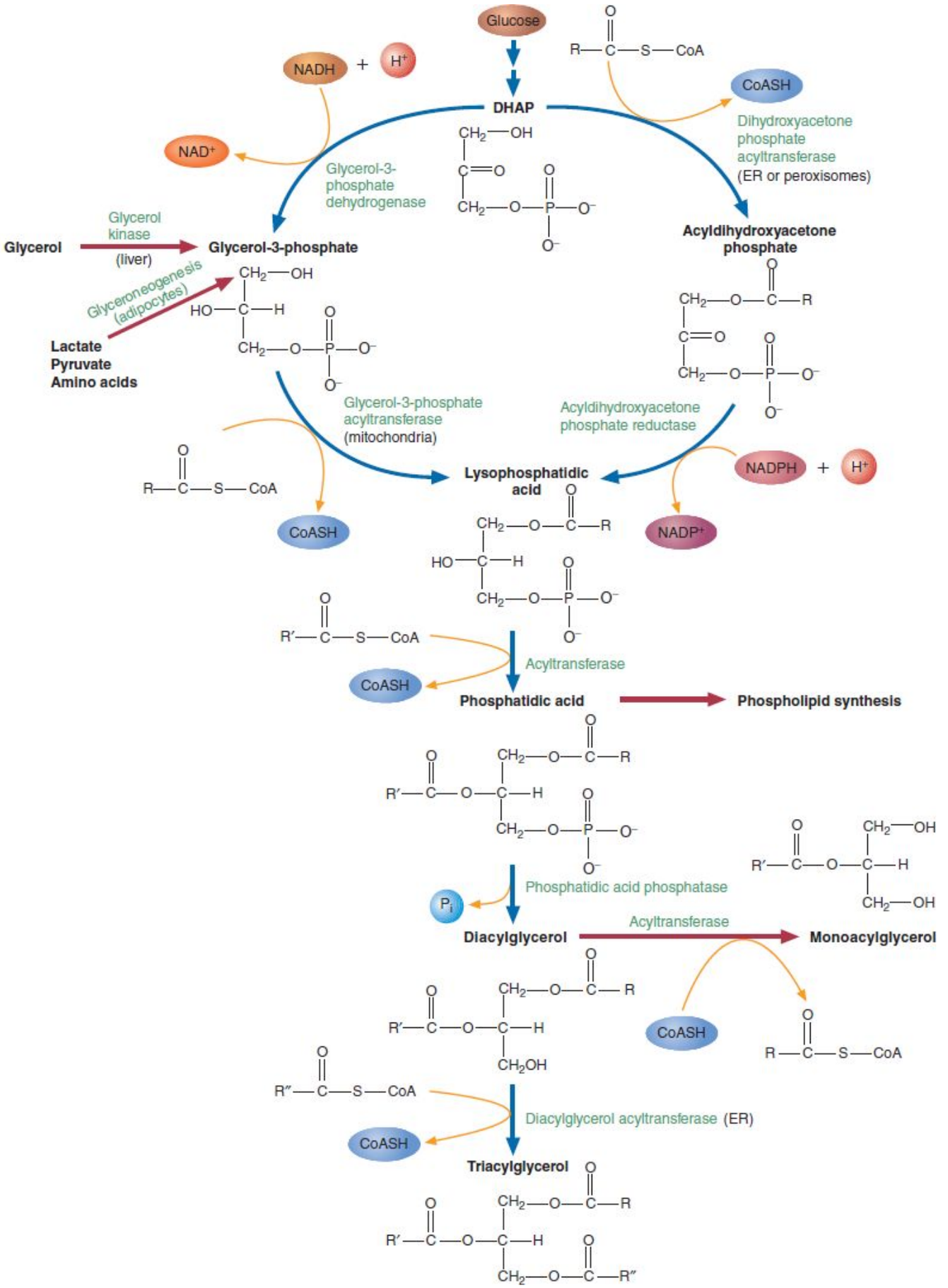


FIGURE 12.3

Triacylglycerol Synthesis

Most triacylglycerols are synthesized in liver and stored in adipose tissue. Glycerol-3-phosphate is required for de

novo synthesis of triacylglycerols in the liver and reassembly of triacylglycerols (storage) in adipose tissue. A significant percentage of the glycerol-3-phosphate required in TG synthesis in liver and adipose tissue is supplied by glyceroneogenesis (see p. 448) using lactate, pyruvate, and glucogenic amino acids such as alanine. The condensation product of glycerol-3-phosphate and two acyl-CoAs is phosphatidic acid, which is used in phospholipid synthesis. In human triacylglycerols, palmitate is often attached at C-1 and oleate at C-2.

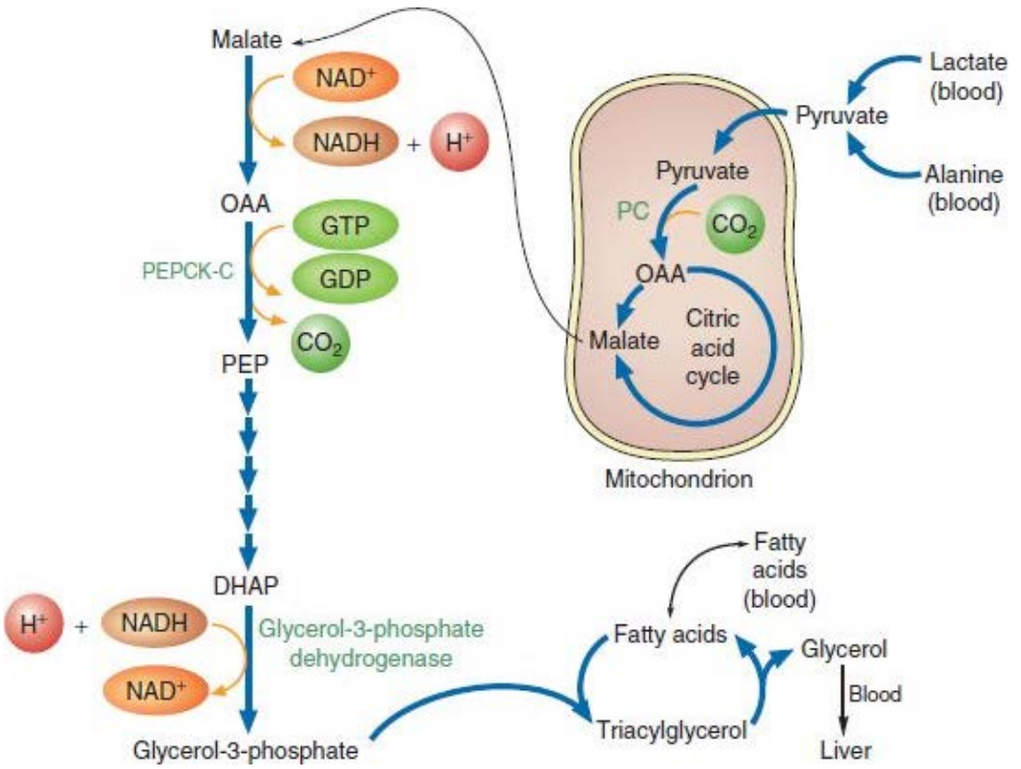


FIGURE 12.4

Glyceroneogenesis in Adipocytes

Glyceroneogenesis, an abbreviated form of gluconeogenesis, is a major source of glycerol-3-phosphate required for TG synthesis. Substrates for this pathway include lactate, pyruvate, and glucogenic amino acids such as alanine or glutamine. Pyruvate is converted to OAA within the mitochondrion by PC (pyruvate carboxylase). After OAA is reduced by NADH, the product, malate, is transported out of the mitochondrion where the reaction is reversed to form OAA. OAA is then phosphorylated and decarboxylated by phosphoenolpyruvate carboxykinase (PEPCK-C) in a GTP-requiring reaction to form phosphoenolpyruvate (PEP). PEP is then converted via gluconeogenesis to DHAP. DHAP is reduced by glycerol-3-phosphate dehydrogenase to glycerol-3-phosphate, which is then utilized in TG synthesis.

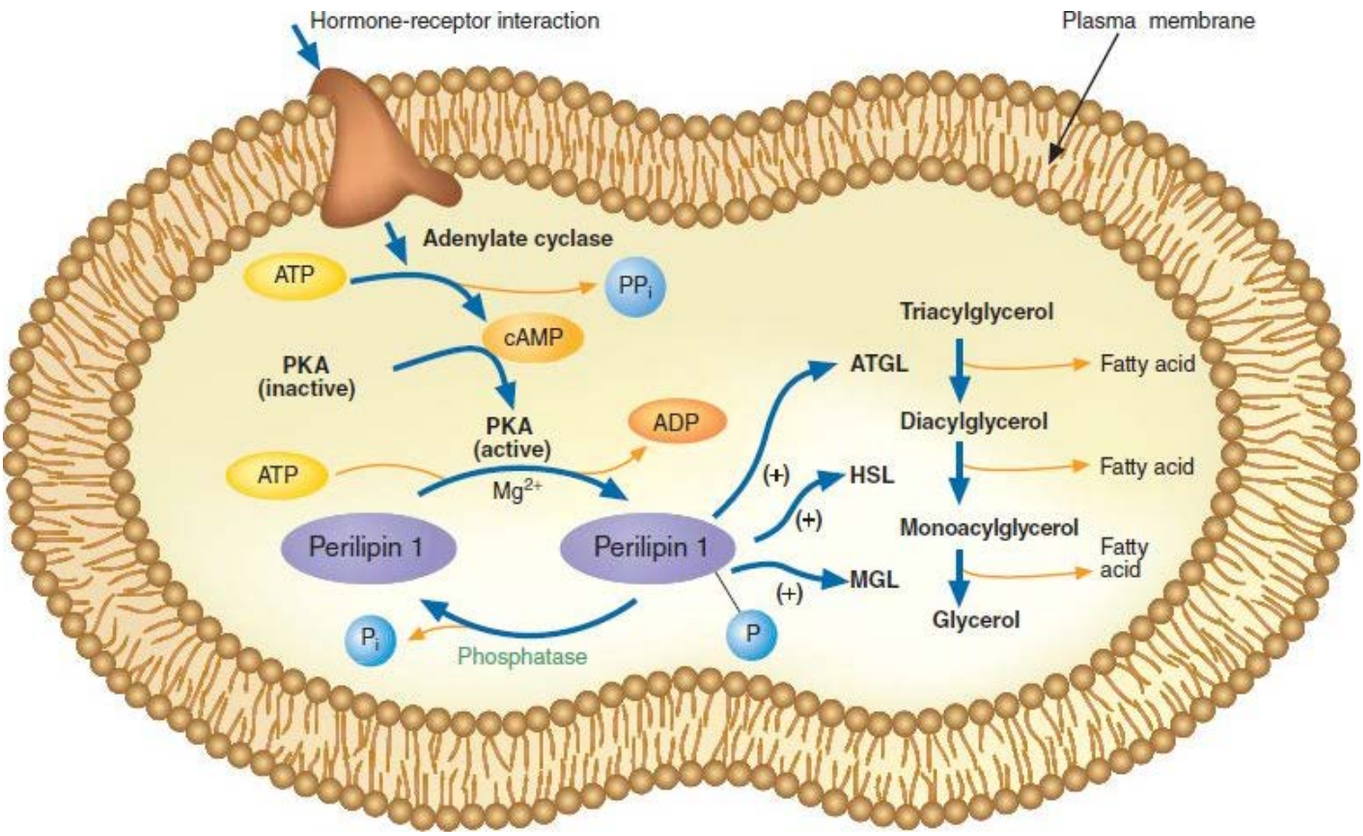


FIGURE 12.5

Diagrammatic View of Lipolysis in Adipocytes

TGs are sequentially hydrolyzed by three lipolytic enzymes, each with its own substrate preferences. In adipocytes, the process begins with the binding of a hormone molecule such as one of the catecholamines (p. 551) or glucagon to its cognate cell-surface receptor. Hormone-receptor binding initiates a cAMP-mediated mechanism that activates protein kinase A (PKA, p. 607). Activated PKA phosphorylates perilipin-1, a protein on the surface of the cell's lipid droplet that is associated with a protein called CGI-58 (not shown), and hormone-sensitive lipase (HSL), allowing the latter enzyme to move to the droplet surface. Phosphorylated perilipin-1 releases CGI-58, which then binds to and activates adipose triglyceride lipase (ATGL). ATGL then hydrolyzes TGs to yield diacylglycerol and a fatty acid. HSL subsequently hydrolyzes diacylglycerol to yield monoacylglycerol and a fatty acid. The final lipase, monoacylglycerol lipase (MGL), catalyzes the conversion of monoacylglycerol to glycerol and a fatty acid. Insulin depresses lipolysis (not shown) by promoting the phosphorylation of protein kinase B (PKB, pp. 612–13), which then phosphorylates a phosphatase that converts cAMP to its inactive derivative 5'AMP.

After their transport across the adipocyte plasma membrane, fatty acids become bound to serum albumin. The albumin-bound fatty acids are carried to tissues throughout the body, where they are released from albumin and taken up by cells across the plasma membrane. The amount of fatty acid that is transported depends on its concentration in blood and the relative activity of the fatty acid transport mechanism. Cells vary widely in their capacity to transport and use fatty acids. Some cells (e.g., brain and red blood cells) cannot use fatty acids as fuel, although others (e.g., cardiac muscle) rely on them for a significant portion of their energy requirements. Once they enter a cell, fatty acids must be transported to their destinations (mitochondria, ER, and other organelles). Several **fatty acid–binding proteins** (water-soluble proteins whose sole function is to bind and transport hydrophobic fatty acids) are responsible for this transport.

Most fatty acids are degraded to form acetyl-CoA within mitochondria in a process referred to as *β*-oxidation. *β*-Oxidation also occurs in peroxisomes. Other oxidative mechanisms are also available to degrade certain nonstandard fatty acids.

Fatty acids are synthesized (pp. 460–67) when an organism has met its energy needs and its nutrient levels are high. (Glucose and several amino acids are substrates for fatty acid synthesis.)

Fatty acids are synthesized from acetyl-CoA in a process that is similar to the reverse of β-oxidation. Although most fatty acids are supplied in the diet, most animal tissues can synthesize some saturated and unsaturated fatty acids. In addition, animals can elongate and desaturate dietary fatty acids. For example, arachidonic acid is produced by adding a two-carbon unit and introducing two double bonds to linoleic acid.

KEY CONCEPTS

- In the exogenous pathway, triacylglycerols and other lipid nutrients are absorbed into the body and distributed to the tissues by chylomicrons.
- When energy reserves are high, triacylglycerols are stored in the process called lipogenesis.
- When energy reserves are low, triacylglycerols are degraded to form fatty acids and glycerol. This process is called lipolysis.
- Triacylglycerols are constantly being synthesized and hydrolyzed to yield fatty acids and glycerol. The recycling rate is stimulated by epinephrine and norepinephrine and depressed by insulin.

QUESTION 12.1

You have just consumed a cheeseburger. Trace the fat molecules (triacylglycerol) from the cheeseburger to your adipocytes (fat cells).

Fatty Acid Degradation

Most fatty acids are degraded by the sequential removal of two-carbon fragments from the carboxyl end. During this process, referred to as ***β*-oxidation**, the β-carbon (second carbon from the carboxyl group) is oxidized, and acetyl-CoA is released as the bond between the α- and β-carbons is cleaved. This process is repeated until the entire fatty acid chain has been processed. Other mechanisms for degrading fatty acids are known. Branched-chain molecules usually require an α-oxidation step (pp. 459–60) in which the fatty acid chain is shortened by one carbon by means of a stepwise oxidative decarboxylation. In some organisms, the carbon farthest from the carboxyl group may be oxidized by a process called ω-oxidation, which generates short-chain dicarboxylic acids. In ω-oxidation, the terminal methyl group is converted to an alcohol by an O_2- and NADPH-requiring ER enzyme called *cytochrome* P_{450} (p. 485). The alcohol is subsequently converted to a carboxylate group by two sequential reactions catalyzed by ADH and aldehyde dehydrogenase. The resulting dicarboxylic acids are then shortened by β-oxidation in mitochondria to short-chain water-soluble dicarboxylic acids such as succinate and adipic acid (p. 460). In humans, ω-oxidation is a minor pathway that becomes relevant only when β-oxidation is impaired. The degradation of odd-chain, branched-chain, and unsaturated fatty acids is described later in the chapter, following a discussion of β-oxidation.

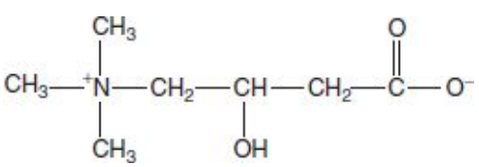
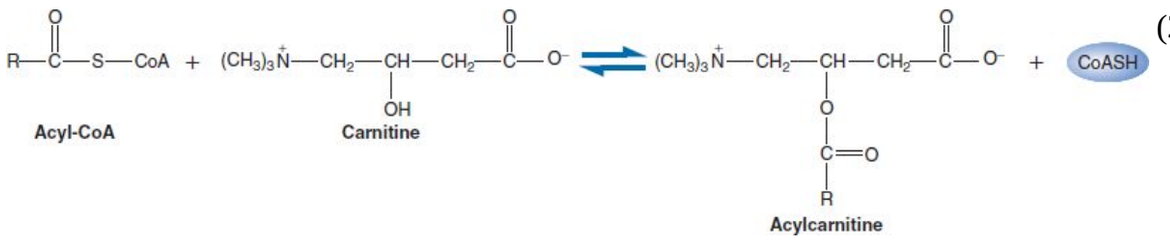


FIGURE 12.6
Structure of Carnitine

3D animation of carnitine

β-OXIDATION *β*-Oxidation of long chain fatty acids (14 to 20 carbon atoms) and medium chain fatty acids (6 to 12 carbon atoms) occurs within mitochondria. Before *β*-oxidation begins, each fatty acid is activated in a reaction with ATP and CoASH (see p. 448). The enzyme that catalyzes this reaction, acyl-CoA synthetase, is found in the outer mitochondrial membrane. Because the mitochondrial inner membrane is impermeable to most acyl-CoA molecules, a special carrier called *carnitine* is used to transport acyl groups into the mitochondrion (**Figure 12.6**). Carnitine-mediated transfer of acyl groups into the mitochondrial matrix is accomplished through the following mechanism (**Figure 12.7**):

1. Each acyl-CoA molecule is converted to an acylcarnitine ester derivative:



(2)

This reaction is catalyzed by carnitine acyltransferase I (CAT-I).

2. A carrier protein within the mitochondrial inner membrane transfers acylcarnitine into the mitochondrial matrix.
3. Carnitine acyltransferase II (CAT-II) reverses the ester-forming reaction to yield acyl-CoA and carnitine.
4. Carnitine is transported back into the intermembrane space by the carrier protein. It then reacts with another acyl-CoA.

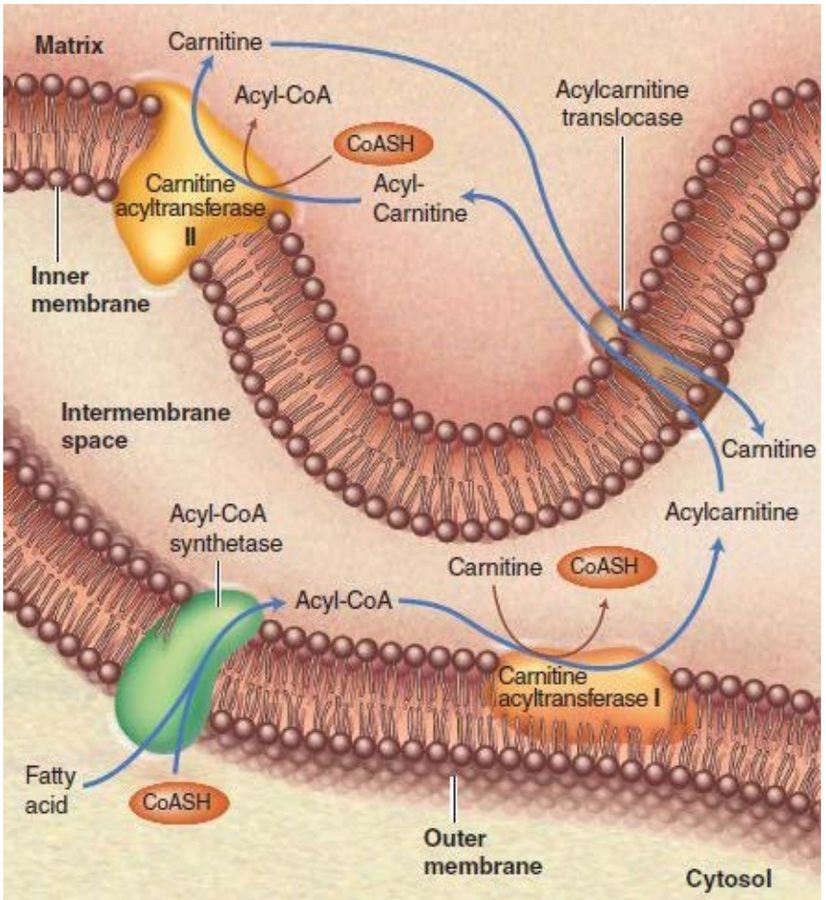


FIGURE 12.7

Fatty Acid Transport into the Mitochondrion

Fatty acids are activated to form acyl-CoA by acyl-CoA synthetase, an enzyme in the outer mitochondrial membrane. Acyl-CoA then reacts with carnitine to form an acylcarnitine derivative. Carnitine acyltransferase I catalyzes this reaction. Acylcarnitine is transported across the inner membrane by acylcarnitine translocase and then reconverted to carnitine and acyl-CoA by carnitine acyltransferase II. Note that acylcarnitine translocase is an antiport protein (i.e., it transports one acylcarnitine into the mitochondrial matrix for every carnitine that is transported into the intermembrane space).

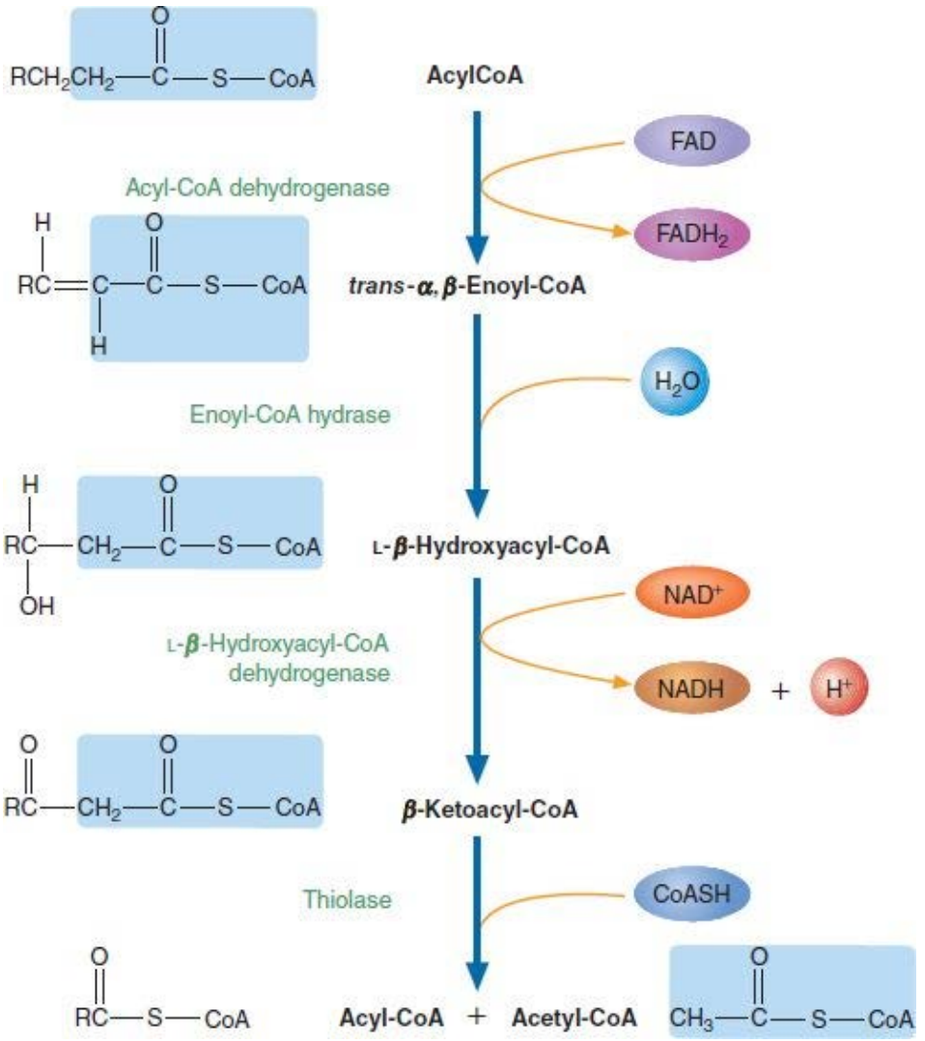
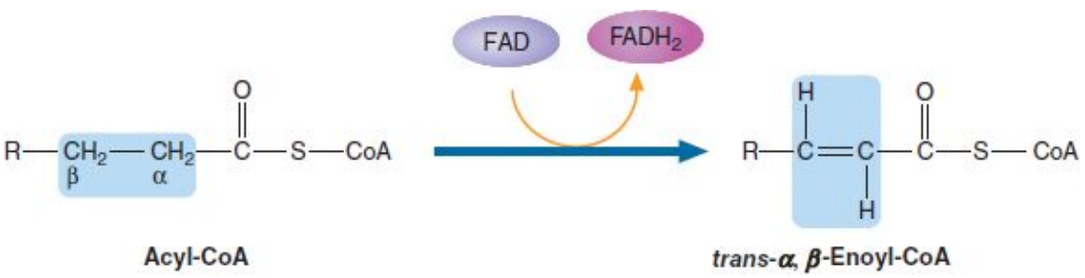


FIGURE 12.8

***β*-Oxidation of Acyl-CoA**

The *β*-oxidation of acyl-CoA molecules consists of four reactions that occur in the mitochondrial matrix. Each cycle of reactions forms acetyl-CoA and an acyl-CoA that is shorter by two carbons.

A summary of the reactions of the *β*-oxidation of saturated fatty acids is shown in **Figure 12.8**. The pathway begins with an oxidation-reduction reaction, catalyzed by acyl-CoA dehydrogenase (a flavoprotein associated with the matrix side of the inner membrane), in which one hydrogen atom each is removed from the *α*- and *β*-carbons and transferred to the enzyme-bound FAD:



(3)

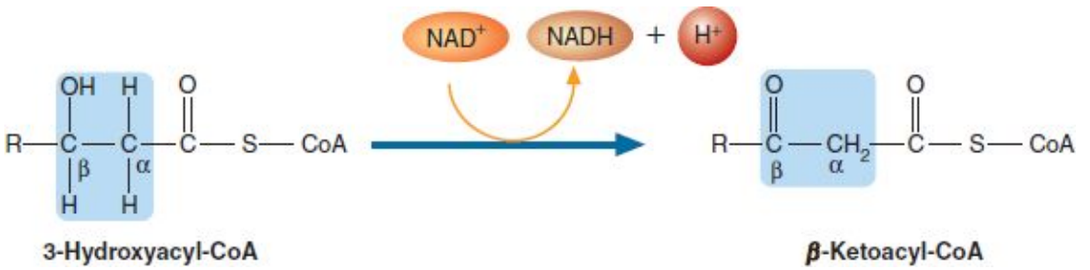
The $FADH_2$ produced in this reaction then donates two electrons to ubiquinone (UQ) in the mitochondrial electron transport chain (ETC; see **Figure 10.6**). There are several isozymes of acyl-CoA dehydrogenase, each specific to a different fatty acid chain length. The product of this reaction is *trans-α,β*-enoyl-CoA.

The second reaction, catalyzed by enoyl-CoA hydrase, involves a hydration of the double bond between the α- and β-carbons:

(4)

The β-carbon is now hydroxylated.

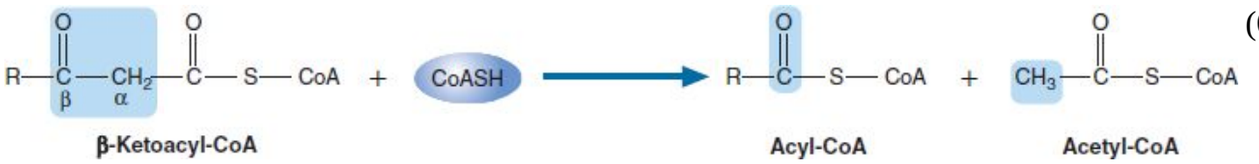
In the next reaction, catalyzed by β-hydroxyacyl-CoA dehydrogenase, this newly formed hydroxyl group is oxidized to yield a β-ketoacyl-CoA:



(5)

The electrons transferred to NAD^+ are later donated to complex I of the ETC.

Finally, thiolase (sometimes referred to as β-ketoacyl-CoA thiolase) catalyzes a C_α-C_β cleavage:

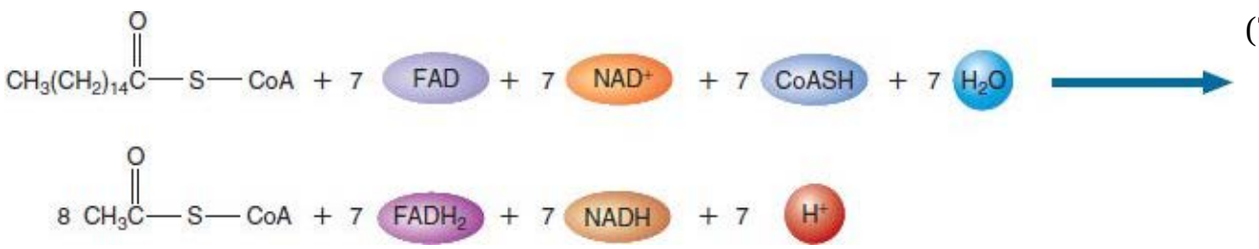


(6)

In this reaction, sometimes called a **thiolytic cleavage**, the enzyme thiolase in the presence of CoASH converts β-ketoacyl-CoA to an acetyl-CoA molecule and an acyl-CoA with two fewer C atoms.

The four steps just outlined constitute one cycle of β-oxidation. During each later cycle, a two-carbon fragment is removed. In a process called the *β-oxidation spiral*, the β-oxidation cycle is repeated until, in the last cycle, a four-carbon acyl-CoA is cleaved to form two molecules of acetyl-CoA.

The following equation summarizes the oxidation of palmitoyl-CoA:



(7)

In muscle, the rate of β-oxidation depends on the availability of its substrate (i.e., the concentration of fatty acids in blood) and the tissue's current energy requirements. When $NADH/NAD^+$ ratios are high, β-hydroxyacyl–CoA dehydrogenase is inhibited. High acetyl-CoA levels depress the activity of thiolase. In liver, where fatty acids are also used in the synthesis of triacylglycerols and

phospholipids, the rate of β-oxidation depends on how quickly these molecules are transported into mitochondria. When glucose levels are high and excess glucose molecules are being converted into fatty acids, malonyl-CoA, the product of the first committed step in fatty acid synthesis, prevents a futile cycle by inhibiting CAT-I. Regulation of β-oxidation by the enzyme AMP-activated protein kinase (AMPK) and select hormones and transcription factors is described on pp. 470–71.

The acetyl-CoA molecules produced by fatty acid oxidation are converted via the *citric acid cycle* to CO_2 and H_2O as additional NADH and $FADH_2$ are formed. A portion of the energy released as NADH and $FADH_2$ is oxidized by the ETC and is later captured in ATP synthesis via *oxidative phosphorylation*. The complete oxidation of acetyl-CoA is discussed in Chapter 10. The calculation of the total number of ATP that can be generated from palmitoyl is reviewed next.

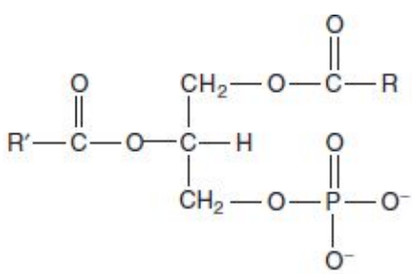
MCAD Deficiency

QUESTION 12.2

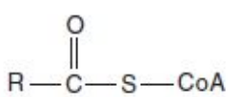
Medium-chain acyl-CoA dehydrogenase (MCAD) is a mitochondrial enzyme that catalyzes the first reaction in the β-oxidation cycle. Its substrates are 12-carbon acyl-CoAs. MCAD deficiency is an autosomal recessive disorder caused by a mutated version of the *MCAD* gene. Symptoms of this disorder, which include fatigue, vomiting, and hypoglycemia, are triggered by fasting. Apparently, the accumulation of toxic amounts of acyl-CoAs depresses gluconeogenesis. Considering that liver glycogen is depleted quickly in children, suggest a treatment that would help MCAD-deficient patients manage their symptoms.

QUESTION 12.3

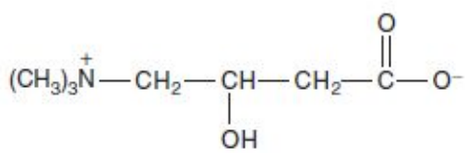
Identify each of the following biomolecules:



(a)



(b)



(c)

QUESTION 12.4

In the absence of oxygen, cells can produce small amounts of ATP from the anaerobic oxidation of glucose. This is not true for fatty acid oxidation. Explain.

The Complete Oxidation of a Fatty Acid

The aerobic oxidation of a fatty acid generates a large number of ATP molecules. As previously described (see p. 378), the oxidation of each FADH during electron transport and oxidative

phosphorylation yields approximately 1.5 molecules of ATP. Similarly, the oxidation of each NADH yields approximately 2.5 molecules of ATP. The yield of ATP from the oxidation of palmitoyl-CoA, which generates 7 $FADH_2$, 7 NADH, and 8 acetyl-CoA molecules to form CO_2 and H_2O, is calculated as follows:

$$\begin{aligned} 7\ FADH_2 \times 1.5\ ATP/FADH_2 &= 10.5\ ATP \\ 7\ NADH \times 2.5\ ATP/NADH &= 17.5\ ATP \\ 8\ Acetyl\text{-}CoA \times 10\ ATP/acetyl\text{-}CoA &= \underline{80\ \ ATP} \\ &\quad 108\ \ ATP \end{aligned}$$

The formation of palmitoyl-CoA from palmitic acid uses two ATP equivalents. The net synthesis of ATP per molecule of palmitoyl-CoA is therefore 106 molecules of ATP.

The yield of ATP from the oxidations of palmitic acid and glucose can be compared. Recall that the total number of ATP molecules produced per glucose molecule is approximately 30. If glucose and palmitic acid molecules are compared in terms of the number of ATP molecules produced per carbon atom, palmitic acid is a superior energy source. The ratio for glucose is 30/6 or 5 ATP molecules per carbon atom. Palmitic acid yields 106/16 or 6.6 ATP molecules per carbon atom. The oxidation of palmitic acid generates more energy than that of glucose because palmitic acid is a more reduced molecule. (Glucose with its six oxygen atoms is a partially oxidized molecule.)

QUESTION 12.5

Determine the number of moles of NADH, $FADH_2$, and ATP molecules that can be synthesized from 1 mol of stearic acid.

β-OXIDATION IN PEROXISOMES *β*-Oxidation of fatty acids also occurs within peroxisomes. In animals, peroxisomal *β*-oxidation shortens very-long-chain fatty acids (22 or more carbon atoms) without ATP synthesis, yielding medium-chain fatty acids. Peroxisomal membrane possesses an acyl-CoA synthetase activity that is specific for very-long-chain fatty acids. Mitochondria apparently cannot activate very long-chain fatty acids such as tetracosanoic (24:0) and hexacosanoic (26:0) acids. Peroxisomal carnitine acyltransferases catalyze the transfer of these molecules into peroxisomes, where they are oxidized to form acetyl-CoA and medium-chain acyl-CoA molecules that are further degraded via *β*-oxidation within mitochondria.

Although the reactions of peroxisomal *β*-oxidation are similar to those in mitochondria, there are some notable differences. First, the initial reaction in the peroxisomal pathway is catalyzed by an acyl-CoA oxidase. The reduced coenzyme $FADH_2$ then donates its electrons directly to O_2 instead of UQ. The H_2O_2 produced when $FADH_2$ is oxidized is converted to H_2O by catalase. Second, the next two reactions in peroxisomal *β*-oxidation are catalyzed by two enzyme activities (enoyl-CoA hydrase and 3-hydroxyacyl CoA dehydrogenase) found on the same protein molecule. Finally, the last enzyme in the pathway (*β*-ketoacyl-CoA thiolase) and its mitochondrial version have different substrate specificities. The former does not efficiently bind medium-chain acyl-CoAs.

THE KETONE BODIES Most of the acetyl-CoA produced during fatty acid oxidation is used by the citric acid cycle or in isoprenoid synthesis (Section 12.3). Under normal conditions, fatty acid metabolism is so carefully regulated that only small amounts of excess acetyl-CoA are produced. In a process called **ketogenesis**, excess acetyl-CoA molecules are converted to acetoacetate, *β*-hydroxybutyrate, and acetone. Together, these molecules are called the **ketone bodies** (**Figure 12.9**).

Ketosis

Ketone bodies form within the matrix of liver mitochondria. The process begins when two acetyl-CoAs condense to form acetoacetyl-CoA. Acetoacetyl-CoA then condenses with another acetyl-CoA to form β-hydroxy-β-methylglutaryl-CoA (HMG-CoA). In the next reaction, HMG-CoA is cleaved to form acetoacetate and acetyl-CoA. Acetoacetate is then reduced to form β-hydroxybutyrate. Acetone is formed by the spontaneous decarboxylation of acetoacetate when the latter molecule's concentration is high. This condition, referred to as **ketosis**, occurs during starvation and in uncontrolled diabetes. Since two ketone bodies contain carboxylate groups, an excess of ketone bodies in the body is also referred to as ketoacidosis. (Diabetes mellitus, a metabolic disease, is discussed in Chapter 16: see the Biochemistry in Perspective essay entitled Diabetes Mellitus.) In both starvation and diabetes, there is a heavy reliance on fat stores and β-oxidation of fatty acids to supply energy.

KEY CONCEPTS

- In β-oxidation, fatty acids are degraded by breaking the bond between the α- and β-carbon atoms.
- The ketone bodies are produced in liver mitochondria from excess molecules of acetyl-CoA.

Several tissues, most notably cardiac and skeletal muscle, use ketone bodies to generate energy. During prolonged starvation (i.e., in the absence of sufficient glucose), the brain uses ketone bodies as an energy source, thereby reducing its dependence on glucose. The oxidation of ketone bodies also spares skeletal muscle protein, a source of substrates for gluconeogenesis (glucose–alanine cycle, p. 306). Other cells that use ketone bodies to generate energy during starvation include enterocytes and adipocytes. The mechanism by which acetoacetate and β-hydroxybutyrate are converted to acetyl-CoA is illustrated in **Figure 12.10**.

A *ketogenic* (ketone body-producing) *diet*, high in fat with low amounts of protein and carbohydrate, has been successfully used in the treatment of refractory (difficult-to-control) childhood epilepsy. The cause of epilepsy, a group of neurological disorders that are characterized by epileptic seizures (disturbances in the brain's electrical activity), is unknown. Although glucose is the brain's principal fuel, it can partially adapt to using ketone bodies to generate energy. Although the ketogenic diet often results in significant reductions in seizures, it must be carefully monitored because of possible adverse effects, including low-grade ketoacidosis, high cholesterol levels, slowed growth, bone fractures, and kidney stones. The mechanism whereby the ketogenic diet reduces seizures is currently unknown.

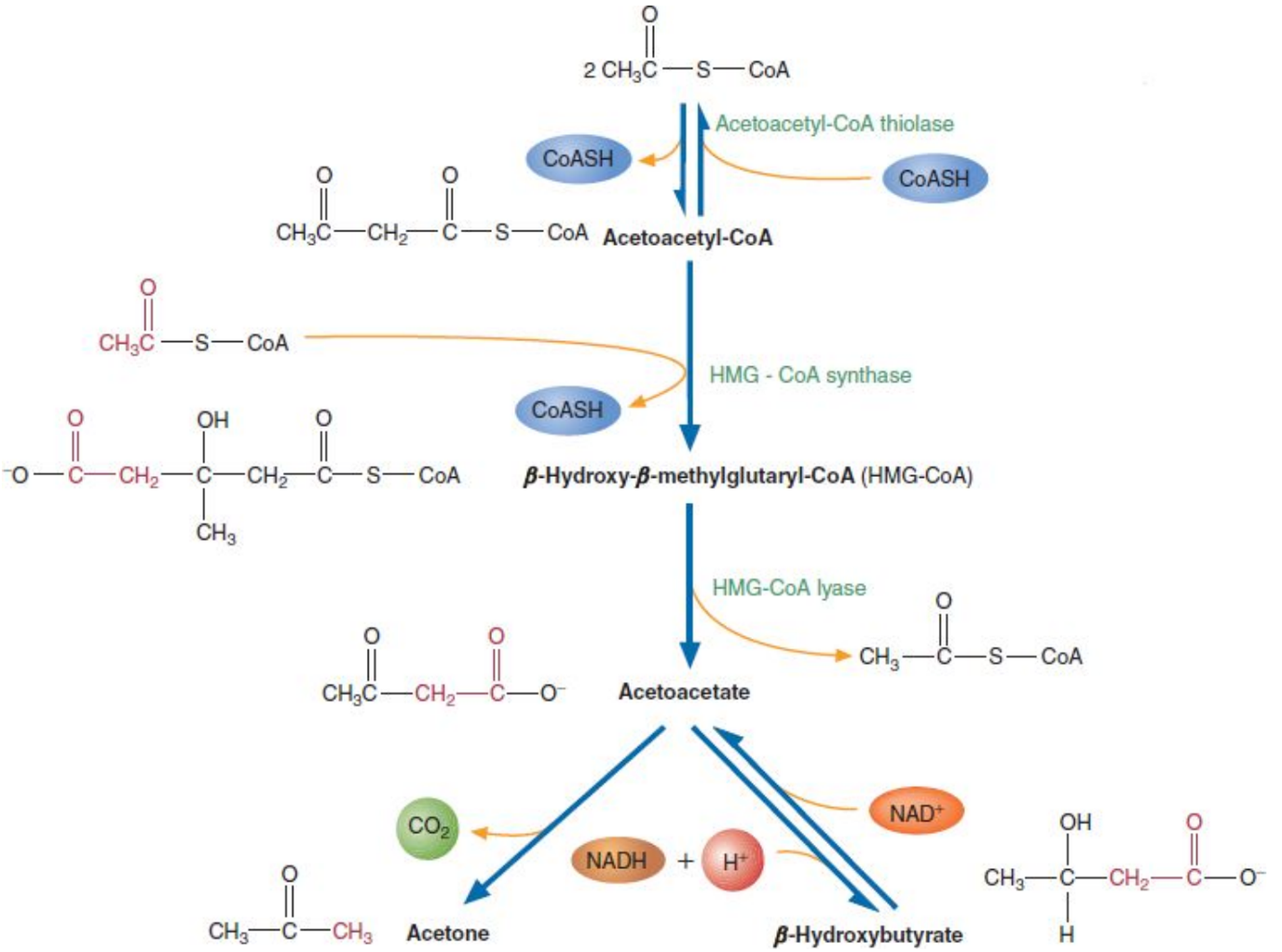


FIGURE 12.9

Ketone Body Formation

Ketone bodies (acetoacetate, *β*-hydroxybutyrate, and acetone) are produced within liver mitochondria when excess acetyl-CoA is available. Under normal circumstances, only small amounts of ketone bodies are produced.

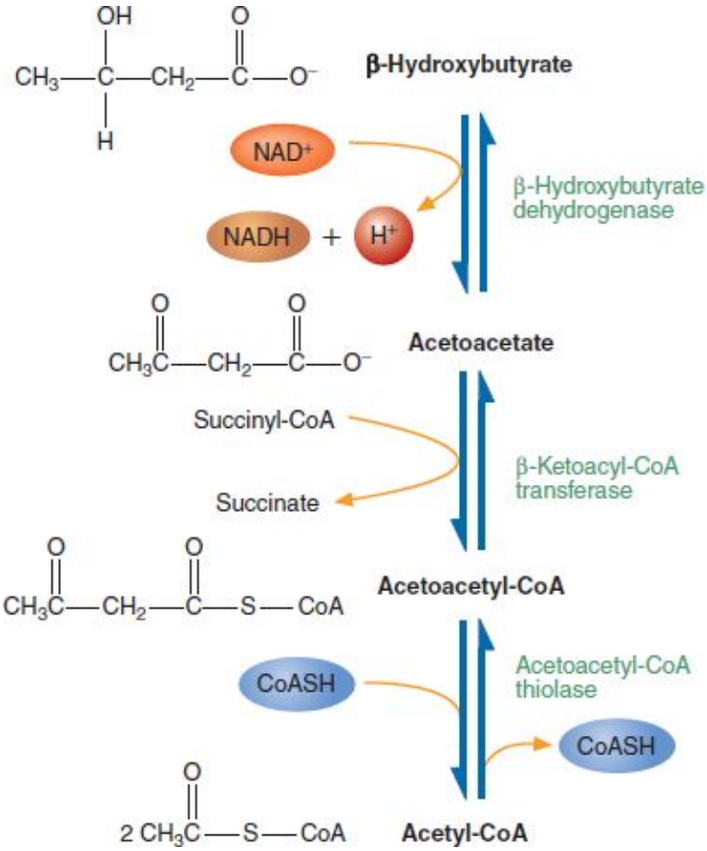


FIGURE 12.10

Conversion of Ketone Bodies to Acetyl-CoA

Some organs (e.g., heart and skeletal muscle) can use ketone bodies (*β*-hydroxybutyrate and acetoacetate) as an energy source under normal conditions. During starvation, however, ketone bodies become an important fuel source for the brain. Because liver does not have *β*-ketoacid-CoA transferase, it cannot use ketone bodies as an energy source. These reactions are reversible. The energy yield from the catabolism of *β*-hydroxybutyrate, 21.5 ATPs, is calculated as follows. Two acetyl CoA products yield 20 ATPs in the citric acid cycle, electron transport, and oxidative phosphorylation. An additional NADH, produced by the oxidation of *β*-hydroxybutyrate to form acetoacetate, yields 2.5 ATPs. Because of the activation of acetoacetate by succinyl CoA, one ATP equivalent is subtracted from the sum of 22.5 ATPs thereby yielding 21.5 ATPs.

Fatty Acid Oxidation: Double Bonds and Odd Chains

The *β*-oxidation pathway degrades saturated fatty acids with an even number of carbon atoms. Certain additional reactions are required to degrade unsaturated, odd-chain, and branched-chain fatty acids.

UNSATURATED FATTY ACID OXIDATION The oxidation of unsaturated fatty acids such as oleic acid requires additional enzymes. They are needed because, unlike the trans double bonds introduced during *β*-oxidation, the double bonds of most naturally occurring unsaturated fatty acids have a cis configuration. The enzyme enoyl-CoA isomerase converts the *cis*-*β*,*γ* double bond to a *trans*-*α*,*β* double bond. **Figure 12.11** illustrates the *β*-oxidation of oleic acid.

ODD-CHAIN FATTY ACID OXIDATION Although most fatty acids contain an even number of carbon atoms, certain organisms (e.g., some plants and microorganisms) produce odd-chain fatty acid molecules. *β*-Oxidation of such fatty acids proceeds normally until the last *β*-oxidation cycle, which yields one acetyl-CoA molecule and one propionyl-CoA molecule. Propionyl-CoA is then converted to succinyl-CoA, a citric acid cycle intermediate (**Figure 12.12**). Ruminant animals such as cattle and sheep derive a substantial amount of energy from the oxidation of odd-chain fatty acids.

These molecules are produced by microbial fermentation processes in the rumen (stomach).

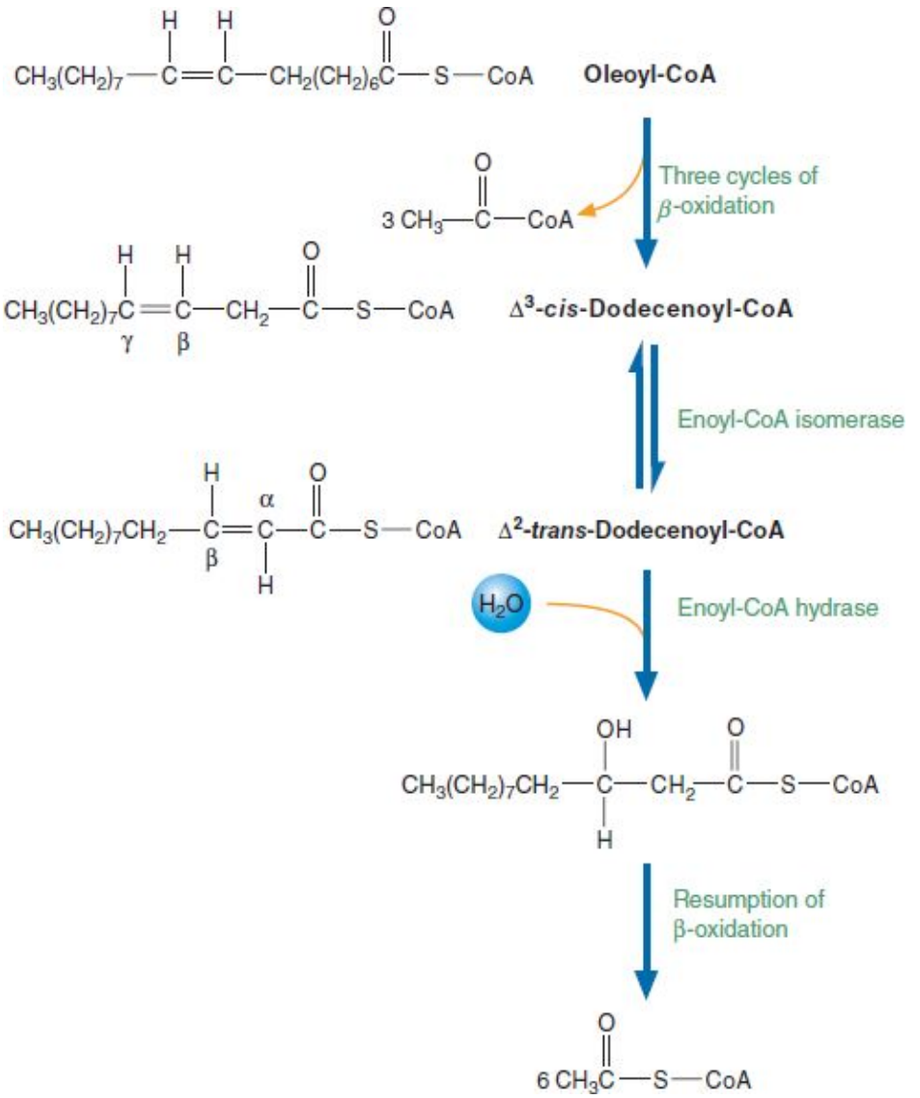


FIGURE 12.11

***β*-Oxidation of Oleoyl-CoA**

β-Oxidation of the CoA derivative of oleic acid progresses until Δ^3-*cis*-dodecenoyl-CoA is produced. This molecule is not a suitable substrate for *β*-oxidation because it contains a cis double bond. After conversion of the *β*,*γ*-*cis* double bond to an *α*,*β*-*trans* double bond, *β*-oxidation resumes.

α-OXIDATION *α-Oxidation* is a mechanism for degrading branched-chain fatty acid molecules such as phytanic acid. Phytanic acid, a 20-carbon fatty acid, is an oxidation product of phytol, a diterpene alcohol esterified to the photosynthetic pigment chlorophyll. Phytol, found in green vegetables, is converted to phytanic acid after ingestion. Phytanic acid is a component of dairy products and other foods derived from herbivorous (plant-eating) animals. In humans, *α*-oxidation takes place in peroxisomes.

β-Oxidation of phytanic acid is blocked by the methyl group substituent on C-3 (the *β*-position). Consequently, the first step in phytanic acid catabolism is an *α*-oxidation in which the molecule is converted to a *α*-hydroxy fatty acid. This reaction is followed by the removal of the carboxyl group (**Figure 12.13**). After activation to a CoA derivative, the product, pristanic acid, can be further degraded by *β*-oxidation. All subsequent side chain methyl groups will now be in the *α*-position, which is not a problem for *β*-oxidation enzymes. Phytanic acid oxidation is critical because large quantities of this molecule are found in the diet.

Refsum's disease (also called *phytanic acid storage syndrome*) is a rare autosomal recessive condition caused by a missing or defective gene that codes for phytanoyl-CoA hydroxylase, resulting in serious neurological problems. Nerve damage occurs when phytanic acid accumulates to the point where it interferes with myelination (myelin sheath formation; see p. 413). Nerve damage can be significantly reduced by eating less phytanic acid-containing foods (i.e., dairy products).

KEY CONCEPTS

Several reactions in addition to β-oxidation are required to degrade unsaturated, odd-chain, and branched-chain fatty acids.

Refsum's Disease

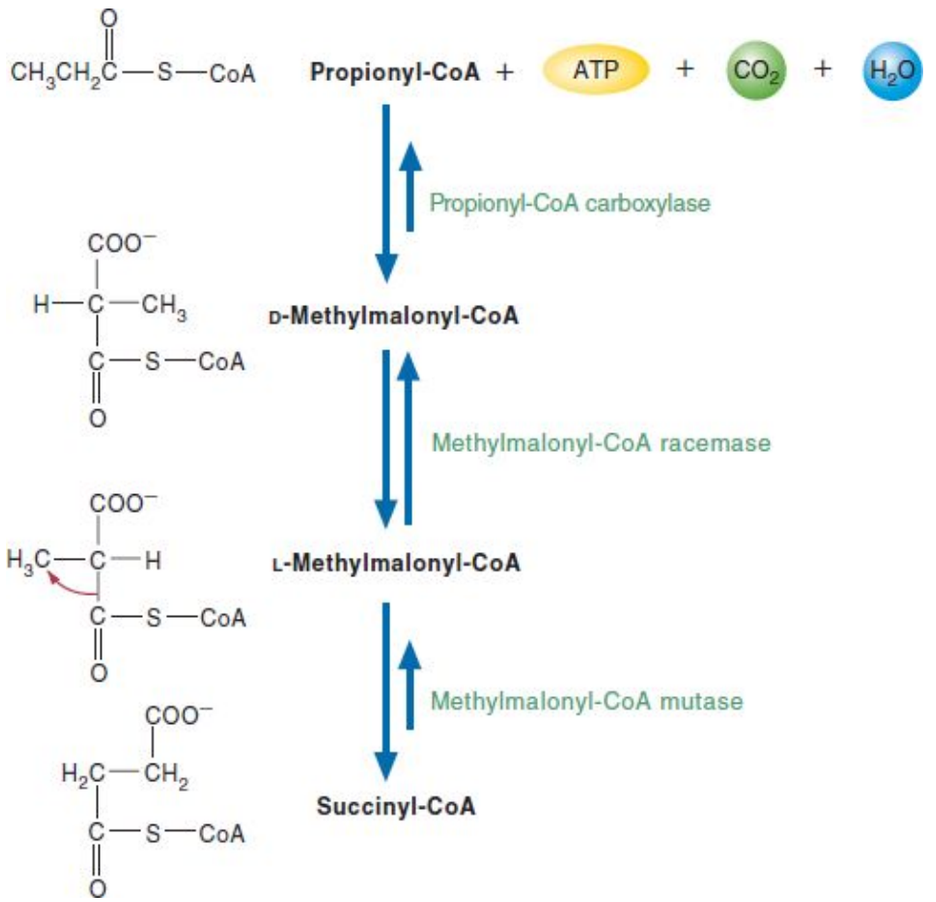


FIGURE 12.12

Conversion of Propionyl-CoA to Succinyl-CoA

In the first step, propionyl-CoA is carboxylated by propionyl-CoA carboxylase, an enzyme with a biotin cofactor (see p. 463). The product, D-methylmalonyl-CoA, is isomerized by methylmalonyl-CoA racemase to form L-methylmalonyl-CoA. In the last step, a hydrogen atom and the carbonyl-CoA group exchange positions. This unusual reaction is catalyzed by methylmalonyl-CoA mutase, an enzyme that requires 5′-deoxyadenosylcobalamin, usually designated as vitamin B_{12}.

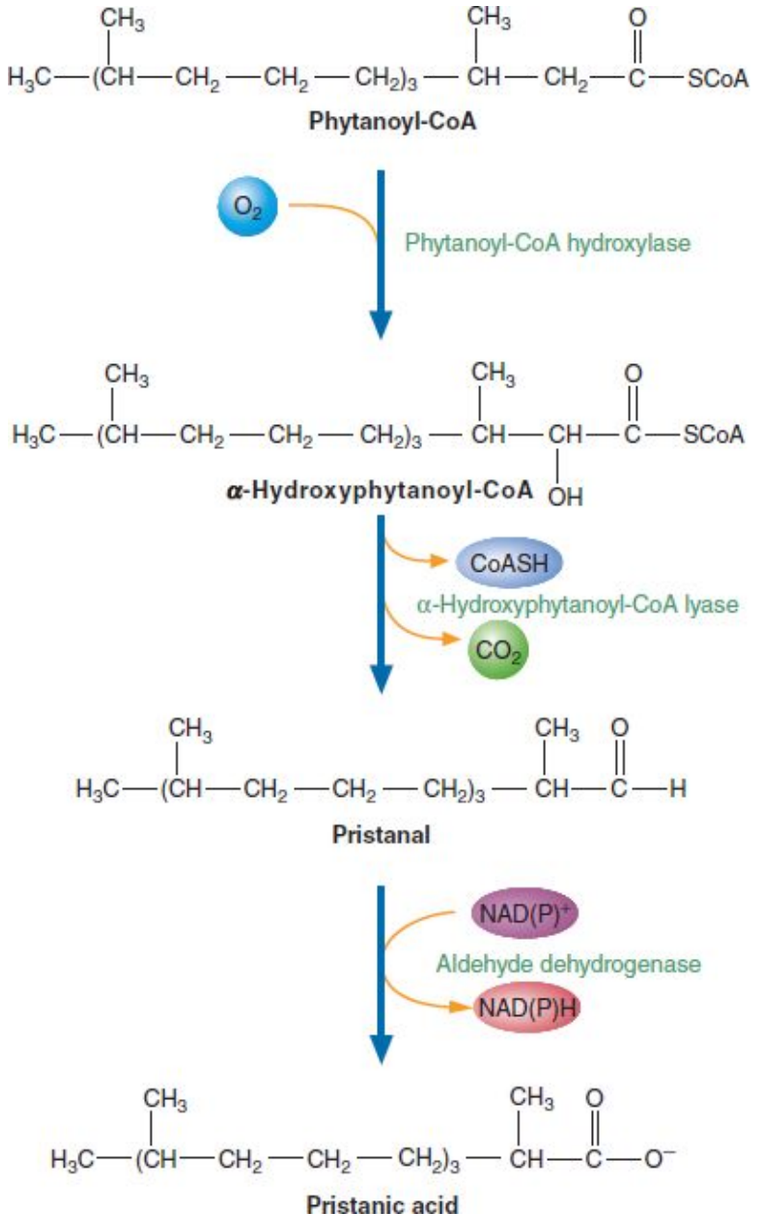


FIGURE 12.13

***α*-Oxidation of Phytanic Acid in Peroxisomes**

In *α*-oxidation, phytanoyl-CoA is converted to *α*-hydroxyphytanoyl-CoA in a reaction requiring O_2 and *α*-ketobutyrate and catalyzed by an Fe^{2+} and ascorbate-requiring enzyme to yield *α*-hydroxyphytanoyl-CoA, succinate, and CO_2. *α*-Hydroxyphytanoyl-CoA is then converted to pristanal in a thiamine pyrophosphate (TPP)–requiring decarboxylation reaction in which the *α*-carbon is oxidized. The thioester bond of the other product, formyl-CoA, is subsequently cleaved to form CoASH and formic acid (HCOOH), which is then oxidized to yield CO_2. Pristanal is oxidized in an $NAD(P)^+$-requiring reaction to form pristanic acid. Pristanic acid is further degraded by *β*-oxidation after esterification with CoASH. The products of this process are three acetyl-CoAs, three propionyl-CoAs, and one isobutyl-CoA.

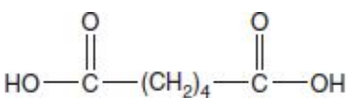


FIGURE 12.14

Adipic Acid

3D animation of adipic acid

QUESTION 12.6

In the past, mammals were believed to be unable to use fatty acids in gluconeogenesis. (Acetyl-CoA cannot be converted to pyruvate because the reaction catalyzed by pyruvate dehydrogenase is irreversible.) Recent experimental evidence indicates that certain unusual fatty acids (i.e., those with odd chains or two carboxylic acid groups) can be converted to glucose in small but measurable quantities. One molecule of propionyl-CoA is produced when an odd-carbon chain fatty acid is oxidized. Describe a possible biochemical pathway by which a liver cell might synthesize glucose from propionyl-CoA. [*Hint*: Refer to **Figure 12.12**.]

QUESTION 12.7

One product of the β-oxidation of dicarboxylic acids is succinyl-CoA. Propose a biochemical pathway for the conversion of the molecule illustrated in **Figure 12.14** to glucose.

Fatty Acid Biosynthesis

Although fatty acid synthesis occurs within the cytoplasm of most animal cells, liver is the major site for this process. (Recall, for example, that liver produces VLDL. See p. 421.) Fatty acids are synthesized when the diet is low in fat and/or high in carbohydrate or protein. Most fatty acids are synthesized from excess dietary carbohydrate. As discussed, glucose is converted to pyruvate in the cytoplasm. After entering the mitochondrion, pyruvate is converted to acetyl-CoA, which condenses with OAA, a citric acid cycle intermediate, to form citrate. When mitochondrial citrate levels are sufficiently high (i.e., cellular energy requirements are low), citrate enters the cytoplasm, where it is cleaved to form acetyl-CoA and OAA. Acetyl-CoA is then used in fatty acid biosynthesis. The net reaction for the synthesis of palmitic acid from acetyl-CoA is as follows:

$$\begin{aligned} &8\ \text{Acetyl-CoA} + 14\ \text{NADPH} + 14\ \text{H}^+ + 7\ \text{ATP} \longrightarrow \\ &\text{Palmitate} + 14\ \text{NADP}^+ + 7\ \text{ADP} + 7\ \text{P}_i + 8\ \text{CoASH} + 6\ \text{H}_2\text{O} \end{aligned} \tag{8}$$

A relatively large quantity of NADPH is required in fatty acid synthesis. A substantial amount of NADPH is provided by the pentose phosphate pathway (see p. 309). Reactions catalyzed by isocitrate dehydrogenase (see p. 345) and malic enzyme (see p. 352) provide smaller amounts.

The biosynthesis of fatty acids is outlined in **Figure 12.15**. At first glance, fatty acid synthesis appears to be the reverse of the β-oxidation pathway. For example, fatty acids are constructed by the sequential addition of two-carbon groups supplied by acetyl-CoA. Additionally, the same intermediates (i.e., β-ketoacyl-, β-hydroxyacyl-, and α,β-unsaturated acyl groups) are found in both pathways. A closer examination, however, reveals several notable differences between fatty acid synthesis and β-oxidation. First, fatty acid synthesis occurs predominantly in the cytoplasm. (Recall that β-oxidation occurs within mitochondria and peroxisomes.) Second, the enzymes that catalyze

fatty acid synthesis are significantly different in structure from those in β-oxidation. In eukaryotes, most of these enzymes are components of a multienzyme complex referred to as fatty acid synthase. Third, the intermediates in fatty acid synthesis are linked through a thioester linkage to **acyl carrier protein** (ACP), a component of fatty acid synthase. (Recall that acyl groups are attached to CoASH through a thioester linkage during β-oxidation.) Note that a phosphopantetheine prosthetic group (**Figure 12.16**) is involved in the linkage acyl groups to both ACP and CoASH). And finally, in contrast to β-oxidation, which produces NADH and $FADH_2$, fatty acid synthesis consumes NADPH. Fatty acid synthesis occurs in two phases: the carboxylation of acetyl-CoA to form malonyl-CoA by acetyl-CoA carboxylase and the synthesis of palmitate by the sequential addition of two carbon units to a growing fatty acyl chain by fatty acid synthase.

ACETYL-COA CARBOXYLASE The carboxylation of acetyl-CoA to form malonyl-CoA is an irreversible reaction that is catalyzed by acetyl-CoA carboxylase (ACC) (**Figure 12.17**). The first phase of this reaction is the ATP-dependent carboxylation of biotin to give carboxybiotin. Subsequent decarboxylation results in the transfer of an activated CO_2 from biotin to acetyl-CoA. Acetyl-CoA carboxylation, the rate-limiting step in fatty acid synthesis, is an activating reaction that is necessary because this carbon–carbon condensation is thermodynamically unfavorable. As a result of resonance stabilization, free carboxylate groups are insufficiently reactive.

Acetyl-CoA carboxylase is found in most organisms. In eukaryotes, ACC contains three domains: BCCP (biotin carboxyl carrier protein), BC (biotin carboxylase), and CT (carboxyltransferase). Biotin, a coenzyme that carries carboxyl groups, is bound to BCCP via an amide linkage to the side chain of a lysine residue. The flexible lysine side chain is, in effect, a swinging arm that transfers the newly carboxylated biotin from the active site of the BC domain to the active site of the CT domain (a 7-Å distance). CT then catalyzes the transfer of the carboxyl group from biotin to acetyl-CoA to form the product malonyl-CoA.

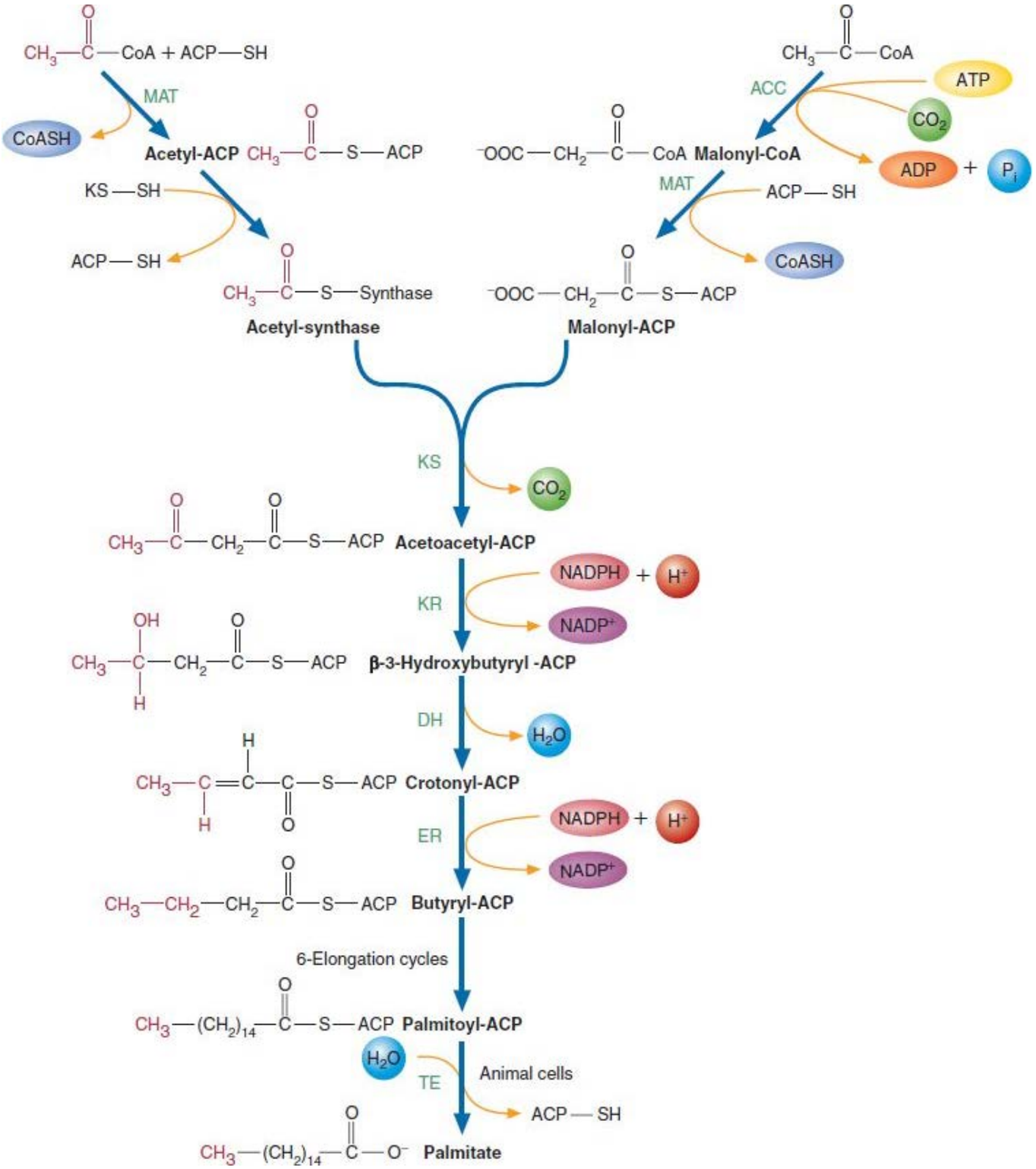


FIGURE 12.15

Fatty Acid Biosynthesis

The substrates for fatty acid synthesis are acetyl-ACP and malonyl-ACP, formed from ACP and acetyl-CoA and malonyl-CoA, respectively. Both reactions are catalyzed by malonyl/acetyl transferase (MAT). Formation of a fatty acid chain begins with a condensation reaction. Catalyzed by *β*-ketoacyl synthase (KS), the acetyl group (linked to KS via a thioester linkage) is transferred to the malonyl group to form acetoacetyl-ACP. The reduction of the *β*-carbonyl group, catalyzed by *β*-ketoacyl-ACP reductase (KR), forms an alcohol. The removal of water to form a carbon–carbon double bond is catalyzed by *β*-hydroxyacyl-ACP dehydratase (DH). Reduction by enoyl-ACP reductase (ER) yields a saturated four-carbon acyl group. This acyl group is then transferred from ACP to the SH group of KS to begin a new elongation cycle. The acyl chain lengthens by two carbons as it condenses with another ACP-linked malonyl group. In animal cells, fatty acid synthesis ends with the release of palmitate from

ACP, catalyzed by thioesterase (TE). In each elongation cycle, two NADPH provide the reducing equivalents (high-energy electrons) required for the reactions catalyzed by KR and ER.

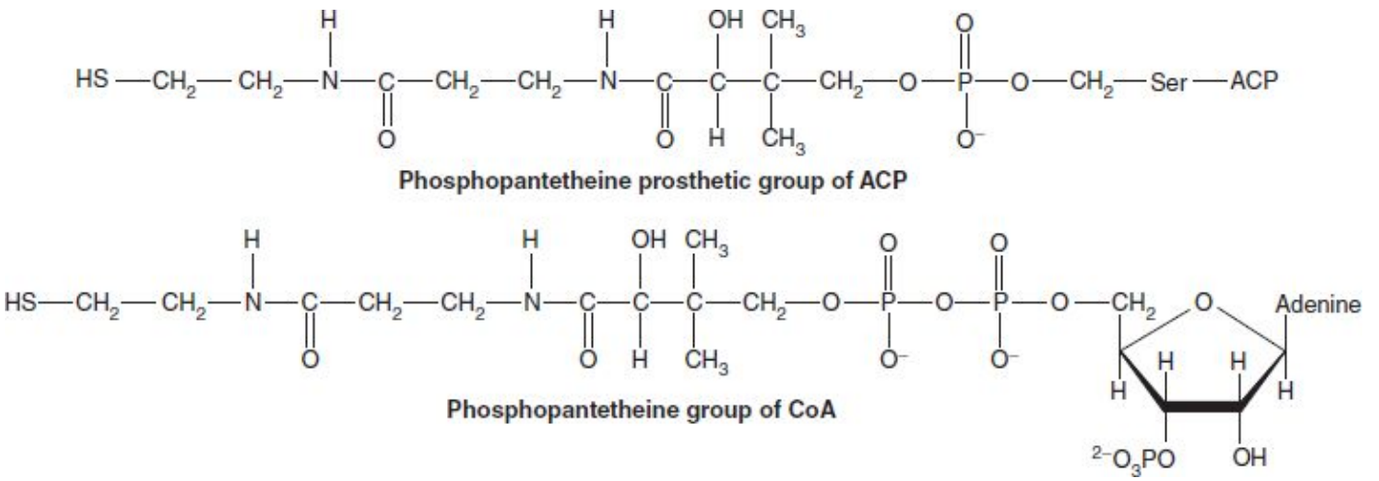


FIGURE 12.16

Comparison of the Phosphopantetheine Group in Acyl Carrier Protein (ACP) and in Coenzyme A (CoASH)

Fatty acids are attached to this prosthetic group on ACP during fatty acid biosynthesis and on CoASH during β-oxidation.

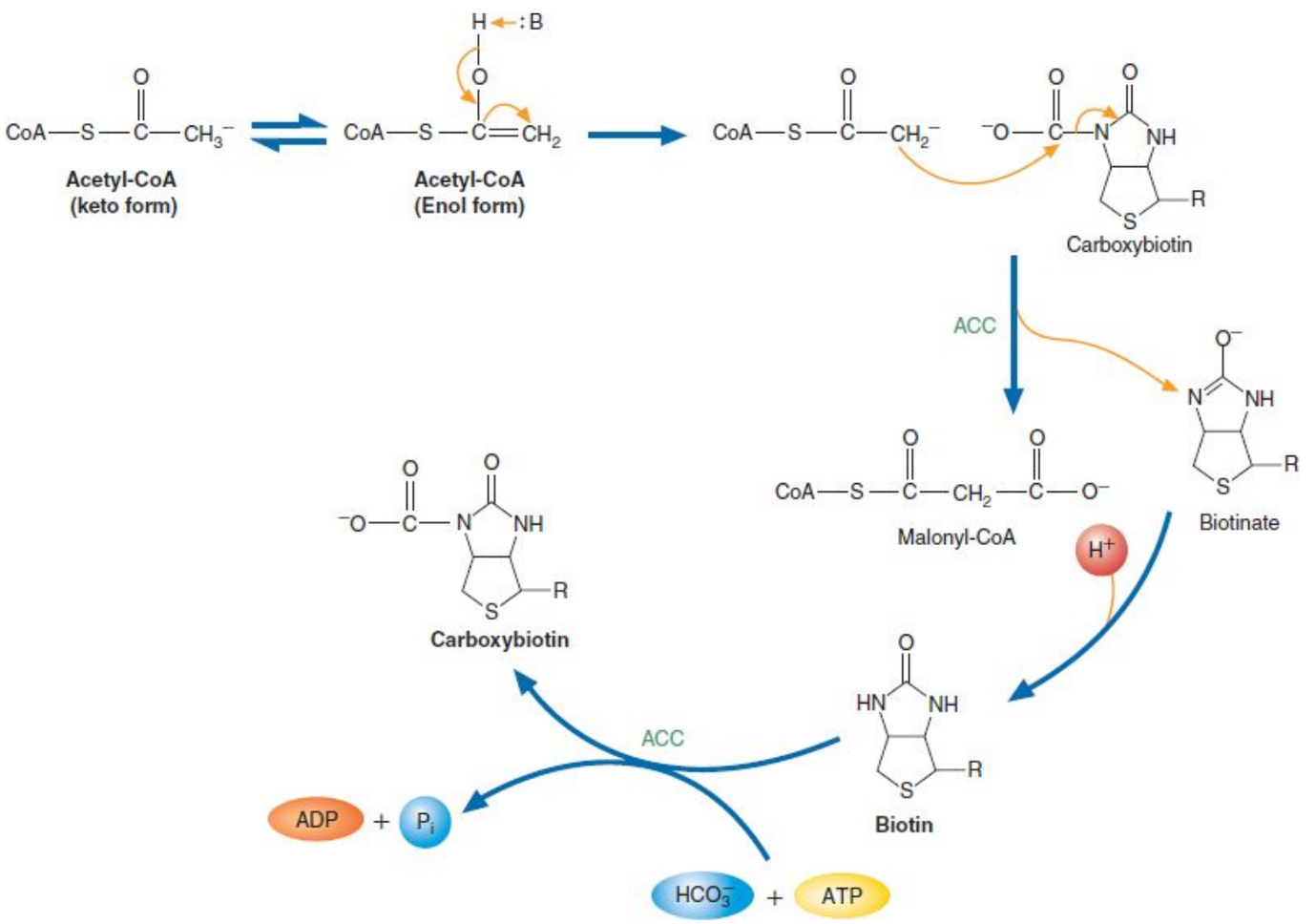


FIGURE 12.17

Synthesis of Malonyl-CoA

The reaction begins with the ATP-dependent carboxylation of the biotin cofactor of acetyl-CoA carboxylase (ACC). The carboxylase abstracts a proton from the α-carbon of the enol form of acetyl-CoA to generate a reactive carbanion. The carbanion attacks the carbonyl carbon of carboxybiotin to yield malonyl-CoA and biotinate. The biotinate is protonated by the enzyme to regenerate its biotin form. ACC then regenerates carboxybiotin, which reacts with another acetyl-CoA.

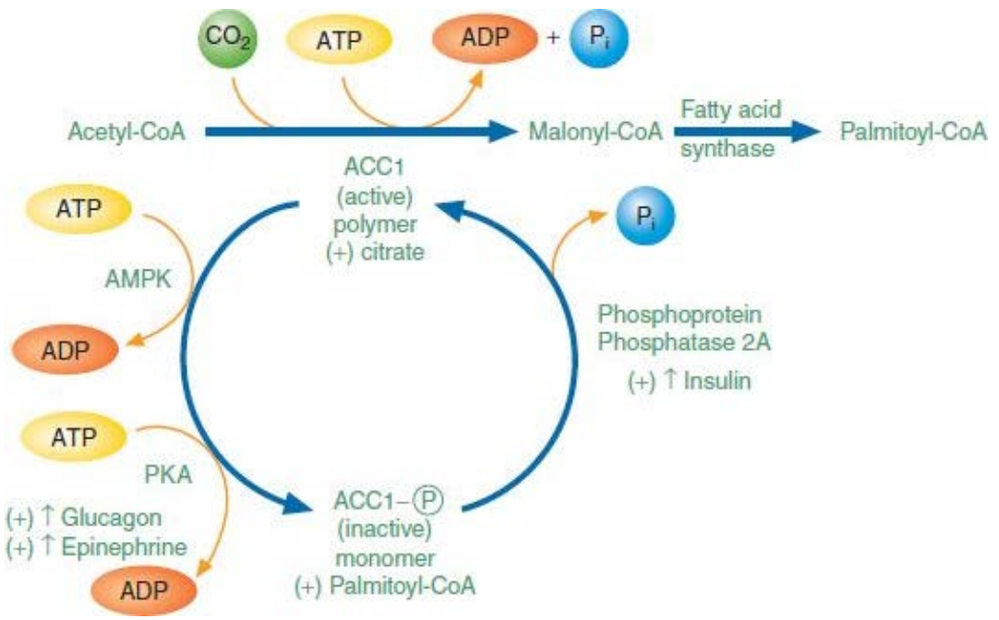


FIGURE 12.18

Regulation of Acetyl-CoA Carboxylase 1

ACC1 is inactivated (depolymerized) by phosphorylation reactions catalyzed by AMPK (as a result of high AMP levels), PKA (stimulated by glucagon or epinephrine), and palmitoyl-CoA accumulation. ACC1 is activated (polymerized) by phosphoprotein phosphatase 2A (PP-2A), which in turn is activated by high insulin levels (glucose readily available) and deactivated by high glucagon levels (low blood glucose).

In mammals, there are two forms of ACC. A cytoplasmic enzyme, ACC1, is expressed in lipogenic tissues such as liver, adipose tissue, and lactating mammary gland. ACC2, associated with the outer mitochondrial membrane, occurs in oxidative tissues such as cardiac and skeletal muscle, where its product, malonyl-CoA, acts as a strong inhibitor of carnitine acyltransferase I. ACC2, therefore, serves a regulatory function in fatty acid oxidation (p. 468). The liver, which oxidizes and synthesizes fatty acids, contains both forms of ACC.

Mammalian ACC contains two subunits, each with a bound biotin cofactor. ACC becomes active when ACC dimers aggregate to form high-molecular-weight polymers (4 million to 8 million Da) composed of 10 to 20 dimers. ACC, a key enzyme in fatty acid metabolism, is highly regulated by allosteric modulators and phosphorylation reactions (**Figure 12.18**). The allosteric effects of citrate, a feed-forward activator that promotes polymerization, and palmitoyl-CoA, an end product inhibitor that causes depolymerization, are dependent on the phosphorylation state of the enzyme. ACC is phosphorylated and therefore inhibited (depolymerized) by AMPK, an important regulatory enzyme in energy metabolism (p. 470).

Phosphorylation by the cAMP-activated protein kinase PKA, stimulated by glucagon and epinephrine, also plays a role in ACC inhibition. Depolymerization is also favored by the presence of palmitoyl-CoA, which binds to and stabilizes the dimeric form of the enzyme. Glucagon and epinephrine help to maintain the phosphorylated inactive form of ACC by inactivating phosphoprotein phosphatase-2A (PP-2A), the enzyme that mediates the dephosphorylation of a

number of target proteins, including ACC. ACC activation occurs when it is dephosphorylated by PP-2A, a reaction promoted by insulin. The polymerized form of ACC is stabilized by binding to citrate, which accumulates when acetyl-CoA levels are high.

FATTY ACID SYNTHASE In humans, the remaining reactions in fatty acid synthesis take place on the fatty acid synthase multienzyme complex (FAS) (**Figure 12.19**). FAS is an X-shaped head-to-head homodimer composed of two identical 272-kDa polypeptides. Each polypeptide has seven catalytic domains and ACP. As a result, FAS synthesizes two fatty acids simultaneously. During fatty acid synthesis, the acyl intermediates are covalently bound to the 2-nm-long phosphopantetheine group of ACP through a thioester linkage. The flexibility of ACP, a relatively unstructured domain of FAS, and the phosphopantetheine group allow the transfer of attached acyl intermediates from one active site to another in the complex so efficiently that each multienzyme complex can synthesize two palmitates in less than a second.

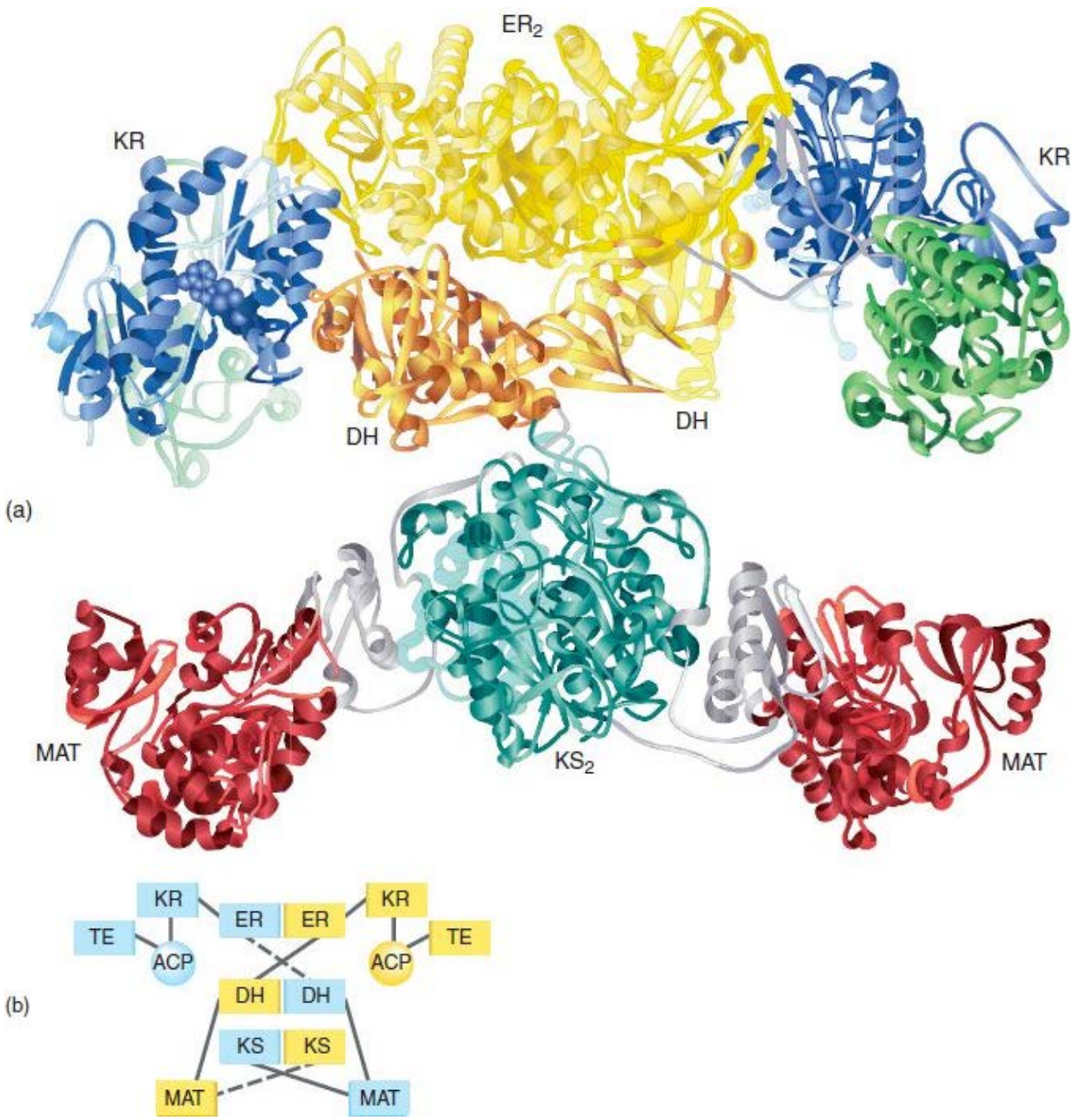


FIGURE 12.19

Fatty Acid Synthase Structure

(a) The structure of mammalian FAS is based on X-ray crystal studies. Note that because ACP is unstructured, its structure (not shown) remains unresolved. Each ACP domain is probably located near a KR domain. (b) This model illustrates the domain organization of the FAS homodimer. [KS = β-ketoacyl synthase; MAT = malonyl/acetyltransferase; DH = β-hydroxyacyl-ACP dehydratase; ER = enoyl-ACP reductase; KR = β-ketoacyl reductase; ACP = acyl carrier protein; TE = thioesterase.]

The synthesis of a fatty acid (**Figure 12.15**) is initiated with the transfer of the acetyl group of acetyl-CoA and the malonyl group of malonyl-CoA to ACP. Both reactions are catalyzed by malonyl/acetyl transferase (MAT). The acetyl group is then transferred from acetyl-ACP to a cysteinyl side chain of β-ketoacyl synthase (KS). KS then catalyzes a condensation reaction (**Figure 12.20**) in which the decarboxylation of the malonyl group creates a carbanion. The carbanion attacks the carbonyl carbon of the acetyl group to yield the product acetoacetyl-ACP.

During the next three steps, consisting of two reductions and a dehydration, the acetoacetyl group is converted to a butyryl group. β-Ketoacyl-ACP reductase (KR) catalyzes the reduction of acetoacetyl-ACP to form β-hydroxybutyryl-ACP. β-Hydroxyacyl-ACP dehydratase (DH) later catalyzes a dehydration, thus forming crotonyl-ACP. Butyryl-ACP is produced when 2,3-*trans*-enoyl-ACP reductase (ER) reduces the double bond in crotonyl-ACP. In the last step of the first cycle of fatty acid synthesis, the butyryl group is transferred from the phosphopantetheine group to the cysteine residue of KS. The newly freed ACP-SH group now binds to another malonyl group. The process is then repeated until, eventually, palmitoyl-ACP is synthesized. The palmitoyl group is released from fatty acid synthase when thioesterase cleaves the thioester bond. Depending on cellular conditions, palmitate can be used directly in the synthesis of several types of lipid (e.g., triacylglycerol or phospholipids), or it can enter the mitochondrion, where several enzymes catalyze elongating and desaturating reactions. ER possesses similar enzymes.

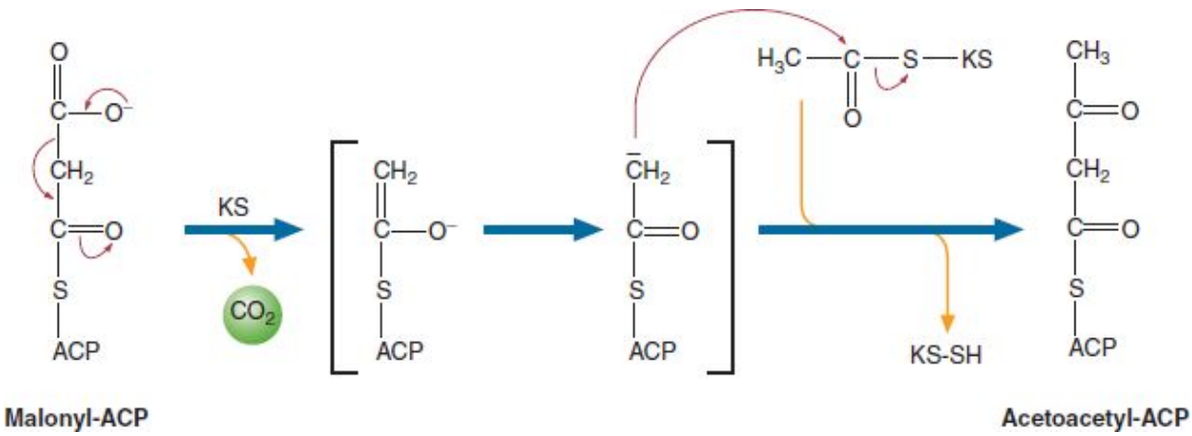


FIGURE 12.20

Formation of Acetoacetyl-ACP

The decarboxylation of malonyl-ACP, catalyzed by β-ketoacyl-ACP synthase (KS), results in the formation of a carbanion. The attack of the carbanion on the carbonyl carbon of an acetyl group linked via a thioester bond to the enzyme yields acetoacetyl-ACP.

FATTY ACID ELONGATION AND DESATURATION Elongation and desaturation of fatty acids synthesized in cytoplasm or obtained from the diet are accomplished primarily by ER enzymes. Fatty acid elongation and desaturation (the formation of double bonds) are especially important in the regulation of membrane fluidity and the synthesis of the precursors for a variety of fatty acid derivatives, such as the eicosanoids. For example, myelination (p. 459) depends especially on the ER fatty acid synthetic reactions. Very-long-chain saturated and monounsaturated fatty acids are

important constituents of the cerebrosides and sulfatides found in myelin. Cells regulate membrane fluidity by adjusting the types of fatty acids that are incorporated into membrane lipids. For example, in cold weather, more unsaturated fatty acids are incorporated. (Recall that unsaturated fatty acids have a lower freezing point than do saturated fatty acids. See p. 404.) If the diet does not provide a sufficient number of these molecules, fatty acid synthetic pathways are activated. Although elongation and desaturation are closely integrated processes, for the sake of clarity they are discussed separately here.

ER fatty acid chain elongation, which uses two-carbon units provided by malonyl-CoA, is a cycle of condensation, reduction, dehydration, and reduction reactions similar to those observed in cytoplasmic fatty acid synthesis. In contrast to the cytoplasmic process, the intermediates in the ER elongation process are CoA esters. These reactions can lengthen both saturated and unsaturated fatty acids. Reducing equivalents are provided by NADPH.

Acyl-CoA molecules are desaturated in ER membrane in the presence of NADH and O_2. Cytochrome b_5 reductase (a flavoprotein), cytochrome b_5, and oxygen-dependent desaturases, functioning together as an electron transport system, efficiently introduce double bonds into long-chain fatty acids (**Figure 12.21**). Both the flavoprotein and cytochrome b_5 (found in a ratio of approximately 1:30) have hydrophobic peptides that anchor the proteins into the ER membrane. Animals typically have Δ^9, Δ^6, and Δ^5 desaturases that use electrons supplied by NADH to activate the oxygen needed to create the double bond. Because elongation and desaturation systems are in close proximity to each other, a variety of long-chain polyunsaturated acids are typically produced. One example of this interaction is the synthesis of arachidonic acid ($20{:}4^{\Delta 5,8,11,14}$) from linoleic acid ($18{:}2^{\Delta 9,12}$).

KEY CONCEPTS

- In animals, fatty acids are synthesized in the cytoplasm from acetyl-CoA and malonyl-CoA.
- Mitochondrial and ER enzymes elongate and desaturate newly synthesized fatty acids as well as those obtained in the diet.

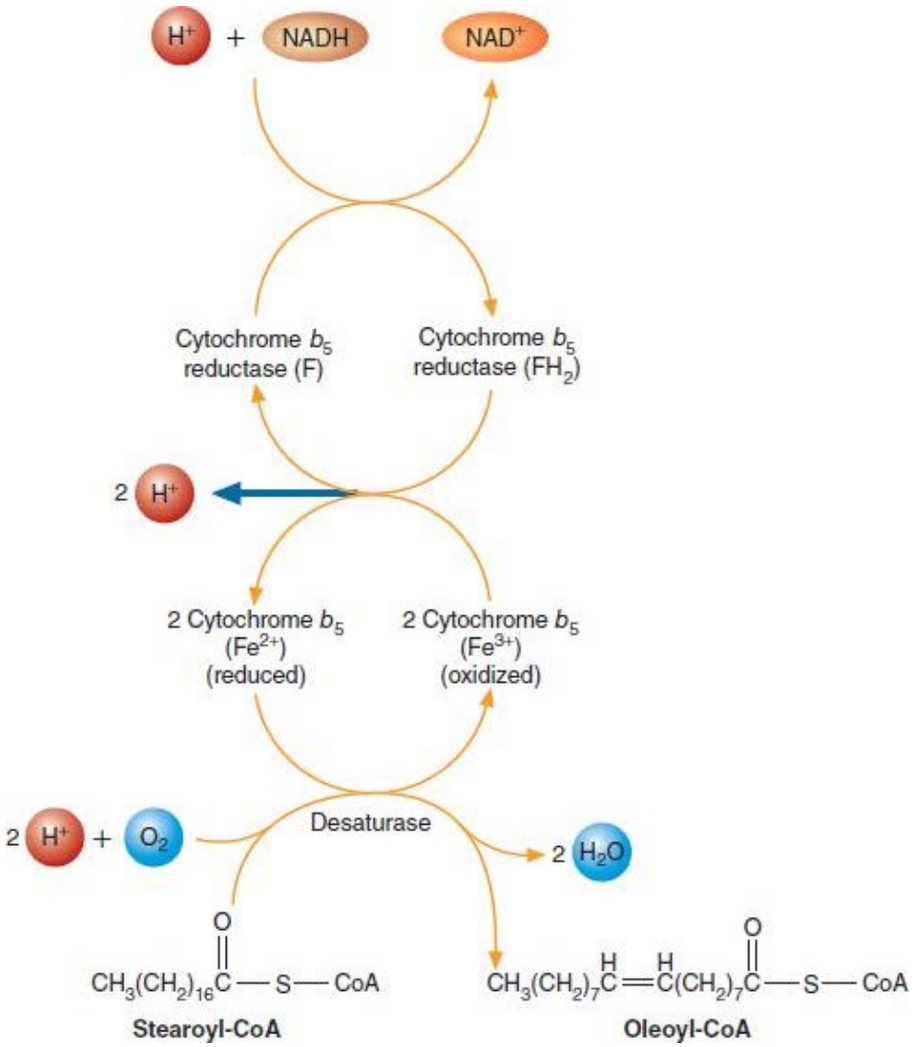


FIGURE 12.21

Desaturation of Stearoyl-CoA

The desaturase uses electrons provided by an electron transport system composed of cytochrome b_5 reductase and cytochrome b_5 to activate the oxygen (not shown) needed to create the double bond. NADH is the electron donor.

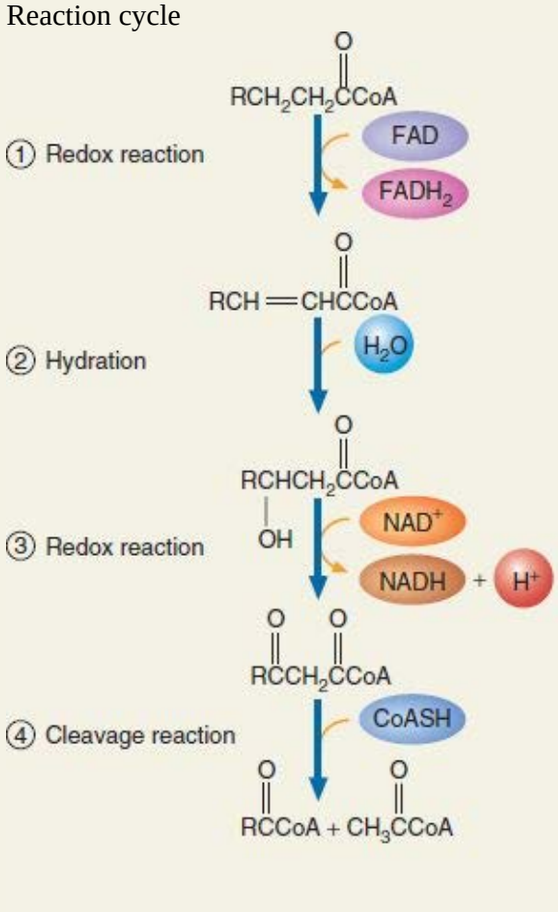
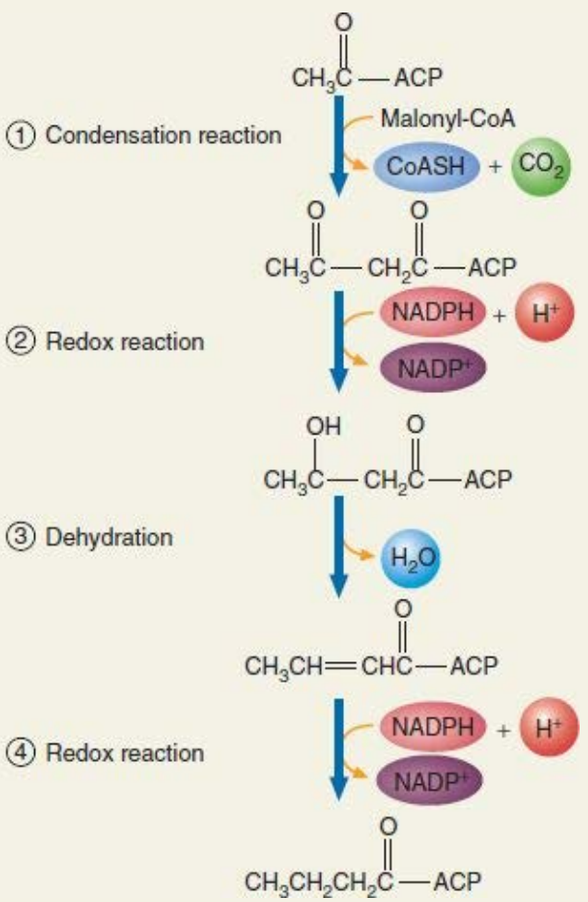
COMPARISON OF FATTY ACID OXIDATION AND FATTY ACID SYNTHESIS The functions of β-oxidation and fatty acid synthesis are clearly different. β-Oxidation degrades fatty acids to yield acetyl-CoA, the substrate of the energy-generating citric acid cycle. In contrast, energy is stored when acetyl-CoA is converted into fatty acids. Although the cellular locations, the redox coenzymes, the acyl group carriers, and the enzymes involved in the β-oxidation and fatty acid synthetic pathways are quite different, the reactions are similar enough to cause confusion. **Table 12.1** outlines the differences between the two processes.

QUESTION 12.8

Excessive consumption of fructose has been linked to obesity and to the condition referred to as *hypertriglyceridemia* (high blood levels of triacylglycerols). Common sources of fructose for most Americans are sucrose and high-fructose corn syrup. Over the past several decades, high-fructose corn syrup has replaced sucrose in many processed foods and beverages because it is inexpensive in comparison to sucrose. It now comprises at least 40% of caloric sweeteners. (The fructose content of fresh fruits and vegetables is so low that it would be difficult to consume sufficient

quantities to induce hypertriglyceridemia.) Sucrose is digested in the small intestine by the enzyme sucrase to yield one molecule each of fructose and glucose. Digestion is so rapid that the blood concentrations of these sugars can become quite high when large quantities are consumed (e.g., sugar-containing soft drinks). Whatever its source, once fructose reaches the liver it is converted to fructose-1-phosphate (see p. 313). In addition, whereas high blood levels of glucose trigger the release of insulin and *leptin* (a hormone secreted by adipose tissue), both of which curb appetite, this does not happen with fructose. After reviewing fructose metabolism and fatty acid and triacylglycerol synthesis, suggest how hypertriglyceridemia and obesity might result from a diet that is rich in sucrose and high-fructose corn syrup.

TABLE 12.1 Comparisons of Fatty Acid *β*-Oxidation and Fatty Acid Synthesis

	Fatty Acid *β*-Oxidation	Fatty Acid Synthesis
Subcellular location	Mitochondrion	Cytoplasm
Substrates	Acyl-CoA, FAD, NAD^+, CoASH	Acetyl-CoA, NADPH
Products	Acetyl-CoA, $FADH_2$, NADH	Palmatate, $NADP^+$, CO_2, CoASH
Reaction cycle		

Hypertriglyceridemia

Regulation of Fatty Acid Metabolism in Mammals

Animals have such varying requirements for energy that the metabolism of fatty acids (*the* major energy source in animals) is carefully regulated. Both short- and long-term regulatory mechanisms are used. In short-term regulation (measured in minutes), the activities of existing molecules of key regulatory enzymes are modified by allosteric regulators (**Figure 12.22**), covalent modification, and hormones. When energy levels are high, β-oxidation is depressed by the binding of the allosteric modulators NADH and acetyl-CoA to β-hydroxyacyl–CoA dehydrogenase and thiolase, respectively. Similarly, malonyl-CoA, the product of ACC, is an allosteric regulator of CAT-I. In liver when insulin/glucagon ratios are high, malonyl-CoA levels rise and cause the inhibition of β-oxidation, thus preventing a futile cycle. High cellular concentrations of long-chain fatty acyl-CoA esters inhibit ACC by promoting its depolymerization.

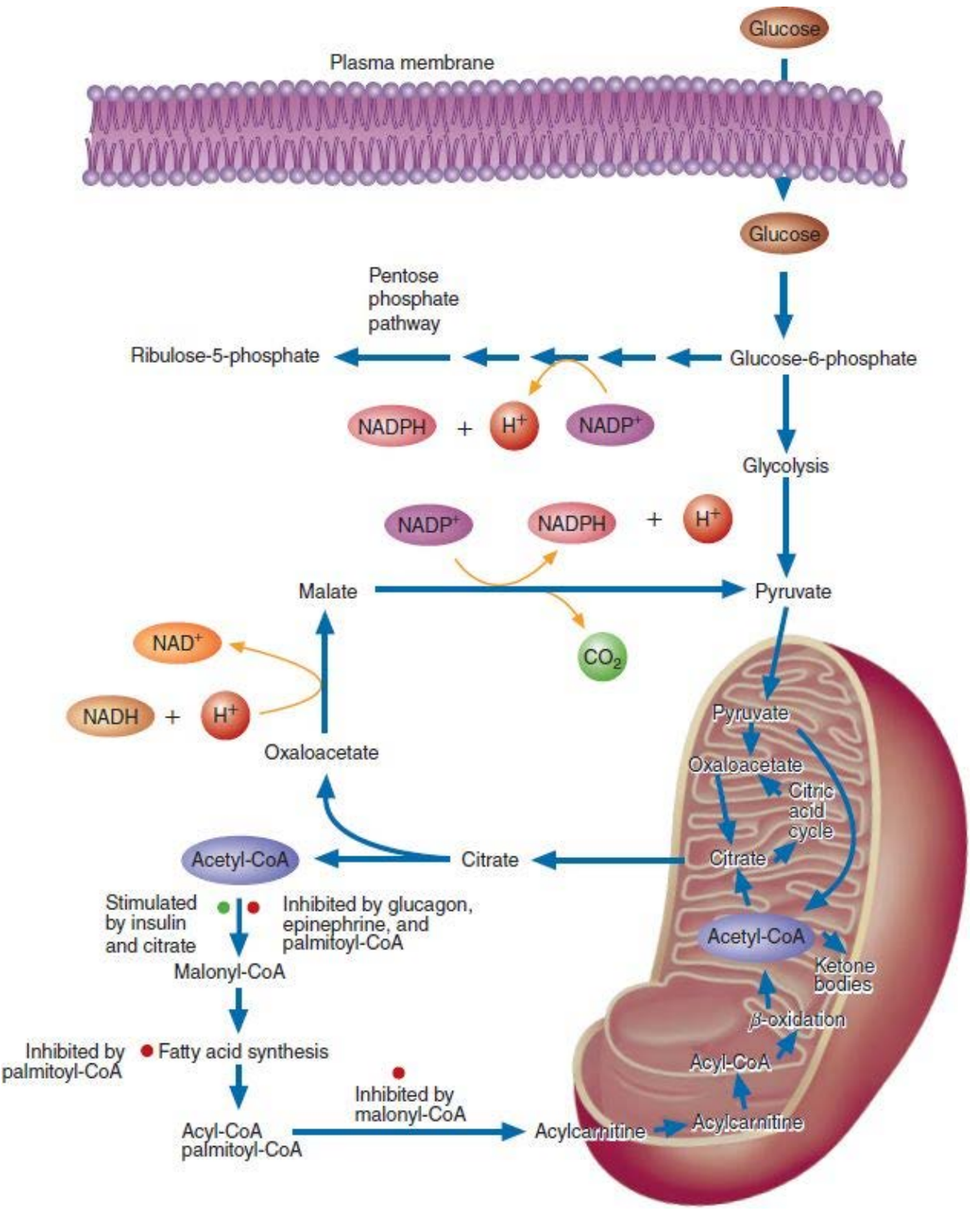


FIGURE 12.22

Regulation of Intracellular Fatty Acid Metabolism

Fatty acids are synthesized in cytoplasm from acetyl-CoA, which is formed within the mitochondrion. Because the inner membrane is impermeable to acetyl-CoA, it is transferred out as citrate. Citrate is produced from acetyl-CoA and oxaloacetate in the citric acid cycle, a reaction pathway in the mitochondrial matrix. Citrate is transferred to cytoplasm when β-oxidation is suppressed, that is, when the cell needs little energy. It is then cleaved to form oxaloacetate and acetyl-CoA. When the cell needs more energy, fatty acids are transported into the mitochondrion as acylcarnitine derivatives. Then acyl-CoA is degraded to acetyl-CoA via β-oxidation. Note that glucagon facilitates fatty acid oxidation, possibly by stimulating CAT-I. AMPK, the hormones glucagon and epinephrine and the substrates citrate, malonyl-CoA, and palmitoyl-CoA are important regulators of fatty acid metabolism. A portion of the NADPH, the reducing agent required in fatty acid synthesis, is generated by several reactions in the pentose phosphate pathway. NADPH is also produced by converting malate, formed by the reduction of oxaloacetate, to pyruvate.

AMPK Fatty acid synthesis and β-oxidation are also rapidly regulated by changes in energy demand by AMPK. **AMPK** (pp. 300–01) is a trimeric enzyme composed of an α subunit (catalytic) and β and γ subunits (regulatory). As AMP/ATP ratios begin to rise, AMPK is activated by upstream AMPK kinases and by its allosteric modulator, AMP. In addition to acting as an allosteric activator, AMP promotes the activating phosphorylation reactions and inhibits dephosphorylation by protein phosphatases. AMP levels are a sensitive indicator of cellular energy status, and they rise in response to stresses that deplete cellular ATP levels, such as nutrient deprivation, hypoxia, heat shock, and prolonged exercise. Once activated, AMPK switches off anabolic pathways, such as fatty acid and triacylglycerol synthesis by phosphorylating ACC1 (p. 464) and glycerol-3-phosphate acyltransferase, respectively. At the same time, AMPK switches on catabolic pathways: For example, β-oxidation is stimulated by AMPK-induced activation of malonyl-CoA decarboxylase (MCD), the enzyme that decreases the concentration of malonyl-CoA and inhibition of ACC2. AMPK's influence over the body's major metabolic processes is outlined in **Figure 12.23**.

HORMONES Hormones play an important role in both short- and long-term regulation of fatty acid metabolism. The short-term effects of insulin that promote fat synthesis are caused by rapid signal transduction mechanisms. Insulin activates phosphoprotein phosphatase 2A, which dephosphorylates and activates ACC1, in combination with allosteric regulators and covalent modifications (p. 464). Insulin also activates ATP-citrate lyase (p. 352) and pyruvate dehydrogenase (in adipocytes). Insulin promotes fat synthesis in adipocytes by triggering the movement of GLUT-4 (a glucose transporter; see p. 619) to the cell surface, thus facilitating the entry into the cell of glucose (the precursor of glycerol-3-phosphate and fatty acids). Insulin simultaneously depresses fat mobilization in adipocytes by stimulating the phosphorylation of hormone-sensitive lipase. Epinephrine increases lipolysis by stimulating the dephosphorylation and inactivation of hormone-sensitive lipase via signal transduction-mediated phosphorylation reactions. Glucagon increases fatty acid oxidation by an unresolved mechanism, possibly by activating CAT-I.

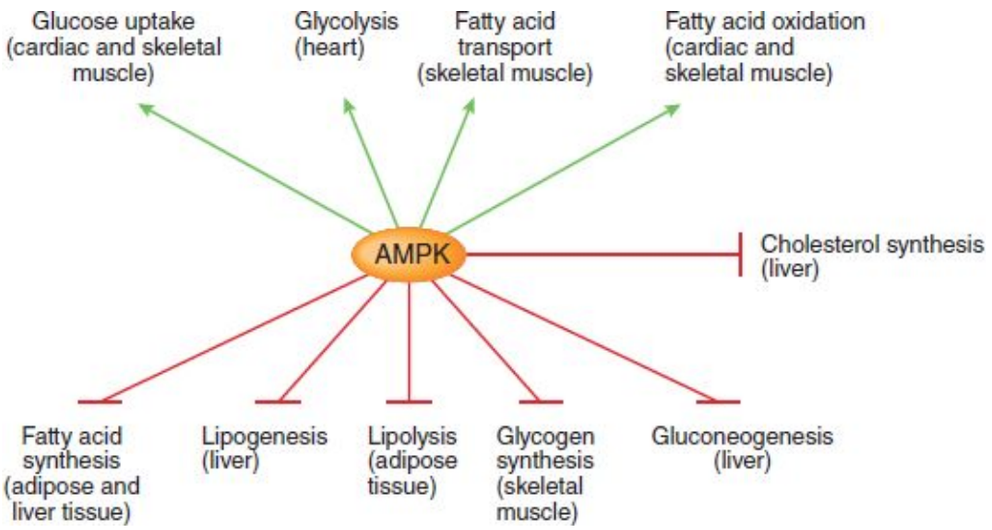


FIGURE 12.23

AMPK-Regulated Pathways in Lipid and Carbohydrate Metabolism

Elevated cellular AMP levels activate AMPK, a major metabolic switch that regulates numerous biochemical pathways. AMPK shifts metabolism from energy-consuming processes to energy-generating processes by phosphorylating target proteins. The effects of AMPK in the principal metabolic pathways in cardiac and skeletal muscle, liver, and adipose tissue are indicated in the diagram. AMPK also regulates the body's metabolism via its effects on insulin secretion from pancreatic β-cells (inhibition) and the appetite center in the brain (stimulates feeding behavior; see p. 625). The main role of activated AMPK in lipid metabolism is to increase fatty acid uptake into cells and their transport into mitochondria to be degraded by β-oxidation. The role of AMPK in lipolysis is unclear. Although AMPK has been observed in experiments to induce the inhibitory phosphorylation of HSL in adipocytes, fatty acid release from cells was observed.

TRANSCRIPTION FACTORS Changes in the long-term regulation of fatty acid metabolism, which occur in response to fluctuating nutrient availability and energy demand, are effected by alterations in gene expression. Two classes of transcription factors are prominent components of an intricate regulatory process: the SREBPs and peroxisome proliferator-activated receptors (PPARs). Each type of transcription factor, when activated, binds to a regulatory element near target genes, a process that triggers the binding of coactivator molecules and, subsequently, transcription.

The **SREBPs** are a group of three proteins (SREBP1a, SREBPlc, and SREBP2) that are coded for by two genes. SREBPla and (more prominently) SREBPlc regulate the expression of genes involved in fatty acid metabolism. (SREBPla, which activates all SREBP-responsive genes, is continuously expressed at low levels in most animal tissues.) SREBP2 regulates genes in cholesterol metabolism (see p. 483). In liver and adipose tissue, the activation of SREBP1c in response to insulin upregulates the transcription of the genes that code for enzymes in fatty acid synthetic pathway and NADPH synthesis. Glucagon and high levels of long-chain fatty acids inhibit SREBPlc. **PPARs**, named for the ability of certain synthetic compounds to cause the proliferation of liver cell peroxisomes, are ligand-activated transcription factors that bind to PPAR response elements associated with target genes. PPARα controls the expression of several genes in lipid metabolism. In liver and adipose tissues, under fasting conditions, it stimulates fatty acid catabolism and ketogenesis. PPARγ, primarily expressed in adipose tissue, in combination with insulin and SREBP1, promotes fat storage by stimulating glucose uptake and fatty acid and triacylglycerol synthesis. The activity of PPARs is stimulated by the binding of several lipid molecules (e.g., saturated and unsaturated fatty acids, prostaglandins, and leukotrienes).

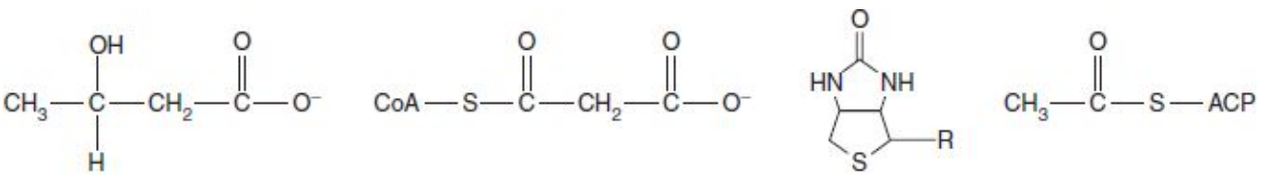
When there are high blood glucose levels, the glucose-responsive transcription factor ChREBP (p. 300) in both liver and adipose tissue is activated. ChREBP/Mlx heterodimer promotes the synthesis of enzymes that convert excess sugar molecules into fatty acids (e.g., ACC and fatty acid synthase).

KEY CONCEPTS

- The metabolism of fatty acids, the major energy source in animals, is regulated in the short term by allosteric modulators, covalent modification, and hormones.
- Long-term regulation, which occurs in response to fluctuating nutrient availability and energy demand, is effected by changes in gene expression.

QUESTION 12.9

Identify each of the following biomolecules:



What is the function of each?

Lipoprotein Metabolism: The Endogenous Pathway

The endogenous lipoprotein pathway, which transports recently synthesized lipids throughout the body, begins in the liver, where VLDLs are assembled on the cytoplasmic surface of hepatocyte ER. Nascent VLDLs contain apolipoprotein B-100, TGs, phospholipids, cholesterol, and cholesterol esters. Once VLDLs are secreted into blood, they are transformed into mature VLDL with the acquisition of apolipoproteins C-II and E, which are transferred from HDL. VLDLs then proceed to unload TGs as they encounter LPL (activated by apolipoprotein C-II) located primarily near the surface of target cells. Fatty acids transported into adipocytes are reconverted into TGs that coalesce into fat droplets. The fatty acids that are transported into muscle cells are oxidized to generate the energy required to drive muscle contraction. Once the TG content of VLDLs has been depleted, they are referred to as IDLs. The removal of IDLs from the blood by endocytosis is mediated by the binding of lipoprotein E to its receptor on the surface of hepatocytes. The TG content of IDLs is further reduced by hepatic lipase. Once the cholesterol content of IDLs exceeds that of TGs, the lipoproteins are referred to as LDLs. LDLs are released from the liver into the bloodstream, which carries them to target tissues. After the binding of apolipoprotein B-100 to LDL receptors, LDLs are internalized via endocytosis into cells, where they release their content, primarily cholesterol.

12.2 MEMBRANE LIPID METABOLISM

The lipid bilayer of cell membranes is composed primarily of phospholipids and sphingolipids. The metabolism of phospholipids is briefly described here. Sphingolipid metabolism is briefly described in an online reading: Sphingolipid Metabolism.

Phospholipid Metabolism

The membrane of the eukaryotic cell's endomembrane system (p. 45) originates in the SER with the synthesis of phospholipid at the interface of SER and cytoplasm. The fatty acid composition of the SER membrane subsequently changes, with unsaturated fatty acids replacing the original saturated

fatty acids. This remodeling, which is accomplished by phospholipases and acyltransferases, allows cells to adjust the fluidity of their membranes.

The syntheses of phosphatidylethanolamine (PE) and phosphatidylcholine (PC) are similar (**Figure 12.24**). PE synthesis begins in the cytoplasm when ethanolamine enters the cell and is immediately phosphorylated. Subsequently, phosphoethanolamine reacts with CTP (cytidine triphosphate) to form the activated intermediate CDP-ethanolamine. (CDP derivatives have an important role in the transfer of polar head groups in phosphoglyceride synthesis.) PE is the product of the reaction between CDP-ethanolamine with diacylglycerol, The biosynthesis of PC begins with the phosphorylation of choline, obtained in the diet. After the reaction of phosphocholine with CTP, the product, CDP-choline, reacts with diacylglycerol to form PC. PC is also synthesized in the liver from PE. PE is methylated in three steps by the enzyme phosphatidylethanolamine-*N*-methyltransferase to form the trimethylated product phosphatidylcholine. *S-Adenosylmethionine* (SAM) is the methyl donor in this set of reactions. (The role of SAM in cellular methylation processes is discussed in Chapter 14.)

In animals, phosphatidylserine (PS) is synthesized in the SER by the reaction of CDP-diacylglycerol with serine. It can also be produced by the reversible polar head group exchange with PE mediated by PE–serine transferase. The decarboxylation of PS to yield PE occurs in mitochondrial membranes.

Phospholipid turnover is rapid. (**Turnover** is the rate at which molecules in a structure are degraded and replaced with newly synthesized molecules.) For example, in animal cells, approximately two cell divisions are required for the replacement of one-half of the total number of phospholipid molecules. Phosphoglycerides are degraded by the phospholipases, each of which catalyzes the cleavage of a specific bond in phosphoglyceride molecules. Phospholipases A_1 and A_2 hydrolyze the ester bonds of phosphoglycerides at C-1 and C-2, respectively (see **Figure 11.11**, p. 412).

Biochemistry IN PERSPECTIVE

Atherosclerosis

What is the biochemical basis of arterial damage in the disease process called atherosclerosis? **Atherosclerosis** is a chronic disease in which soft masses called *atheromas* accumulate within arterial walls, eventually compromising their functional structure. Normal arterial walls are strong and flexible. They consist of three well-defined layers: the *intima* (a single layer of endothelial cells attached to an underlying extracellular matrix); the *media* (layers of smooth muscle cells embedded in an extracellular matrix consisting of elastic fibers, collagen, and proteoglycan); and the *adventitia* (the outermost layer, which consists of fibroblasts, smooth muscle cells, collagen, and elastin).

Once believed to be little more than passive conduits of blood, blood vessels are now known to be physiologically active. The endothelial cells that line arteries perform several vital functions. Among these functions are providing a barrier that prevents toxic substances from penetrating into the vessel wall and regulating the response of arteries to shear stress (the fluctuating mechanical force created by blood flow). Nitric oxide ($NO\cdot$, p. 178), the vasodilator produced by endothelial NO synthase, relaxes smooth muscle cells and inhibits their abnormal proliferation. Endothelial cells also produce molecules that provide arterial linings with smooth Teflon-like surfaces that prevent white blood cells

from adhering. In the course of aging, the endothelial barrier becomes leaky and vulnerable, a result that is accelerated by poor diet, diabetes, hypertension, smoking, and a sedentary lifestyle.

The atherosclerotic process is initiated by injury to the surface of endothelial cells, for example, AGE formation caused by high blood glucose levels (p. 256) or the binding of AGEs inhaled in tobacco smoke. Endothelial cells respond by secreting a type of cytokine (a class of signaling proteins) called chemokines (proteins that induce chemotaxis) that attract white blood cells to the injured site and by producing surface adhesion molecules that permit the binding of monocytes, cells that differentiate into macrophages. Monocytes/macrophages bind to AGEs via the AGE receptor (RAGE), a protein in the immunoglobulin superfamily. Oxidative stress ensues as the result of a RAGE-ligand-initiated signaling pathway. The inflammatory process causes a loosening of endothelial cell junctions, thereby allowing sdLDLs (p. 421) to enter the damaged site, where they become ensnared by interactions with proteoglycans. Trapped sdLDLs (small, dense LDLs, a high-risk factor for atherosclerosis, as opposed to lb [light buoyant] LDLs) are then engulfed by macrophages via endocytosis. Afterward these sdLDLs become depleted of antioxidant molecules, causing oxidative damage to accumulate. The attempt of macrophages to clear the site of oxidized LDL (oxLDL) is overwhelmed, and the phagocytes become so filled with lipoprotein that they transform into "foam cells." The oxLDLs within foam cells promote the release of growth factors such as platelet-derived growth factor (PDGF, p. 603) and fibroblast growth factor. These and similar molecules increase smooth muscle cell proliferation and migration and increased extracellular matrix component synthesis, which has the effect of reorganizing the vessel wall architecture. As the process progresses, activated macrophages secrete additional cytokines that recruit other inflammatory immune system cells, which in turn cause further inflammation. As the plaque thickens, it develops a necrotic core, caused in part by the decreasing capacity of macrophages to engulf and destroy dead cells. These dead cells break open, releasing other damaging molecules. In effect, the attempt to heal the injury results in the formation of a highly disorganized mass of accumulated fatty material, dead and dying cells, and scar tissue covered by a fibrous cap containing collagen and elastin, which walls off the damaged tissue.

The atherosclerotic process usually causes the formation of atheromas that extend sideways and not outward, which would cut off blood flow. Atherosclerotic lesions (*plaque*) usually do not cause obvious problems for long periods of time, perhaps decades. Eventually, the inflammatory process weakens the fibrous cap, which may suddenly rupture. It is the subsequent formation of a thrombus (blood clot) that prevents blood flow, especially in smaller vessels such as the coronary arteries. Sudden death, the most common symptom of coronary artery damage, results from such an event. In a small number of myocardial infarctions (heart attacks), plaque formation grows outward, slowly occluding blood flow. In these cases, decreased blood flow through one or more coronary arteries causes angina pectoris (pain and tightness in the chest) by preventing O_2 and nutrients from reaching the myocardial cells. Prompt medical attention to life-threatening occlusions can prevent death.

SUMMARY Atherosclerosis, which may lead to myocardial infarction, is initiated by damage to the endothelial cells that line arteries. The formation of atherosclerotic lesions begins with the accumulation of sdLDL and progresses to an inflammatory process that degrades arterial structure and function.

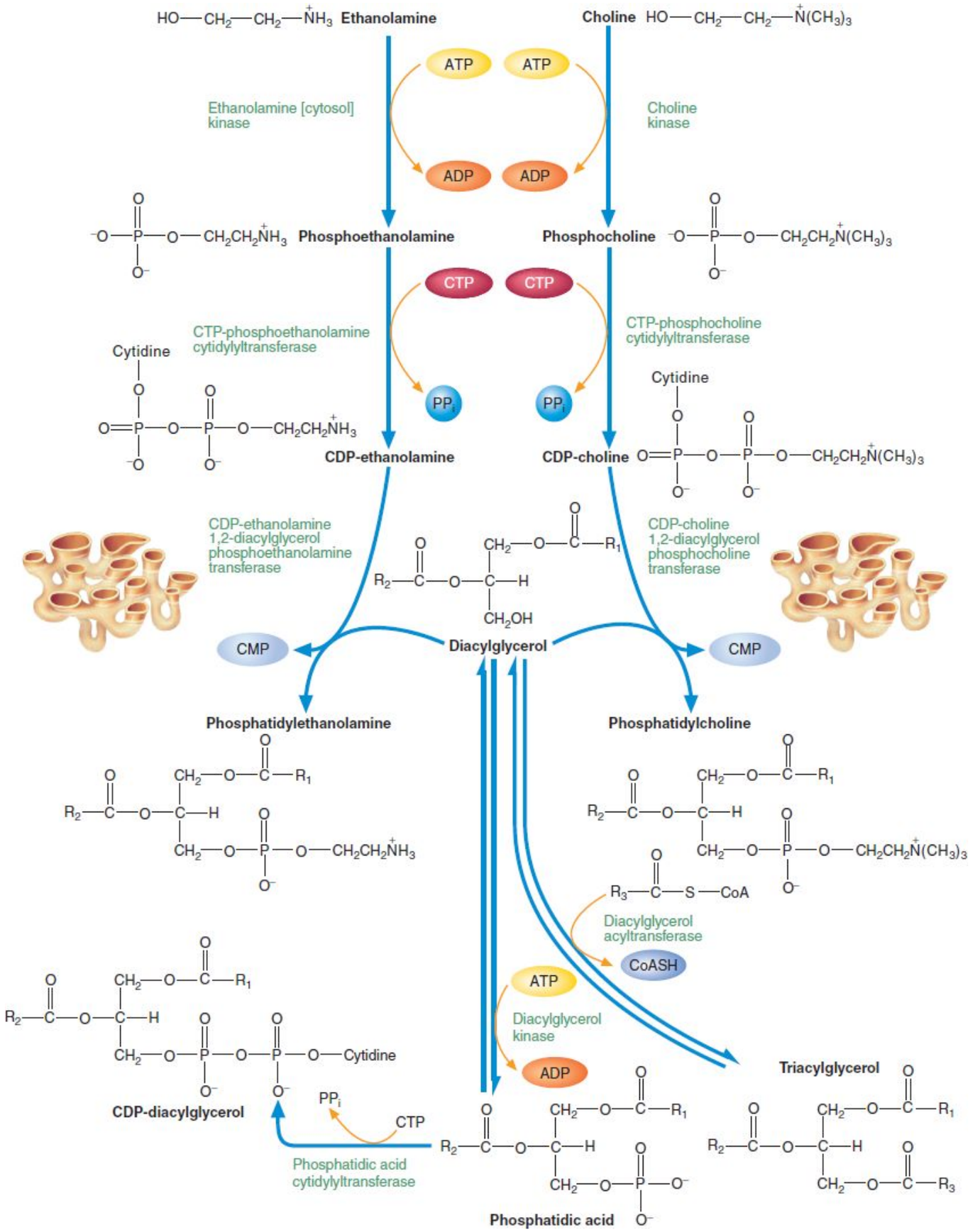


FIGURE 12.24

Phospholipid Synthesis

After ethanolamine or choline has entered a cell, it is phosphorylated and converted to a CDP derivative. Diacylglycerol then reacts with the CDP derivative, and phosphatidylethanolamine or phosphatidylcholine is formed. Triacylglycerol is produced if diacylglycerol reacts with acyl-CoA. CDP-diacylglycerol, formed from

phosphatidic acid and CTP, is a precursor of several phospholipids (e.g., phosphatidylglycerol and phosphatidylinositol).

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on sphingolipid metabolism.

KEY CONCEPTS

- Phospholipid synthesis occurs in the membrane of the SER. After phospholipids have been synthesized, they are remodeled by altering their fatty acid composition.
- Phospholipid degradation is catalyzed by several phospholipases.

12.3 ISOPRENOID METABOLISM

Isoprenoids occur in all eukaryotes. Despite the astonishing diversity of isoprenoid molecules, the mechanisms by which different species synthesize them are similar. In fact, the initial phase of isoprenoid synthesis (the synthesis of isopentenyl pyrophosphate) appears to be identical in all of the species in which this process has been investigated. **Figure 12.25** illustrates the relationships among the isoprenoid classes.

Because of its importance in human biology, cholesterol has received enormous attention from researchers. For this reason, the metabolism of cholesterol is better understood than that of any other isoprenoid molecule.

Cholesterol Metabolism

Cholesterol is derived from two sources: diet and de novo synthesis. When the diet provides sufficient cholesterol, usually about 400 mg per day, the synthesis of this molecule is depressed. In normal individuals, cholesterol delivered by LDL suppresses the synthesis of both cholesterol and LDL receptors. Cholesterol biosynthesis, which averages about 900 mg per day, and LDL receptor synthesis are stimulated when the diet is low in cholesterol. As described previously, cholesterol is a vital cell membrane component and a precursor in the synthesis of important metabolites. Cholesterol is also used to form bile salts.

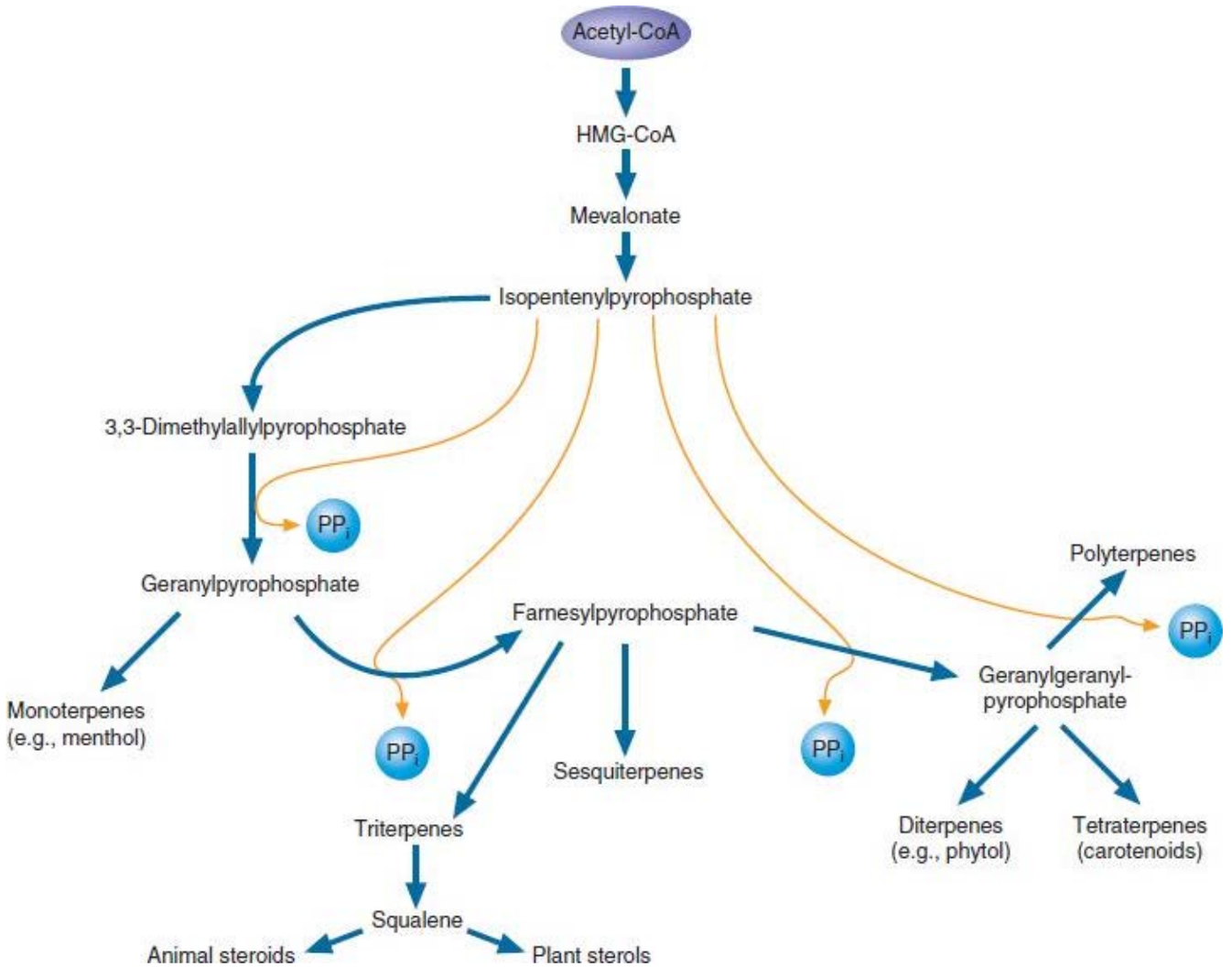


FIGURE 12.25

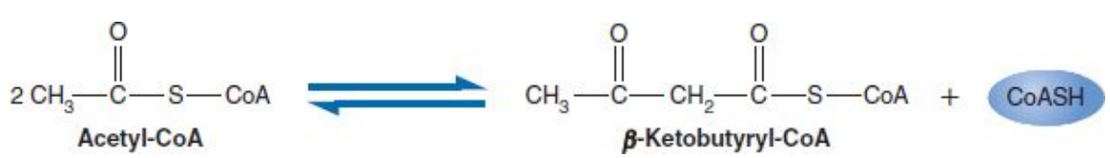
Isoprenoid Biosynthesis

Isoprenoid biosynthetic pathways produce an astonishing variety of products in different cell types and in different species. Despite this diversity, the beginning of isoprenoid biosynthesis appears to be identical in the species investigated (e.g., yeast, mammals, and plants). (HMG-CoA = *β*-hydroxy-*β*-methylglutaryl-CoA)

CHOLESTEROL SYNTHESIS Although all tissues (e.g., adrenal glands, ovaries, testes, skin, and intestine) can make cholesterol, most cholesterol molecules are synthesized in the liver. Cholesterol synthesis can be divided into three phases:

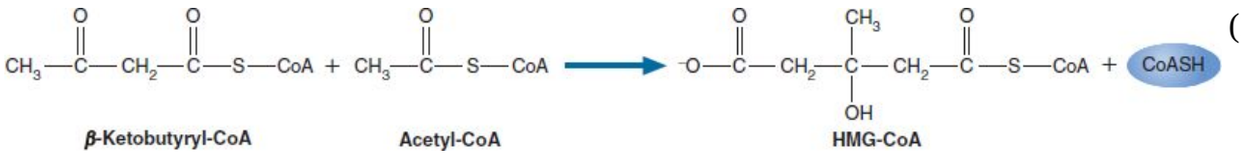
1. Formation of HMG-CoA (*β*-hydroxy-*β*-methylglutaryl-CoA) from acetyl-CoA
2. Conversion of HMG-CoA to squalene
3. Conversion of squalene to cholesterol.

The first phase of cholesterol synthesis is a cytoplasmic process. (Recall that the initial substrate, acetyl-CoA, is produced in mitochondria from fatty acids or pyruvate. Also observe the similarity of the first phase of cholesterol synthesis to ketone body synthesis. Refer to **Figure 12.9**.) The condensation of two acetyl-CoA molecules to form *β*-ketobutyryl-CoA (also referred to as acetoacetyl-CoA) is catalyzed by thiolase, also referred to as acetyl-CoA acetyl transferase.



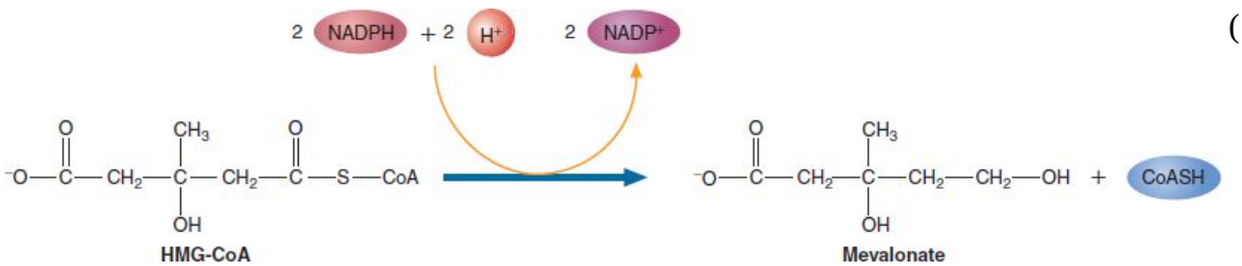
(9)

In the next reaction, *β*-ketobutyryl-CoA condenses with another acetyl-CoA to form *β*-hydroxy-*β*-methylglutaryl-CoA (HMG-CoA). This reaction is catalyzed by *β*-hydroxy-*β*-methylglutaryl-CoA synthase (HMG-CoA synthase).



(10)

The second phase of cholesterol synthesis begins with the reduction of HMG-CoA to form mevalonate. This reaction is catalyzed by HMG-CoA reductase (HMGR), the rate-limiting enzyme in cholesterol synthesis. NADPH is the reducing agent.

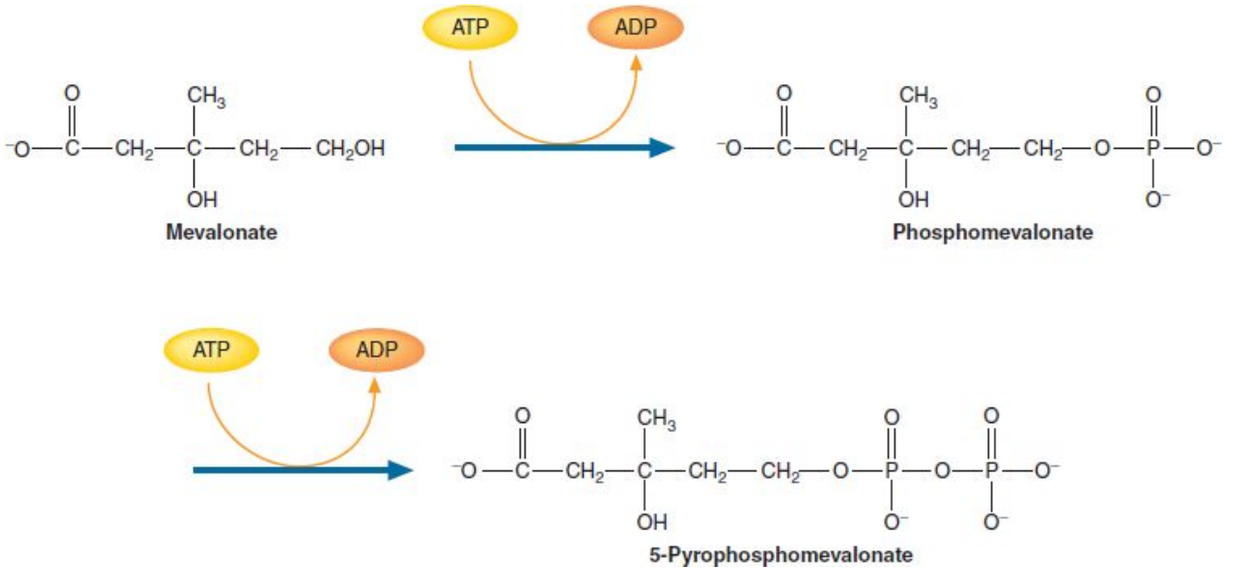


(11)

The HMGR polypeptide consists of three major domains: an N-terminal sterol-sensing transmembrane anchor domain, a catalytic domain, and a linker that connects the membrane and catalytic domains. The enzyme, located on the cytoplasmic face of the SER, consists of two HMGR dimers that associate to form a tetramer. Each dimer has an active site at the interface between the two dimers. The reaction (**Figure 12.26**) begins with the nucleophilic acyl substitution in which there is a hydride transfer from NADPH to the thioester carbonyl group of HMG-CoA. This transfer is assisted by a hydrogen bond between a lysine residue and the thioester carbonyl oxygen. The C—S bond in the product mevaloyl-CoA is hydrolyzed to form mevaldehyde. The CoA-thiolate anion product is protonated by a histidine residue and then released. Protonation of the carbonyl oxygen of mevaldehyde by the lysine residue facilitates a second hydride transfer from NADPH to form mevalonate.

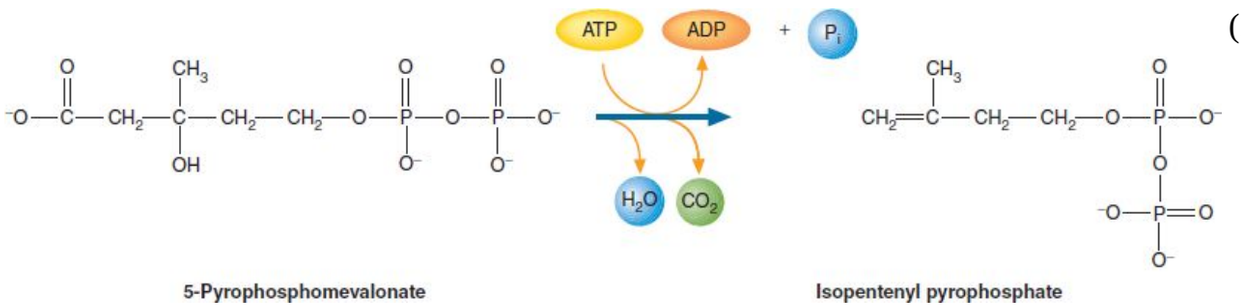
In a series of cytoplasmic reactions, mevalonate is converted to farnesylpyrophosphate. Mevalonate kinase catalyzes the synthesis of phosphomevalonate. A second phosphorylation reaction catalyzed by phosphomevalonate kinase produces 5-pyrophosphomevalonate.

(12)



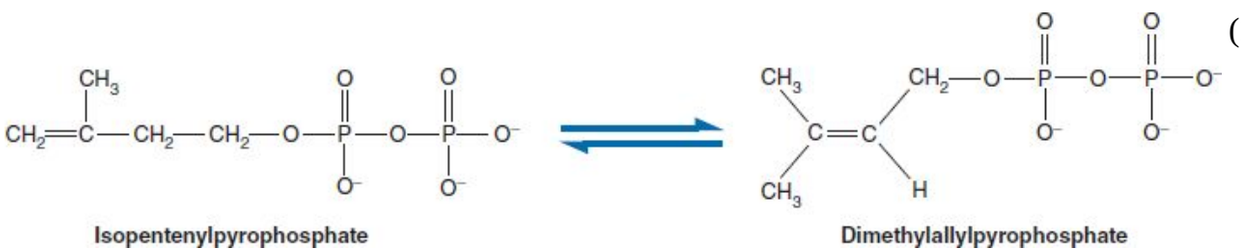
Phosphorylation reactions significantly increase the solubility of these hydrocarbon molecules in the cytoplasm.

5-Pyrophosphomevalonate is converted to isopentenylpyrophosphate in a process involving a decarboxylation and a dehydration:



(13)

Isopentenyl pyrophosphate is next converted to its isomer dimethylallylpyrophosphate by isopentenylpyrophosphate isomerase. (A CH_2=CH—CH_2— group on an organic molecule is sometimes referred to as an *allyl group*.)



(14)

Geranylpyrophosphate is generated during a condensation reaction between isopentenylpyrophosphate and dimethylallylpyrophosphate (**Figure 12.27**). Pyrophosphate is also a product of this reaction and two subsequent reactions. (Recall that reactions in which pyrophosphate is released are irreversible because of subsequent pyrophosphate hydrolysis.) Geranyl transferase catalyzes the head-to-tail condensation reaction between geranylpyrophosphate and isopentenylpyrophosphate that forms farnesylpyrophosphate. Squalene is synthesized when farnesyl transferase catalyzes the head-to-head condensation of two farnesylpyrophosphate molecules. (Farnesyl transferase is sometimes referred to as squalene synthase.) This reaction requires NADPH as an electron donor.

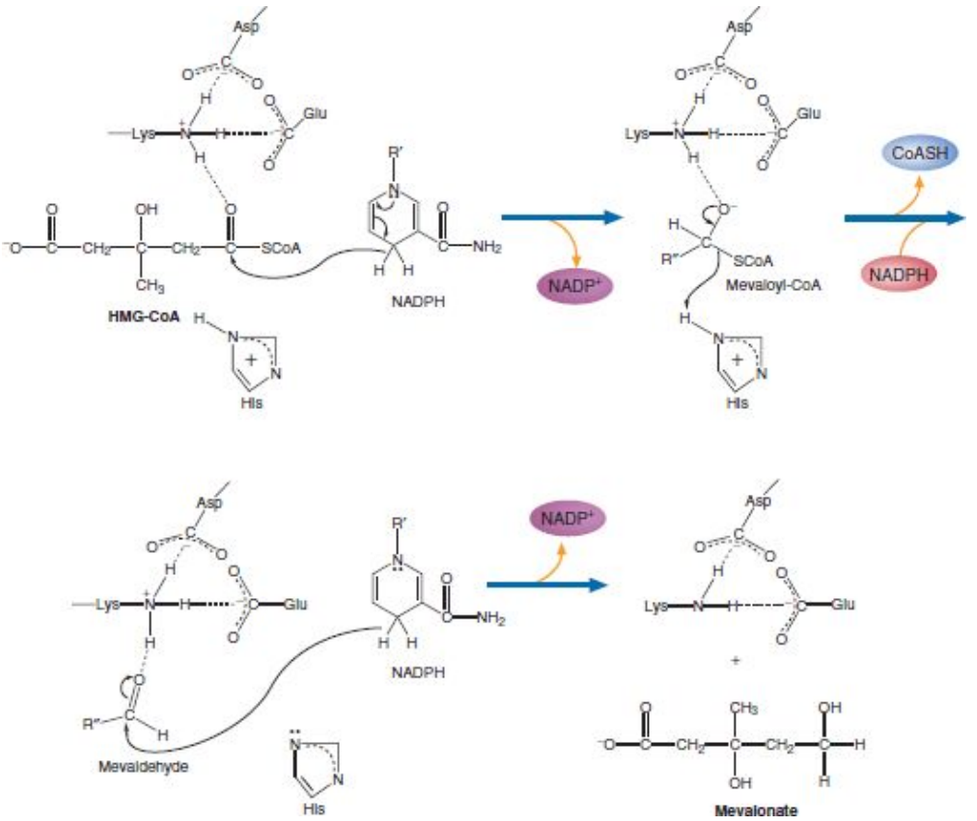


FIGURE 12.26

The HMGR-Catalyzed Reaction

In the HMGR-catalyzed reaction, there is an initial hydride transfer from NADPH to the thioester carbonyl of the substrate HMG-CoA. Subsequently, the C—S bond of mevaloyl-CoA is hydrolyzed to form mevaldehyde, and a histidine residue protonates the CoA-thiolate anion. A second hydride transfer from NADPH, facilitated by the protonation of the carbonyl oxygen of mevaldehyde, results in the formation of mevalonate, the product of the reaction. Note that carboxylate side chain groups of Asp and Glu residues orient the Lys side chain amino group within the active site.

The last phase of the cholesterol biosynthetic pathway (**Figure 12.28**) begins by binding squalene to a cytoplasmic protein called **sterol carrier protein**. The conversion of squalene to lanosterol occurs while the intermediates are bound to this protein. The enzyme activities required for the oxygen-dependent epoxide formation (squalene monooxygenase) and subsequent cyclization (2,3-oxidosqualene lanosterol cyclase) that result in lanosterol synthesis have been localized in microsomes (p. 68). The flavoprotein squalene monooxygenase requires NADPH and FAD for activity. After its synthesis, lanosterol binds to a second carrier protein, to which it remains attached during the remaining reactions. All of the enzyme activities that catalyze the remaining 20 reactions needed to convert lanosterol to cholesterol are embedded in SER membranes. In a series of transformations involving NADPH and oxygen, lanosterol is converted to 7-dehydrocholesterol. This product is then reduced by NADPH to form cholesterol.

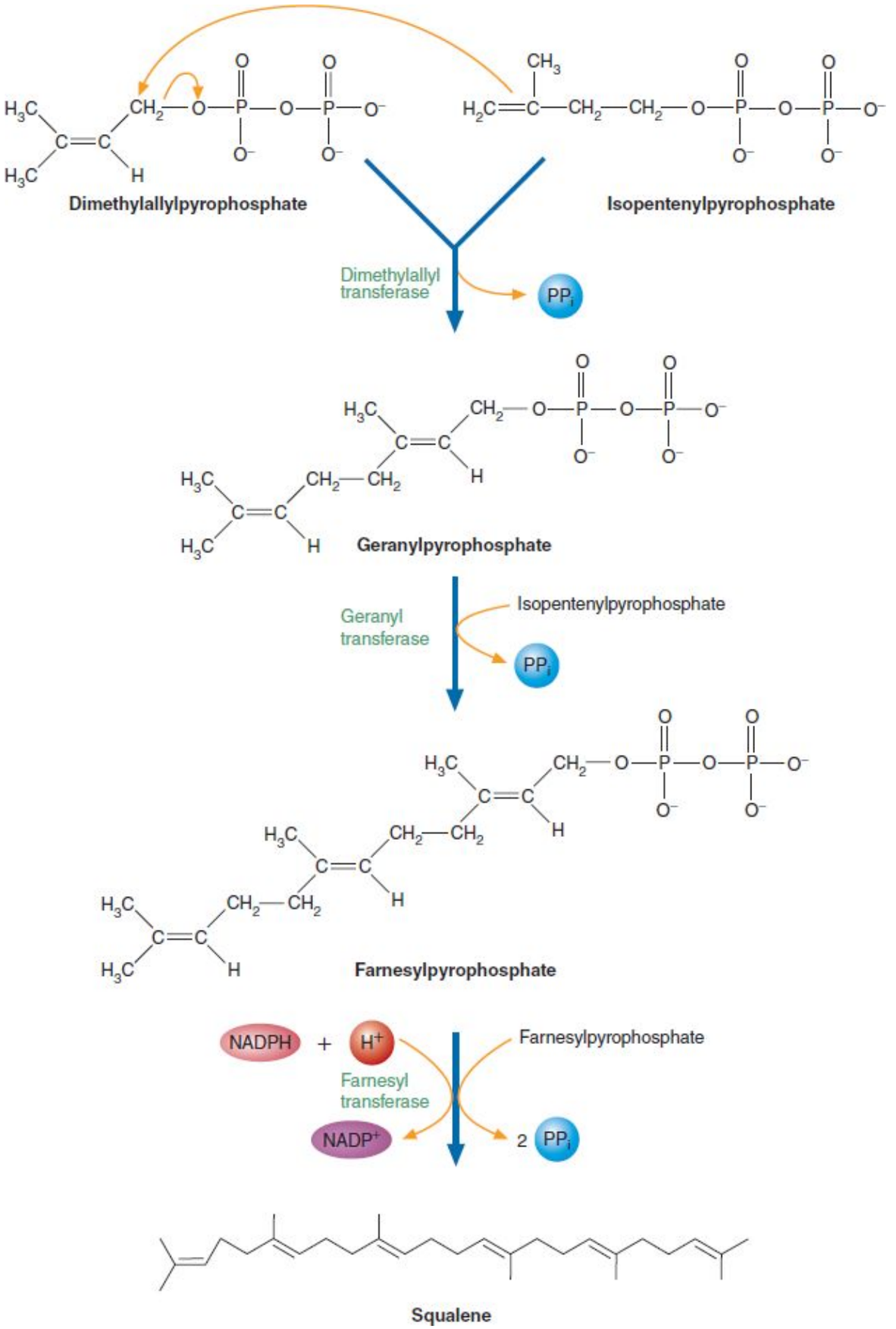


FIGURE 12.27

Synthesis of Squalene

The head-to-tail condensation of dimethylallylpyrophosphate and isopentenylpyrophosphate generates the terpene geranylpyrophosphate. A subsequent head-to-tail condensation with another isopentenylpyrophosphate generates the C farnesylpyrophosphate. Head-to-head condensation of two molecules of farnesylpyrophosphate produces

the C_{30} triterpene, squalene. The geranylgeranyl groups used in prenylation reactions are synthesized in a reaction between farnesyl pyrophosphate and isopentenyl pyrophosphate.

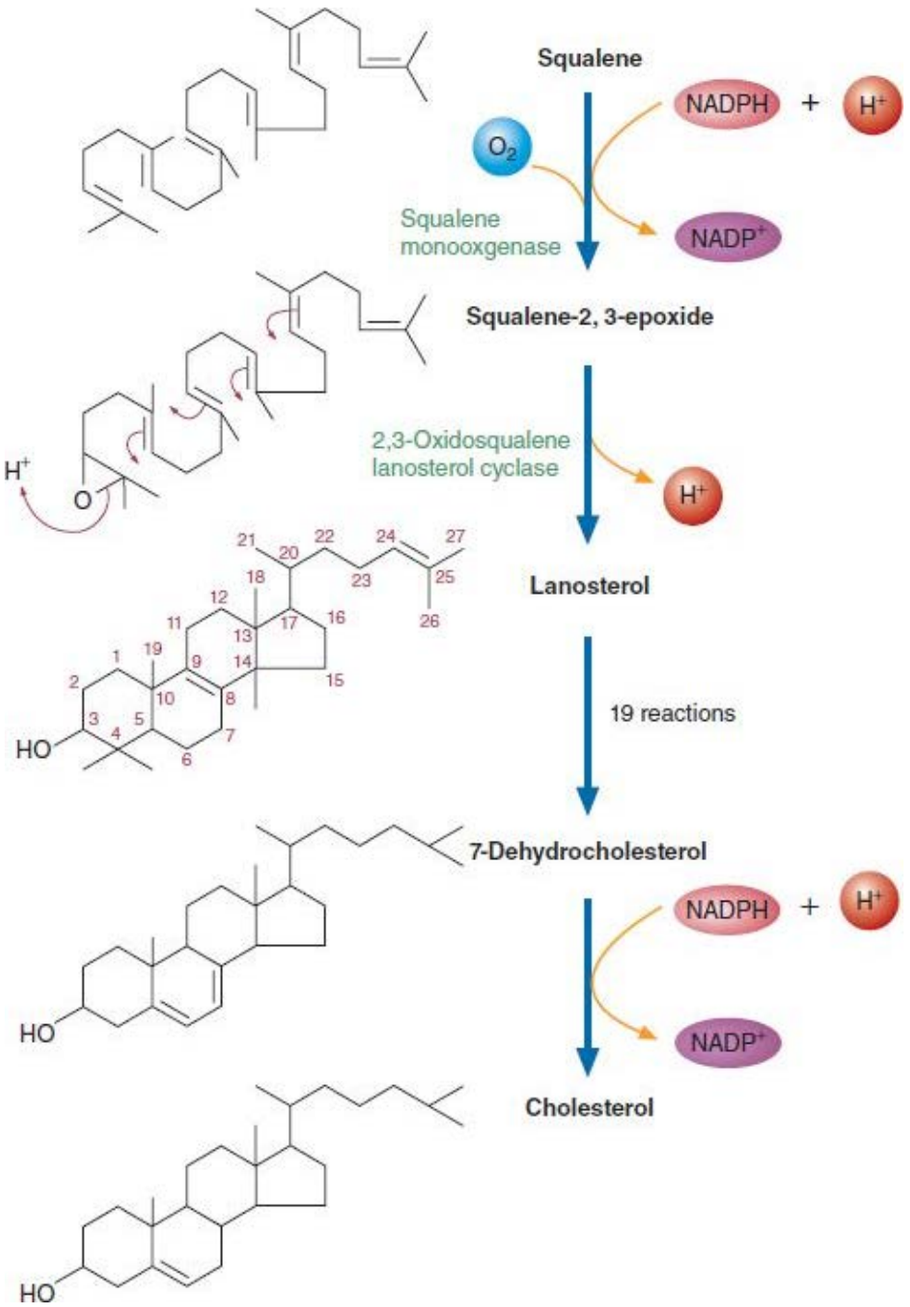


FIGURE 12.28

Synthesis of Cholesterol from Squalene

This is the major route in mammals. In an alternative minor route, squalene is converted to desmosterol, which is then reduced to form cholesterol. The details of these and many reactions in the major route are poorly understood. (Desmosterol differs from cholesterol because of a C=C bond between C-24 and C-25.)

CHOLESTEROL DEGRADATION Unlike many other types of biomolecules, cholesterol and other steroids cannot be degraded to smaller molecules. Instead, downregulation of synthesis along with loss caused by bile acid synthesis and excretion and steroid hormone biotransformation combine to lower circulating cholesterol levels. About one-half of the cholesterol synthesized daily is used to produce bile acids, which occurs in the liver. The synthesis of cholic acid, one of the principal bile acids, is outlined in **Figure 12.29**. The conversion of cholesterol to 7-α-hydroxycholesterol,

catalyzed by cholesterol-7-hydroxylase (a SER enzyme), is the rate-limiting reaction in bile acid synthesis. Cholesterol-7-hydroxylase is a cytochrome P_{450} enzyme (pp. 485–86). In later reactions, the double bond at C-5 is rearranged and reduced, and an additional hydroxyl group is introduced. The products of this process, cholic acid and deoxycholic acid, are converted to bile salts by SER enzymes that catalyze conjugation reactions. (In **conjugation reactions,** a molecule's solubility is increased by converting it into a derivative that contains a water-soluble group. Amides and esters are common examples of these conjugated derivatives.) Most bile acids are conjugated with glycine or taurine ($H_3N^+CH_2CH_2SO^-_3$).

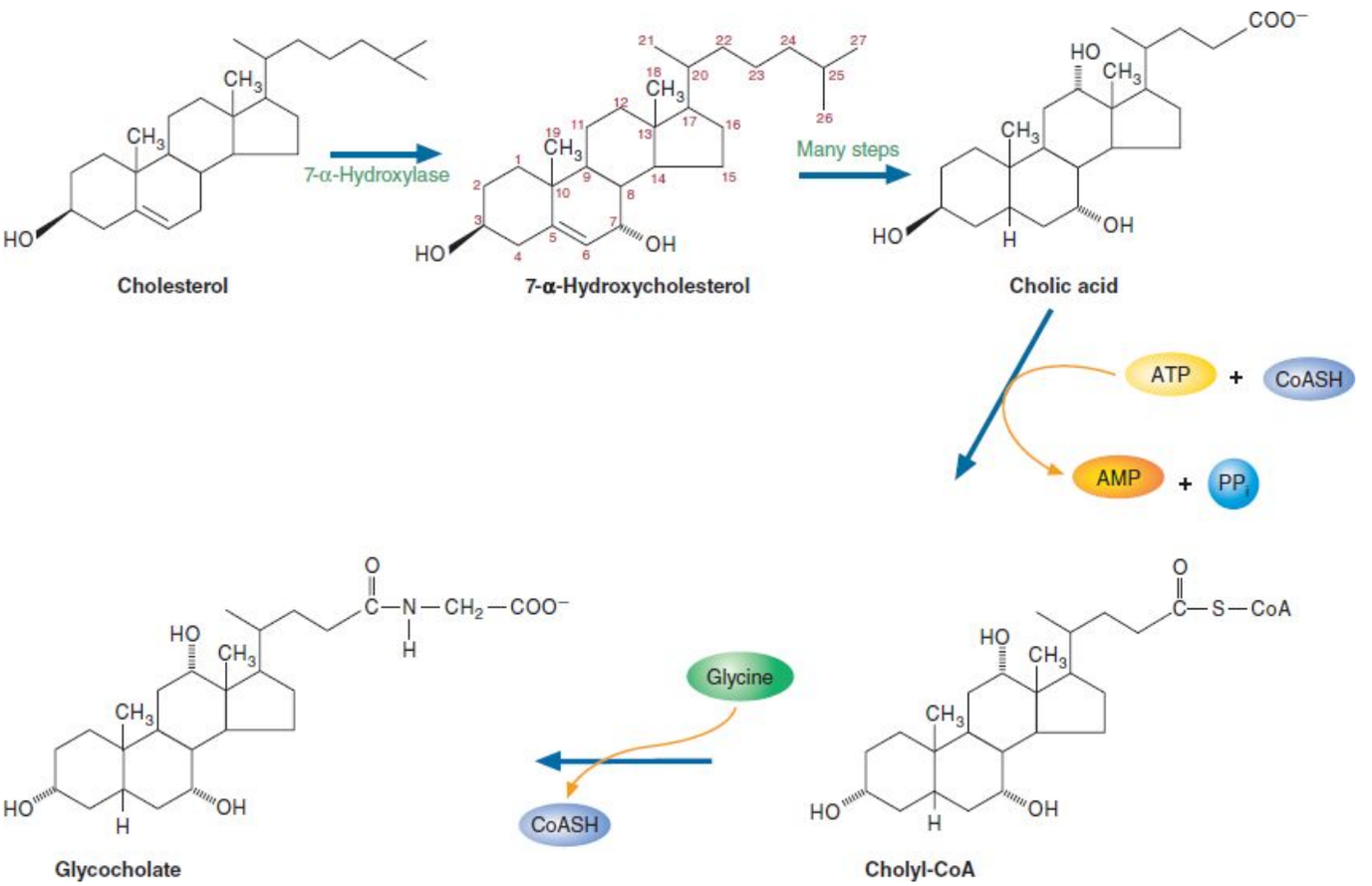


FIGURE 12.29

Bile Salt Synthesis

Bile salts are emulsifying agents that facilitate dietary fat digestion in the small intestine. They are synthesized in the liver from cholesterol. The bile acid cholic acid (cholate) is produced from cholesterol in a series of reactions, two of which are hydroxylation reactions catalyzed by cytochrome P_{450} enzymes: 7-*α*-hydroxylase and 12-*α*-hydroxylase (not shown). The bile salt glycocholate is produced when cholyl-CoA reacts with glycine to form an amide bond linkage.

The bile salts are important components of *bile*, a yellowish-green liquid produced by hepatocytes that aids in the digestion of lipids. In addition to bile salts, bile contains cholesterol, phospholipids, and bile pigments (bilirubin and biliverdin), degradation products of heme. After bile is secreted into the bile ducts and stored in the gall bladder, it is used in the small intestine as an emulsifying agent to form biliary micelles and enhance the absorption of dietary fat and fat-soluble vitamins (A, D, E, and K). Most bile salts (about 90%) are reabsorbed in the distal ilium (near the end of the small intestine). They enter the blood and are transported back to the liver, where they are resecreted into the bile ducts with other bile components. Bile acid conjugation reactions prevent premature absorption of bile acids in the biliary tract (the duct system and gallbladder) and small intestine. The

reabsorption of bile salts in the distal ilium of the small intestine (necessary for effective recycling) is apparently triggered by the glycine or taurine signal. (It has been estimated that bile salt molecules are recycled 18 times before they are finally eliminated.)

Gallstones

QUESTION 12.10

The formation in the gallbladder or bile ducts of gallstones (crystals usually composed of cholesterol and inorganic salts) afflicts millions of people. Predisposing factors for this excruciatingly painful disorder include obesity and infection of the gallbladder (*cholecystitis*). Because cholesterol is virtually insoluble in water, it is solubilized in bile by its incorporation into micelles composed of bile salts and phospholipids. Gallstones tend to form when cholesterol is secreted into bile in excessive quantities. Suggest a reason why an obese person is prone to gallstone formation. [*Hint*: HMG-CoA reductase activity is higher in obese individuals.]

CHOLESTEROL HOMEOSTASIS The critical roles of cholesterol in animal bodies, combined with the potentially toxic properties it exhibits when present in excessive amounts, require that its concentration be maintained within narrow limits. Cholesterol homeostasis is achieved through intricate mechanisms that regulate its biosynthetic pathway, LDL receptor activity, and bile acid biosynthesis. The regulation of cholesterol biosynthesis is accomplished primarily by modulation of existing HMGR molecules, gene expression changes, and enzyme degradation.

The principal means by which the activity of existing HMGR molecules is regulated is downregulation via phosphorylation reactions (**Figure 12.30**). HMGR activity is depressed by phosphorylation by AMPK in response to high cellular concentrations of AMP, which have the effect of integrating cholesterol biosynthesis, a metabolically expensive process, into the cell's energy metabolism. cAMP, which is regulated by hormones such as glucagon and epinephrine, also depresses HMGR by activating phosphoprotein phosphatase inhibitor 1 (PPI-1) by means of a phosphorylation reaction catalyzed by PKA (protein kinase A; see p. 607). Activated PPI-1 inhibits several phosphatases, including PP-2A, which can increase HMGR activity by removing phosphate groups. Insulin increases HMGR activity, in part by inhibiting cAMP synthesis. HMGR is also regulated by a negative feedback mechanism involving various sterols, including cholesterol derived from the endocytosis of LDL receptors, and nonsterol derivatives of mevalonate.

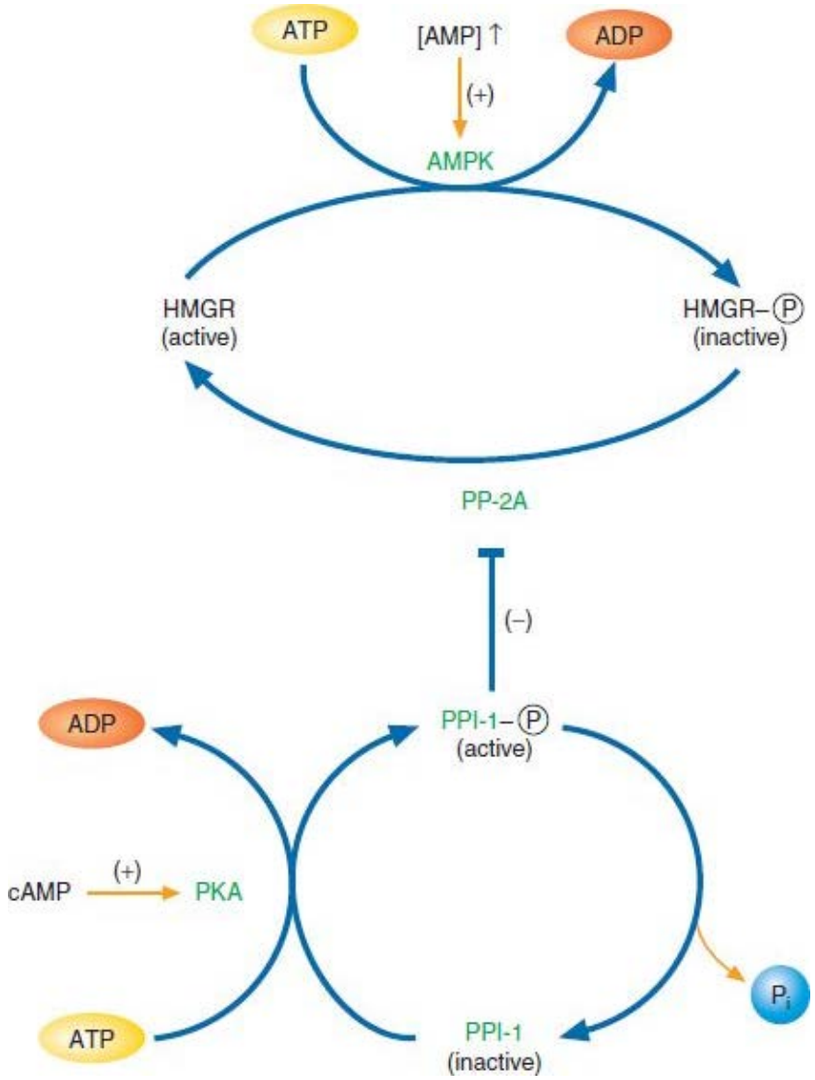


FIGURE 12.30

Regulation of HMGR by Covalent Modification

HMGR is inactivated by phosphorylation reactions catalyzed by AMPK in response to rising AMP levels. Activation of HMGR is effected by PP-2A. When cAMP levels are high, phosphoprotein phosphatase inhibitor 1 (PPI-1), activated by a PKA-catalyzed phosphorylation reaction, inhibits PP-2A. An inhibited PP-2A ensures the inactive status of HMGR.

Sterol-mediated changes in gene expression are a major feature of cholesterol homeostasis. The ER membrane protein SREBP2 is the predominant regulator of cholesterol biosynthesis. In addition to stimulating the expression of cholesterol biosynthesis genes, SREBP2 activates the LDL receptor gene and three genes required for NADPH synthesis (G-6-PD, 6-phosphogluconate dehydrogenase, and malic enzyme). SREBP2 is bound to SREBP cleavage-activating protein (SCAP). When cells have sufficient sterols (cholesterol and other steroid alcohols), the sterol-sensing domain (SSD) of SCAP binds cholesterol and an ER retention protein called Insig (*ins*ulin-*i*nduced *g*ene). In cholesterol-depleted cells, Insig no longer binds to SCAP, which now escorts SREBP-2 from the ER to the Golgi (**Figure 12.31**), where two proteases release the N-terminal domain of SREBP2, the active transcription factor, from the membrane. The activated SREBP2 transcription factor then translocates to the nucleus, where it binds to the SREs (sterol regulatory elements) of the target genes. In addition to SREBP, the expression of some target genes requires the binding of coregulatory transcription factors. For example, transcription of the genes for HMGR and HMG-CoA synthase also requires the binding of nuclear factor-1 and CREB (cAMP response element binding protein).

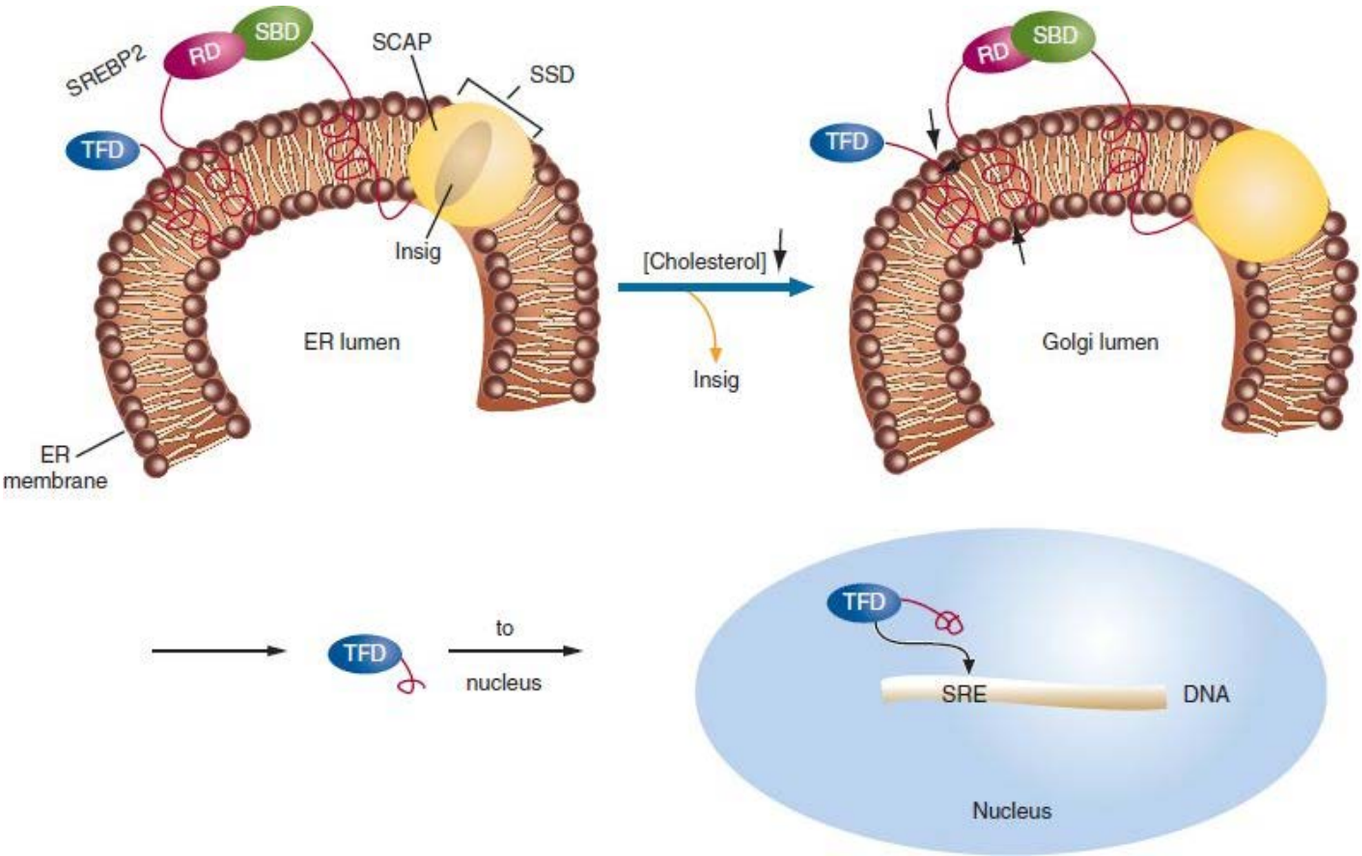


FIGURE 12.31

SREBP2 Regulation

SREBP2, an ER protein, is complexed to SCAP (SREBP cleavage-activating protein) via the binding of the regulatory domain (RD) of SREBP2 with the SREBP-binding domain (SBD) of SCAP. The SSD (sterol-sensing domain) of SCAP binds cholesterol. When cholesterol levels are high and cholesterol is bound to SSD, the ER retention protein Insig is also bound. When cholesterol levels are depleted and cholesterol is no longer bound to SSD, Insig is released, and the SREBP/SCAP complex is transferred to the Golgi complex. When SREBP2 is present in the Golgi complex, two proteases cleave it at two sites (arrows) to release the now active TFD (transcription factor domain). The SREBP transcription factor then moves into the nucleus, where it binds to SREs (sterol regulatory elements) that are associated with sterol-related genes.

As cellular cholesterol levels begin to rise as a result of newly synthesized enzymes and the internalization of LDL from the bloodstream via LDL receptors, a concentration threshold is reached and a few molecules bind to the sterol-sensing domains of SCAP. As the result of a conformation change in the SCAP-SREBP complex, the ER protein Insig replaces cholesterol in the sterol-sensing domain of SCAP and the complex is retained in the ER, thus ending SREBP2 processing and the transcription of target genes. High levels of cholesterol and other metabolites of mevalonate also depress further cholesterol synthesis by inhibiting the translation of existing HMGR mRNA. In liver, excess cholesterol activates ACAT (p. 418), the enzyme that catalyzes the transfer of a fatty acid from a fatty acyl-CoA molecule to the hydroxyl group of cholesterol to generate a cholesterol ester storage molecule.

HMGR degradation, another means by which cholesterol levels are controlled, is mediated by Insig. When cholesterol levels are high, sterols bind to the enzyme's sterol-sensing N-terminal domain. Subsequently, HMGR binds to Insig, which in turn is associated with ubiquitin ligase, an enzyme that initiates a major proteolytic mechanism (described in Chapter 15).

High cholesterol concentrations in the liver also trigger bile acid biosynthesis. When cholesterol begins to accumulate, some molecules become oxidized to form oxysterols (e.g., 25-

hydroxycholesterol). Oxysterols bind to and activate LXR (liver X receptor), which then forms an active heterodimer transcription factor with RXR (retinol X receptor). The heterodimer then causes the transcription of 7-α-hydroxylase, the rate-limiting enzyme in bile synthesis.

KEY CONCEPTS

- Cholesterol is synthesized from acetyl-CoA in a multistep pathway that occurs primarily in the liver.
- Cholesterol is degraded primarily by conversion to the bile salts, which facilitate the emulsification and absorption of dietary fat.

The Cholesterol Biosynthetic Pathway and Drug Therapy

A majority of medical researchers believe that high blood serum levels of total cholesterol (the sum of cholesterol in VLDL, LDL, and HDL) combined with high LDL levels are strongly associated with cardiovascular disease. Currently, a class of drugs called the *statins* is routinely used to lower serum cholesterol in an attempt to reduce the risk of heart attack and stroke. Statins such as lovastatin are competitive inhibitors of HMG-CoA reductase, the rate-limiting enzyme in cholesterol biosynthesis. Because most cholesterol synthesis occurs in the liver, cholesterol serum levels are reduced when liver cells compensate for reduced synthesis by removing cholesterol from serum LDLs. Because most of the body's cholesterol is synthesized at night, statins are usually taken in the evening. Statin drugs also interfere with the synthesis of the mixed terpenoid coenzyme Q (ubiquinone), a critical molecule in energy generation. Consequently, statin therapy must be accompanied by CoQ supplementation.

Cholesterol and Heart Disease

Bisphosphonates are a class of drugs used to treat *osteoporosis*, a bone disease in which bone mineral density is reduced, and malignant bone diseases. Because bisphosphonates bind Ca^{2+}, they are readily absorbed into bone tissue where they kill osteoclasts. *Osteoclasts* are bone-remodeling cells that break down and remove old bone tissue. Bisphosphonates such as alendronate trigger osteoclast apoptosis by inhibiting the enzymatic activities that convert isopentenylpyrophosphate to farnesylpyrophosphate (see **Figure 12.25**). Osteoclast cell death occurs as a result of the absence of geranylpyrophosphate and farnesylpyrophosphate, the substrates for the prenylation reactions that link several cell signal proteins (e.g., Ras) to the plasma membrane.

Biochemistry IN PERSPECTIVE

Biotransformation

How does the body metabolize potentially toxic hydrophobic molecules? In **biotransformation**, a series of enzyme-catalyzed processes in which toxic substances are converted into less toxic metabolites, the enzymes generally possess broad specificities. In mammals,

biotransformation is used principally to convert toxic molecules, which are usually hydrophobic, into water-soluble derivatives so that they may be more easily excreted. The enzymes that catalyze the biotransformation of xenobiotics (foreign molecules such as drugs or organic pollutants) are similar to several of the enzymes that dispose of hydrophobic endogenous molecules. Although biotransformation reactions occur in several locations within the cell (e.g., the cytoplasm and mitochondria), most occur within the SER. Cell types also differ in their biotransforming potential. In general, cells located near the major points of xenobiotic entry into the body (e.g., liver, lung, and intestine) possess greater concentrations of biotransforming enzymes than others.

Biotransformation processes have been differentiated into two major types. During **phase I**, oxidative and/or hydrolytic reactions involving oxidoreductases and hydrolases convert hydrophobic substances into more polar molecules that are usually less biologically active. **Phase II** consists of reactions in which metabolites containing appropriate functional groups are conjugated (covalently bonded) with substances such as glucoronate or sulfate by enzymes such as glucuronosyl transferase or sulfotransferase. In general, conjugation dramatically improves solubility, which then promotes rapid excretion. (In animals, excretion of biotransformed molecules is sometimes referred to as phase III.) Although many substances undergo these phases sequentially, a significant number do not. For example, some molecules are excreted as phase I metabolites, whereas others undergo only phase II reactions. Moreover, variations in enzyme concentrations, availability of cosubstrates, and the order in which the reactions occur may cause certain substances to be converted into more than one end product. However, despite these and other complications, basic biotransformation patterns have emerged. Several well-researched examples of phase I and phase II reactions are described below. In the following discussions, the term **detoxication** refers to the process by which a toxic molecule is converted to a more soluble (and usually less toxic) product. The more familiar term **detoxification** implies the correction of a state of toxicity, that is, the chemical reactions that produce sobriety in an intoxicated person.

Phase I reactions usually convert substrates to more polar forms by introducing or unmasking a functional group (e.g., —OH, —NH_2, or —SH). Many phase I enzymes are located in the SER membrane, but others, such as the dehydrogenases (e.g., alcohol dehydrogenases and peroxidases), occur in the cytoplasm, whereas still others (e.g., monoamine oxidase) are localized in mitochondria. The predominant enzymes of SER oxidative metabolism are the monooxygenases, sometimes referred to as *mixed-function oxidases*. They are so named because, in a typical reaction, one molecule of oxygen is consumed (reduced) per substrate molecule: one oxygen atom appearing in the product and the other in a molecule of water. Monooxygenases can carry out an immense variety of chemical reactions. Some of these reactions form highly unstable (and therefore toxic) intermediates.

There are two major types of SER monooxygenase, both of which require NADPH as an external reductant: the cytochrome P_{450} (cyt P_{450}) system and flavin-containing monooxygenases. The **cytochrome P_{450} system**, which consists of two enzymes (NADPH–cytochrome P_{450} reductase and cytochrome P_{450}), is involved in the oxidative metabolism of many endogenous substances (e.g., steroids and bile acids), as well as the detoxication of a wide variety of xenobiotics. **Flavin-containing monooxygenases** catalyze an NADPH- and an oxygen-requiring oxidation of substances (primarily xenobiotics) bearing functional groups containing nitrogen, sulfur, or phosphorus. The properties of the cyt P_{450} electron transport systems are described next.

Cytochrome P_{450} Electron Transport Systems

In cytochrome P_{450} electron transport systems, found in SER and inner mitochondrial membranes, two electrons are transferred one at a time from NADPH to a cytochrome P_{450} protein by NADPH–cytochrome P_{450} reductase (**Figure 12A**). The latter enzyme is a flavoprotein that contains both FAD and FMN in a ratio of 1:1 per mole of enzyme. In addition to its role in the cytochrome P_{450} system,

the reductase is also believed to be involved in the function of heme oxygenase. (See the Chapter 15 online Biochemistry in Perspective essay Heme Biotransformation.)

The hemoproteins referred to as cytochrome P_{450} are so named because of the complexes they form with carbon monoxide. In the presence of the gas, light is strongly absorbed at a wavelength of 450 nm. More than 6000 *cyt* P_{450} genes have been identified in species as diverse as mammals and bacteria. In humans, 57 *cyt* P_{450} genes, classified into 18 families, have been identified. Each gene codes for a protein with a unique specificity range. Cytochrome P_{450} proteins found in the liver have broad and overlapping specificities. For example, molecules as diverse as alkanes, aromatics, ethers, and sulfides are routinely oxidized. In contrast, cytochrome P_{450} proteins in the adrenal glands, ovaries, and testes that add hydroxyl groups to intermediates during steroid synthesis have narrow specificities. Despite this diversity, all cytochrome P_{450} isozymes contain one molecule of heme and are similar in their physical properties and catalytic mechanisms.

Despite an enormous variety of substrates, all of the oxidative reactions catalyzed by cytochrome P_{450} may be viewed as hydroxylation reactions (i.e., an OH group appears in each reaction) (**Figure 12B**). The general reaction is as follows:

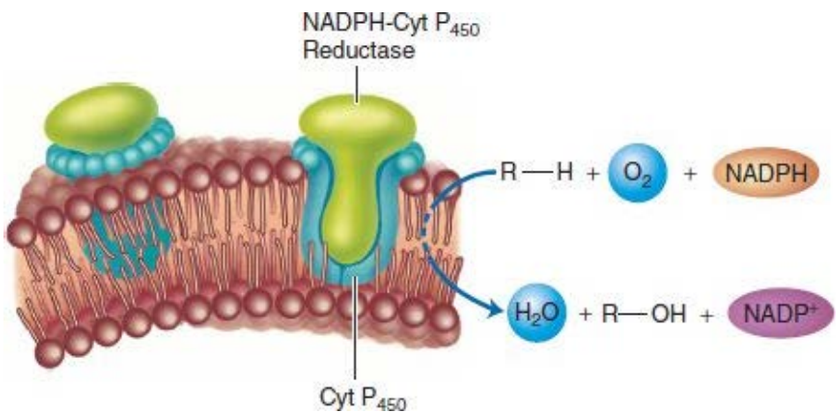


FIGURE 12A

The Cytochrome P_{450} Electron Transport System

Cytochrome P_{450} and cytochrome P_{450} reductase are components of an electron transport system used to oxidize both endogenous and exogenous molecules.

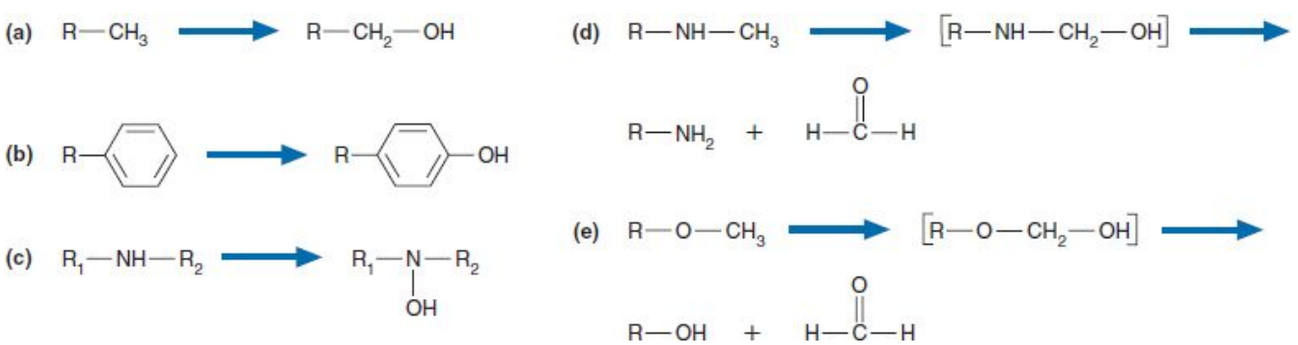


FIGURE 12B

Diverse Substrates Oxidized by Cytochrome P_{450} Isozymes

Among the reactions catalyzed by cytochrome P_{450} are (a) aliphatic oxidation, (b) aromatic hydroxylation, (c) N-hydroxylation, (d) N-dealkylation, and (e) O-dealkylation.

$$R—H + O_2 + NADPH + H^+ \rightarrow ROH + H_2O + NADP^+$$

where R—H is the substrate.

The oxygenation reaction is initiated when the substrate binds to oxidized cytochrome P_{450} (Fe^{3+}). This binding promotes a reduction of the enzyme substrate complex by an electron transferred from NADPH via cytochrome P_{450} reductase (Fe^{2+}—substrate). After reduction, cytochrome P_{450} can bind O_2. Then the electron from heme iron is transferred to the bound O_2, thus forming a transient Fe^{3+}—O_2^-—substrate species. (If the bound substrate is easily oxidized, it can be converted into a peroxy radical.) A second electron transferred from the flavoprotein results in the generation of a Fe^{3+}—O_2^{2-}—substrate complex. This brief association ends when the oxygen–oxygen bond is broken.

One atom is released in a water molecule, while the other remains bound to heme. After a hydrogen atom or electron has been abstracted from the substrate, the oxygen species (now a powerful oxidant) is transferred to the substrate. The cycle ends with release of the product from the active site. Depending on the nature of the substrate, the product is either an **epoxide** (a highly reactive ether in which the oxygen is incorporated into a three-membered ring) or an alcohol.

The function of conjugation (phase II) reactions is to inactivate biologically active substances and/or to form more polar (and therefore more easily excretable) derivatives. During this process, lipophilic metabolites, bearing functional groups that can act as acceptors, undergo enzyme-catalyzed reactions along with second (or donor) substrates. Among the most frequently used donor substrates are glucuronic acid, glutathione (see later, Section 14.3), sulfate, and amino acids, such as glycine.

SUMMARY Biotransformation is the enzyme-catalyzed process in which toxic, hydrophobic molecules are converted to less toxic, water-soluble molecules.

Chapter Summary

1. Acetyl-CoA plays a central role in most lipid-related metabolic processes. For example, acetyl-CoA is used in the synthesis of fatty acids. When fatty acids are degraded to generate energy, acetyl-CoA is the product.
2. In the exogenous lipoprotein pathway, dietary lipid is distributed to the body's tissues. The process begins within enterocytes where TGs, cholesterol, and other lipid molecules and apolipoprotein B-48 are packaged into chylomicrons.
3. Depending on the body's current energy requirements, newly digested fat molecules are used to generate energy or are stored within adipocytes. When the body's energy reserves are low, fat stores are mobilized in a process referred to as lipolysis in which triacylglycerols are hydrolyzed to fatty acids and glycerol. Glycerol is transported to the liver, where it can be used in glucose synthesis. Most fatty acids are degraded to form acetyl-CoA in the β-oxidation pathway, a series of four reactions in which the β-carbon is oxidized, followed by the breakage of the bond between the α- and β-carbons. Peroxisomal β-oxidation shortens very long fatty acids. Additional reactions are required to degrade odd-chain and unsaturated fatty acids and dicarboxylic acids. When the acetyl-CoA product is present in excess, ketone bodies are produced and are used as an energy source in some tissues.
4. Fatty acid synthesis begins with the carboxylation of acetyl-CoA to form malonyl-CoA. ACC, a key enzyme in fatty acid metabolism, is regulated by allosteric modulators and phosphorylation reactions. The remaining reactions of fatty acid synthesis take place on the fatty acid synthase multienzyme complex. Several enzymes are available to elongate and desaturate dietary and newly synthesized fatty acids.
5. Both short- and long-term regulatory mechanisms are used to control fatty acid metabolism. Short-term regulation includes the use of malonyl-CoA as an inhibitor of CAT-I, and AMPK-catalyzed phosphorylation of ACC1 and glycerol-3-phosphate acyltransferase. Hormones such as insulin, glucagon, epinephrine, and cortisol have roles in short- and long-term regulation. Long-term regulation of fatty acid metabolism involves changes in gene expression triggered by transcription factors. Two prominent examples are the

SREBPs and the PPARs.

6. After phospholipids have been synthesized at the interface of the SER and the cytoplasm, they are often "remodeled"; that is, their fatty acid composition is adjusted. The turnover (i.e., the degradation and replacement) of phospholipids, mediated by the phospholipases, is rapid.
7. Cholesterol synthesis can be divided into three phases: formation of HMG-CoA from acetyl-CoA, conversion of HMG-CoA to squalene, and conversion of squalene to cholesterol. Cholesterol is the precursor for all steroid hormones and the bile salts. Bile salts are used to emulsify dietary fat. Cholesterol homeostasis is achieved through regulation of cholesterol biosynthesis, LDL receptor activity, and bile acid biosynthesis.

Take your learning further by visiting the **companion website** for *Biochemistry* at **www.oup.com/us/mckee**, where you can complete a multiple-choice quiz on lipid metabolism to help you prepare for exams.

Chapter 12 Review Quiz

Suggested Readings

Calderon-Dominguez M, et al. 2016. Fatty acid metabolism and the basis of brown adipose tissue function. Adipocyte 5(2):98–118.

Fan W, Evans R. 2016. The quest to burn fat, effortlessly and safely. Science 353:749–50.

Howe V, et al. 2016. How do cells sense sterol excess? Chemistry and Physics of Lipids 199:170–8.

Liu W, et al. 2016. Pathogenesis of nonalcoholic steatohepatitis. Cell Mol Life Sci 73:1969–87.

Rashidi B, et al. 2017. Antiatherosclerotic effects of vitamins D and E in suppression of atherosclerosis. J Cell Physiol 232:2968–76.

Scheele C, Nielsen S. 2017. Metabolic regulation and the anti-obesity perspectives of human brown fat. Redox Biology 12:770–5.

Wang Y, Xu D. 2017. Effects of aerobic exercise on lipids and lipoproteins. Lipids in Health and Disease 16:132–9.

Key Words

acyl carrier protein, *461*

AMPK, *470*

atherosclerosis, *473*

bile salts, *445*

biotransformation, *485*

chylomicron remnants, *446*

conjugation reaction, *481*

cytochrome P_{450} system, *485*

detoxication, *485*

detoxification, *485*

epoxide, *486*

fatty acid–binding protein, *451*

flavin-containing monooxygenase, *485*

glyceroneogenesis, *448*
ketogenesis, *456*
ketone body, *456*
ketosis, *456*
lipogenesis, *448*
lipolysis, *448*
β-oxidation, *451*
phase I reaction, *485*
phase II reaction, *485*
PPAR, *471*
SREBP, *471*
sterol carrier protein, *478*
thiolytic cleavage, *454*
triacylglycerol cycle, *447*
turnover, *472*

Review Questions

SECTION 12.1

Comprehension Questions

1. Define the following terms:
    a. enterocyte
    b. lipoprotein lipase
    c. ApoB-48
    d. ApoE
    e. ApoC-II
2. Define the following terms:
    a. chylomicron remnant
    b. hypertriglyceridemia
    c. glyceroneogenesis
    d. perilipin-1
    e. ATGL
3. Define the following terms:
    a. HSL
    b. CGI-58
    c. fatty acid-binding protein
    d. *β*-oxidation
    e. carnitine
4. Define the following terms:
    a. thiolytic cleavage
    b. ketogenesis
    c. ketone bodies

d. α-oxidation
e. ACC

5. Define the following terms:

a. AMPK
b. ACP
c. SREBP
d. PPAR
e. atherosclerosis

6. Define the following terms:

a. sdLDL
b. oxLDL
c. cytochrome P_{450}
d. atheroma
e. abetalipoproteinemia

Fill in the Blanks

7. The ketone bodies are _____, _____, and _____.
8. After the triacylglycerol content of VLDL has been depleted, the lipoprotein is referred to as a(n) _____.
9. Glycerol generated from TG hydrolysis in adipocytes is converted by the liver into _____, which serves as a substrate for the synthesis of glucose.
10. _____ is the precursor of bile acids.
11. The oxidation of 1 mol of palmitic acid yields a total of _____ mol of metabolic water.
12. _____ is a carrier molecule required for the transport of fatty acids into mitochondrial matrix.
13. The carrier of molecules of carbon dioxide in fatty acid synthesis is _____.
14. β-Hydroxybutyrate is the product of _____ metabolism.

Short-Answer Questions

15. What are the differences between β-oxidation in mitochondria and in peroxisomes? What similarities are there between these processes?
16. Describe the difference between IDLs and LDLs.
17. Explain why low-fat diets supplemented with medium-chain fatty acids are prescribed for patients with abetalipoproteinemia.
18. Compare the energy content of a stearic acid molecule compared with that of glucose.
19. Determine the number of moles of ATP that can be generated from the fatty acids in 1 mol of tristearin, a triacylglycerol composed of glycerol esterified to three stearic acid molecules. What is the fate of the glycerol product in the body?
20. What is the function of each of the following substances: AMPK, lipoprotein lipase, and hormone sensitive lipase?
21. In the intestine, the triacylglycerols must be converted to fatty acids and glycerol by hydrolytic enzymes before transport into enterocytes. Afterward, fatty acids and glycerol are reconverted into triacylglycerols and then packaged into chylomicrons. Suggest why this energy-requiring process is used instead of a direct transport of triacylglycerols into enterocytes.
22. The peroxisomal enzyme β-ketoacyl-CoA thiolase does not bind medium-chain acyl-CoA, in contrast to the analogous mitochondrial enzyme. Explain why this phenomenon is an advantage to the cell.
23. Review the steps in β-oxidation and determine which reactions are oxidation reactions.
24. Many processed and fast foods contain trans fatty acids. Explain why these molecules are a

problem for the body.

25. Provide an explanation for the intracellular separation of fatty acid biosynthesis (cytoplasm) and degradation (mitochondria and peroxisomes).
26. Describe the role of insulin in lipid metabolism.
27. In experiments investigating fatty acid synthesis when $^{14}CO_2$, one of two substrates in the synthesis of malonyl-CoA, is introduced, no label appears in the eventual fatty acid products. Explain.
28. Why is the enzyme-catalyzed introduction of carbon–carbon double bonds into fatty acids called an electron transport system?
29. What function does the enzyme ACC1 serve, and how is it regulated?
30. *β*-Oxidation of naturally occurring monounsaturated fatty acids requires an additional enzyme. What is the name of this enzyme? How does it accomplish this task?
31. Outline the reactions in the *α*-oxidation pathway. Why is *α*-oxidation necessary?
32. Before *β*-oxidation, fatty acids must be converted to their CoASH derivatives. Explain why this reaction is necessary.
33. Define the term *thiolytic cleavage*. In what biochemical process does it occur?
34. Explain how hormones act to modify the short and long term metabolism of fatty acids.
35. List three differences between fatty acid synthesis and *β*-oxidation.
36. Describe the mechanism in which NADH is involved in the activation of oxygen molecules in fatty acid desaturation processes.
37. Under severe starvation conditions, affected individuals develop “acetone breath.” Explain.
38. What are the roles of SREBPs and PPARs in fatty acid metabolism?
39. Describe the endogenous lipoprotein pathway.

Critical-Thinking Questions

40. Describe the possible effects of low levels of carnitine on a person’s metabolism.
41. In type I diabetes mellitus, excessive production of acetyl-CoA may surpass the body’s capacity to oxidize it. As a result, acetoacetate, *β*-hydroxybutyrate, and acetone (ketone bodies) accumulate. When generated in large amounts, blood pH falls, thereby reducing the capacity of red blood cells to carry oxygen. Explain in general terms how high concentrations of ketone bodies may result in a fatal coma.
42. The acyl-CoA dehydrogenase deficiency diseases are a group of inherited defects that impair the *β*-oxidation of fatty acids. Symptoms of the disease range from nausea and vomiting to frequent comas. Symptoms may be alleviated by eating regularly and avoiding periods of starvation (12 hours or more). Why does this simple remedy alleviate the symptoms?
43. During periods of stress or fasting, blood glucose levels fall. In response, fatty acids are released from adipocytes into the blood. Explain how such a blood glucose level decrease results in fatty acid release.
44. The adaptation of desert animals to their environment includes water conservation mechanisms. A number of such organisms conserve water so successfully that they never drink it. They depend instead on water generated by metabolism. Determine how much water can be obtained from the complete oxidation of 1 mole of palmitic acid.
45. Provide an explanation for the fact that most of the body’s fatty acids are 16 or 18 carbons long.

SECTIONS 12.2 & 12.3

Comprehension Questions

47. Define the following terms:

a. turnover
b. cyt P_{450}
c. detoxication
d. detoxification
e. epoxide

48. Define the following terms:

a. carboxybiotin
b. cytochrome b5
c. allyl group
d. conjugation reaction
e. sterol carrier protein

49. Define the following terms:

a. cholecystitis
b. HMGR
c. SCAP
d. statins
e. SRE

50. Define the following terms:

a. SAM
b. CDP-diacylglycerol
c. osteoporosis
d. osteoclast
e. bile acid

Fill in the Blanks

50. Oxidation of phytanic acid produces acetyl-CoA and _____ as end products.
51. _____ is the steroid precursor of the bile salts.
52. SAM is a donor of _____ groups in the synthesis of some phospholipid molecules.
53. The rate-limiting enzyme in bile acid synthesis is _____.
54. The products of bile acid synthesis are _____ and _____.
55. _____ is the transcription factor that activates the transcription of the LDL receptor gene.
56. _____ is the enzyme that catalyzes the condensation of 2 acetyl-CoAs to form β-ketobutyryl-CoA.

Short-Answer Questions

57. How do cells adjust the fluidity of their membranes?
58. How would you describe the three phases of biotransformation?
59. Outline the synthesis of bile acids. What functions do these substances have?
60. How are lipid molecules such as estrogen and β-carotene related to each other? What biosynthetic reactions do these molecules have in common?
61. List and describe the components of the cytochrome P_{450} electron transport system. What is the role of each component?
62. Draw the structure of a cholesterol molecule and indicate the isopentenyl units.
63. Why do gallstones form in susceptible individuals?
64. What conjugation reactions do bile acids undergo and what is their function?

Critical-Thinking Questions

65. Phospholipases show an enhanced activity for a substrate above the critical micelle concentration. (The critical micelle concentration, or cmc, is that concentration of a lipid above which micelles begin to form.) What type(s) of noncovalent interactions is (are) possible between the lipid molecules and the enzyme?
66. What do the interactions in Question 65 suggest about the structure of phospholipases?
67. There is a relatively high concentration of phosphatidylcholine on the lumenal side of the ER membrane. What feature of membranes is responsible for this effect?
68. Describe the effects of a class of pharmaceuticals called the statins on patients.
69. An experimenter using acetyl-CoA with a ^{14}C on the carbonyl group traces the radioactive label in cells synthesizing cholesterol. On what atoms of mevalonate will the label appear?
70. Determine the position of ^{14}C (see Question 69) in isopentenylpyrophosphate.

MCAT Study Questions

71. Statin drugs are used to treat hypercholesterolemia because they

    a. prevent cholesterol from inserting in plasma membranes
    b. prevent cholesterol ester formation
    c. inhibit HMG-CoA reductase
    d. stimulate HMG-CoA synthase

72. _____ occurs within mitochondria.

    a. fatty acid synthesis
    b. β-oxidation
    c. VLDL synthesis
    d. LDL oxidation

73. Which of the following molecules are not involved in lipolysis within adipocytes?

    a. protein kinase A
    b. PEPCK
    c. adenylate cyclase
    d. perilipin-1

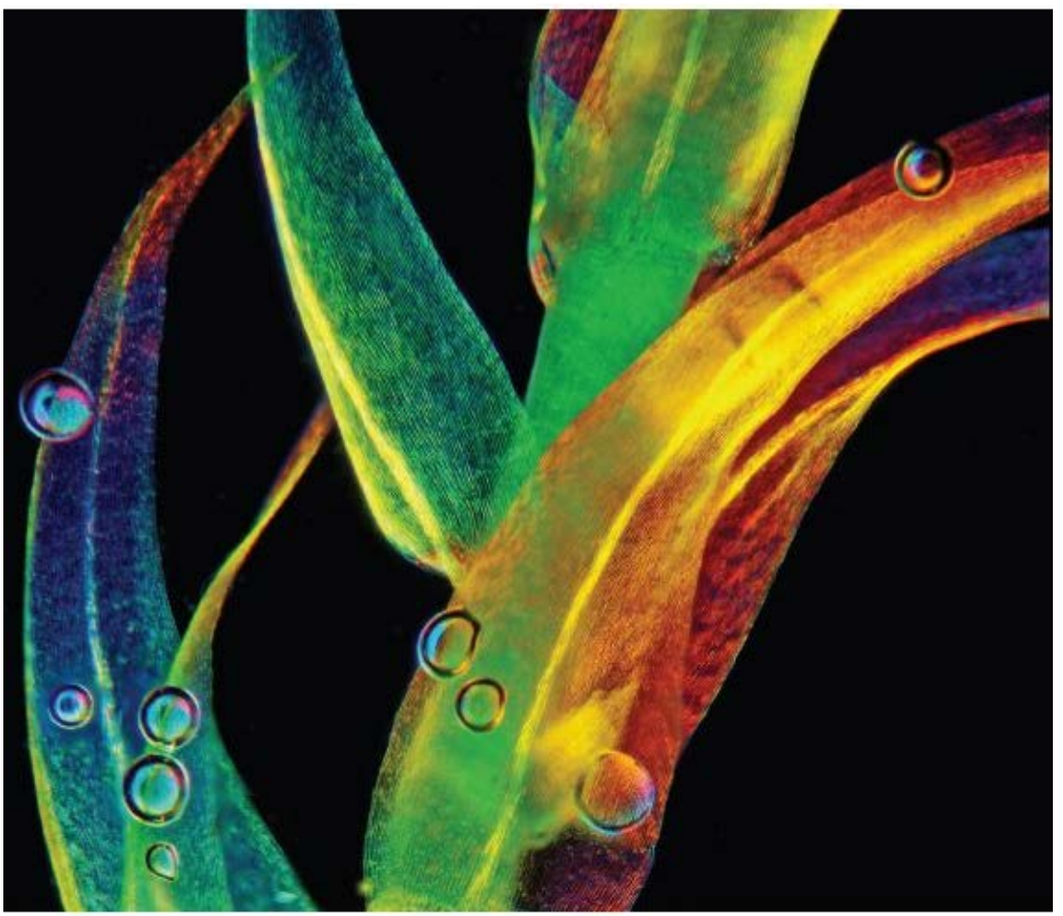
74. Insulin promotes the following processes in its target cells except for _____.

    a. GLUT-4 transport to plasma membranes
    b. ATP-citrate lyase activation
    c. stimulating glucose-6-phosphate lyase
    d. activation of ACC1

75. Atherosclerosis is promoted by all of the following except _____.

    a. sdLDL
    b. nitric oxide
    c. inflammatory chemokines
    d. AGE formation

CHAPTER 13 Photosynthesis



Oxygenic Photosynthesis In this polarized light and dark field micrograph, a small moss plant undergoing photosynthesis is observed releasing oxygen bubbles.

OUTLINE

CLIMATE CHANGE, RENEWABLE ENERGY, AND PHOTOSYNTHESIS

13.1 CHLOROPHYLL AND CHLOROPLASTS

13.2 LIGHT

13.3 LIGHT REACTIONS
Photosystem II and Water Oxidation
Photosystem I and NADPH Synthesis
Photophosphorylation

13.4 THE LIGHT-INDEPENDENT REACTIONS

The Calvin Cycle
Photorespiration
Alternatives to C3 Metabolism

13.5 REGULATION OF PHOTOSYNTHESIS
Light Control of Photosynthesis
Control of Ribulose-1,5-Bisphosphate Carboxylase

Biochemistry in Perspective
The Artificial Leaf: Biomimetic Photosynthesis

AVAILABLE ONLINE
Biochemistry in Perspective
Starch and Sucrose Metabolism
Biochemistry in the Lab
Photosynthetic Studies
Biochemistry in Perspective
Photosynthesis in the Deep
Biochemistry in Perspective
Crassulacean Acid Metabolism

Climate Change, Renewable Energy, and Photosynthesis

Ever-accelerating glacier melting and iceberg formation in both the Arctic Circle and Antarctica combined with rising sea levels are impressive indicators that the Earth's climate is warming up! During the Earth's long history, there have been numerous transitions between glacial "ice ages" and interglacial warm periods. These climate changes often lasted from thousands to tens of thousands of years. Abrupt transitions lasting only a few decades have occurred but were always associated with dramatic natural events. Significant changes in the Earth's orbit or the sun's intensity can alter the amounts of solar radiation that reach Earth. Volcanic eruptions of unusual duration and intensity can emit massive amounts of CO_2 and other greenhouse gases that trap solar heat, and altered ocean current patterns can redistribute heat around the Earth. Since the beginning of the Industrial Revolution about 150 years ago, however, a steady increase in fossil fuel burning and deforestation has resulted in the accumulation of massive amounts of greenhouse gases in the atmosphere. Their impact on the Earth's climate is significant.

Global warming is a difficult and seemingly intractable problem, with consequences that include an increased incidence and intensity of forest fires, drought and flooding, and catastrophic food and water shortages. Other effects include economic damage from more powerful hurricanes and an increased prevalence of insect-borne diseases.

Slowing the progress of global warming will require substantial investments. New technologies are needed to replace fossil fuels with economically competitive fuels that are carbon neutral (i.e., with no net release of CO_2 into the atmosphere). *Biofuels* are renewable energy sources synthesized by living organisms directly or indirectly via photosynthesis, the light energy–driven process that converts CO_2 and H_2O into organic molecules. Currently produced biofuels include ethanol (from corn kernels), biodiesel (long-chain alkyl esters derived from vegetable oil or animal fat), and microbe-generated inorganic molecules such as hydrogen gas (H_2).

Sustainable biofuel production requires that the following three criteria be met. First, biofuels must be economically generated in large enough quantities—about 10 terawatts (TW) of energy per year (1 TW = 10^{12} watts or 3.2×10^{18} J/year). Second, biofuels must not compete with food production. The use of arable farmland for biomass-derived energy production drives up the cost and availability of food. Finally, biofuel

production must not adversely affect the environment. Most biofuel production methods currently in use fail to meet these criteria. For example, corn-based ethanol production is cost ineffective (energy costs alone account for at least 90% of output), and the diversion of corn, a food staple, into energy generation has caused increased food prices. Palm oil-derived biodiesel, made possible by new plantations in Southeast Asia, has resulted in deforestation that not only has destroyed the habitat of endangered animals (e.g., tigers, gibbons, and orangutans), but has also removed a critical "sink" for atmospheric CO_2 removal.

Considering the drawbacks of current biofuel technologies, is there any hope for the development of a successful fossil fuel substitute? Current research demonstrates that biodiesel production involving photosynthetic algae and cyanobacteria, which does not use arable land, can be up to 100 times that obtained from current methods. It is hoped that the combination of this promising technology with improvements in energy generation from agricultural, industrial, and municipal waste can contribute substantially to solving the Earth's human-caused climate change problem.

Overview

WITHOUT QUESTION, OXYGENIC PHOTOSYNTHESIS IS THE MOST IMPORTANT BIOCHEMICAL PROCESS ON EARTH. WITH A FEW MINOR EXCEPTIONS, photosynthesis is the only mechanism by which an external abiotic source of energy is harnessed by the living world, and it is the source of O_2, which sustains all aerobic organisms. Chapter 13, which is devoted to a discussion of the principles of photosynthetic processes, emphasizes the relationship between photosynthetic reactions, the structure of chloroplasts and the relevant properties of light.

Photosynthesis is the light-driven biochemical mechanism whereby CO_2 is incorporated into organic molecules such as glucose (**Figure 13.1**). Captured light energy is used to synthesize ATP and NADPH, which drive this process. The reducing power of NADPH is necessary because a strong electron donor is required to reduce the fully oxidized, low-energy carbon atoms in CO_2 to the carbon units of organic molecules.

KEY CONCEPTS

- Incorporating CO_2 into organic molecules requires energy and reducing power.
- In photosynthesis, both of these requirements are provided by a complex process driven by light energy.

Photosynthesis is performed by biochemical mechanisms, referred to as **photosystems**, which are membrane-bound protein complexes found in chloroplasts (plants and algae) or the cell membrane of photosynthetic bacteria. The protein subunits provide a scaffold for numerous types of pigment molecules, quinones, and Fe-S clusters. There are two photosystems in chloroplasts: photosystem I (PSI) and photosystem II (PSII). Each photosystem is composed of two functional components. The **light-harvesting antenna** captures solar energy and transfers it to a **reaction center**, which uses captured light energy to drive transmembrane electron transport. In PSI, light-driven electron transport is used to synthesize NADPH. Electron flow from PSII to cytochrome b_6f complex pumps protons across a membrane to subsequently drive ATP synthesis. In oxygenic photosynthesis, the electrons of water molecules replace the PSII reaction center electrons, yielding the waste product molecule O_2. Although O_2 is a highly reactive molecule, its toxicity is limited by the fact that it is a gas that rapidly diffuses away from the photosynthesizing organism in contrast to several nonoxygenic photosynthetic waste products (e.g., sulfur, sulfate, or nitrate).

Chapter 13 describes the principles of oxygenic photosynthesis. The discussion begins with a detailed view of chloroplast structure. A brief review of the relevant properties of light is followed by a description of the reactions that constitute modern photosynthesis. These include the light reactions and the light-independent reactions. It is during the light reactions that electrons are energized and eventually used in the synthesis of both ATP and NADPH. Both molecules are then used in the light-independent reactions (often referred to as the dark reactions or Calvin cycle) to drive the synthesis of carbohydrate. C4 metabolism, a photosynthetic variation is also discussed. Another variation, crassulacean acid metabolism, is described in an online essay. The chapter ends with a discussion of several mechanisms that control photosynthesis in plants.

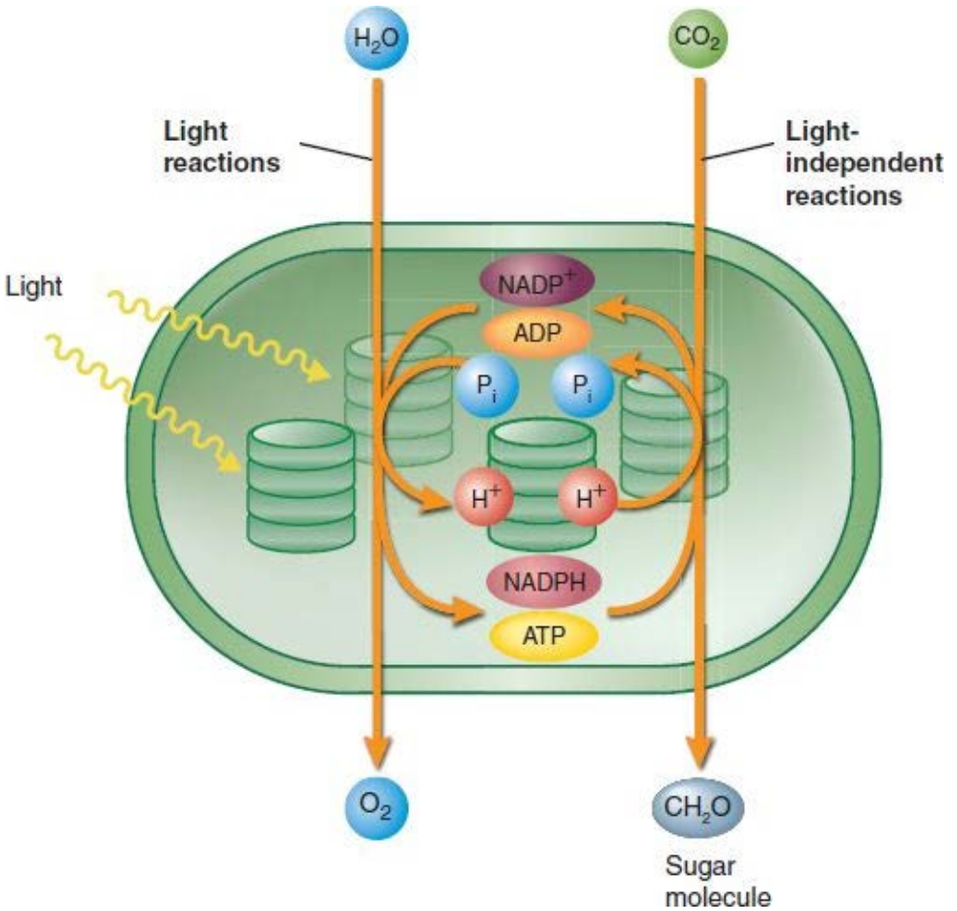


FIGURE 13.1

Overview of Oxygenic Photosynthesis

Oxygenic (oxygen-producing) photosynthesis occurs in two sets of reactions: the light reactions and the light-independent reactions. In chloroplasts (**Figure 2.25**), the light reactions occur within the thylakoid membrane, and the light-independent reactions (the Calvin cycle) occur within the stroma. The light reactions use light energy to drive the synthesis of NADPH and ATP. Water molecules are the source of the electrons and protons used to synthesize these molecules. O_2 is released as a waste product. The light-independent reactions use NADPH and ATP to convert CO_2 into sugar molecules.

13.1 CHLOROPHYLL AND CHLOROPLASTS

Photosynthesis begins with the absorption of light energy by specialized pigment molecules (**Figure 13.2**). The **chlorophylls** are green pigment molecules that resemble heme. *Chlorophyll a* plays the principal role in oxygenic photosynthesis because its absorption of light energy directly drives photochemical events. Chlorophyll a is also involved in **light harvesting**, the process whereby

absorbed energy is channeled to a reaction center. *Chlorophyll b* is a light-harvesting pigment that passes absorbed energy on to chlorophyll a. The orange-colored **carotenoids** are isoprenoid molecules that either function as light-harvesting pigments (e.g., lutein, a xanthophyll, see p. 416) or protect against overexcitation and reactive oxygen species (e.g., *β*-carotene).

Chloroplasts, the photosynthetic organelle in plants and algae, are, like mitochondria, the result of a primordial endosymbiotic event. Chloroplasts are descendants of primitive cyanobacteria; mitochondria are descendants of *α*-proteobacteria. Both organelles have an outer and an inner membrane with different permeability characteristics (**Figure 13.3**). As with mitochondria, the outer membrane of chloroplasts is highly permeable, and the inner membrane possesses specialized carrier molecules that regulate molecular traffic. Unlike the mitochondrial inner membrane, the chloroplast inner membrane does not possess electron transport proteins or ATP synthase. Instead, it is involved in the incorporation of fatty acids into several types of membrane lipid molecules and the synthesis of amino acids, chlorophyll, carotenoids, *α*-tocopherol (p. 417) and plastoquinone (p. 417), and the assimilation of nitrogen atoms into organic biomolecules. The chloroplast inner membrane encloses an inner space, referred to as the **stroma**, that resembles the mitochondrial matrix. The stroma possesses a variety of enzymes (e.g., those that catalyze the light-independent reactions and fatty acid and starch synthesis), DNA, and ribosomes.

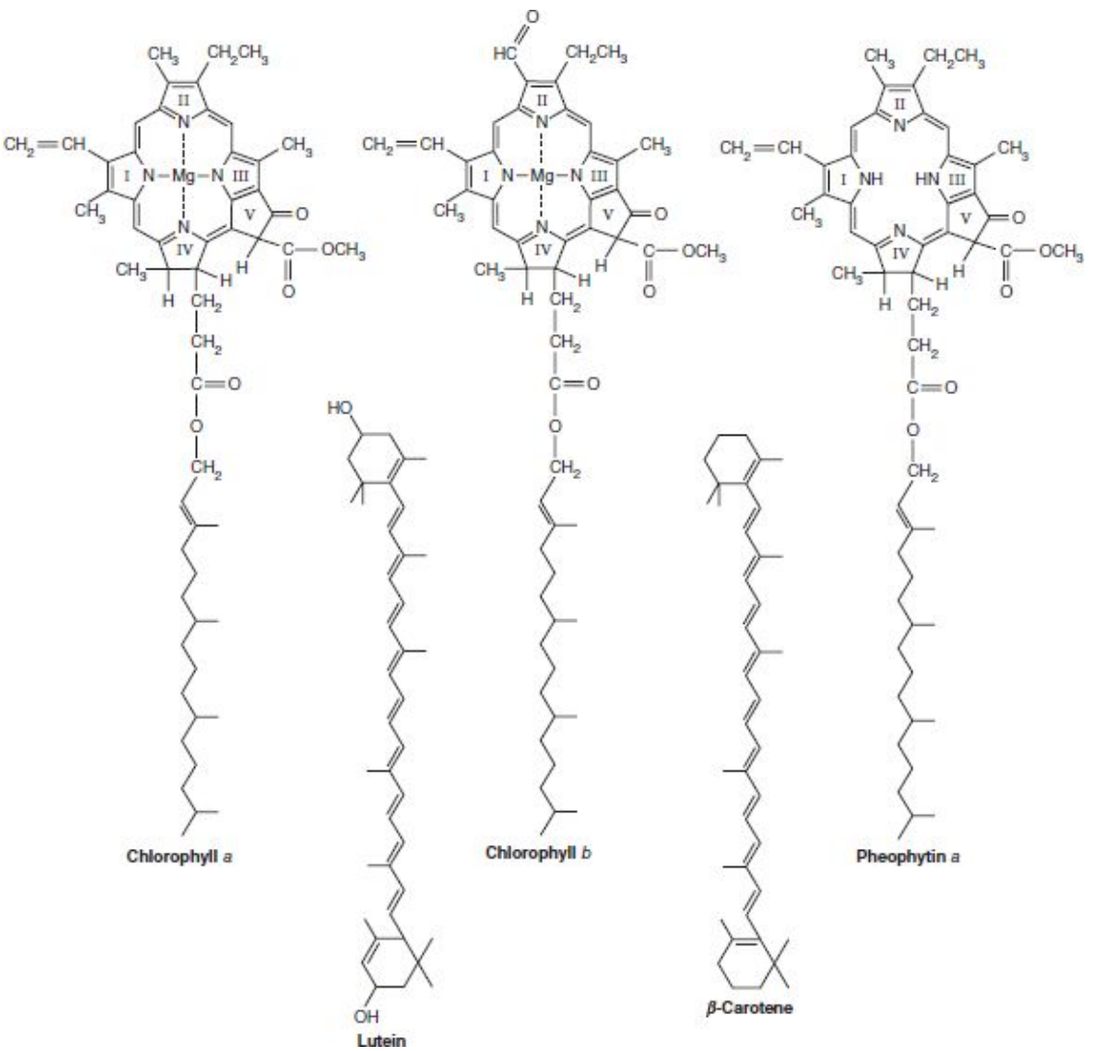


FIGURE 13.2

Pigment Molecules Used in Photosynthesis

Chlorophylls a and b are found in almost all oxygenic photosynthesizing organisms. They possess a complex cyclic structure (called a porphyrin) with an Mg (II) ion at its center. Chlorophyll a possesses a methyl group attached to ring II of the porphyrin, whereas chlorophyll b has an aldehyde group attached to the same site.

Pheophytin a is similar in its structure to chlorophyll a. The Mg (II) ion is replaced by two protons. Chlorophylls a and b and pheophytin a all possess a phytol chain esterified to the porphyrin. The phytol chain extends into and anchors the molecule to the membrane. Lutein and *β*-carotene are the most abundant carotenoids in thylakoid membranes.

3D animation of Lutein

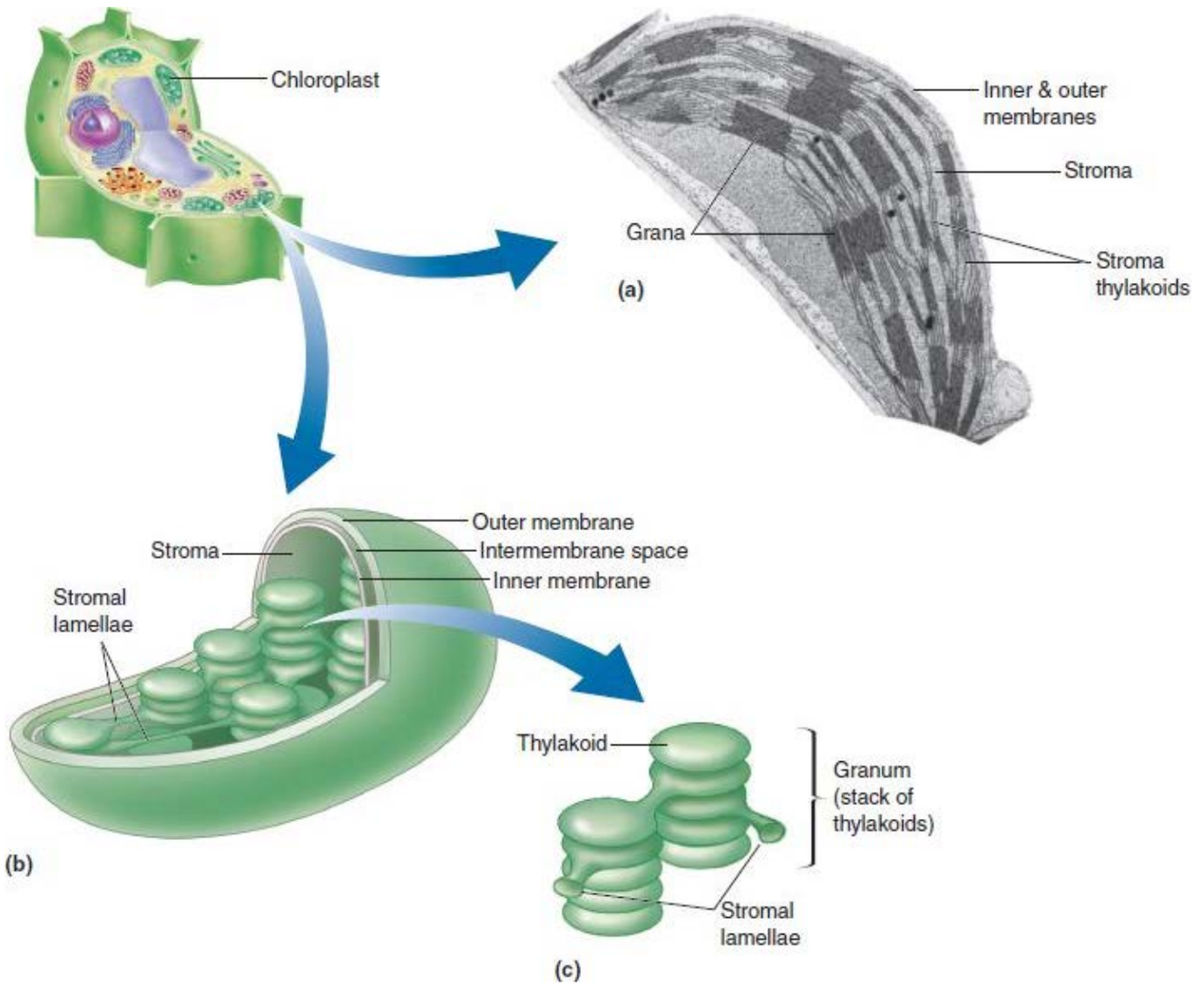


FIGURE 13.3

Chloroplast Structure

Chloroplasts have inner and outer membranes. A third membrane forms within the aqueous, enzyme-rich stroma into flattened sacs called thylakoids. A stack of thylakoids is called a granum. Unstacked, connecting thylakoid membrane is referred to as stromal lamellae. (a) Electron micrograph of a chloroplast. (b) Diagrammatic view of a chloroplast. (c) Cutaway view of a granum. Chloroplasts are substantially larger than plant mitochondria, rod-shape structures approximately 1500 nm long and 500 nm wide. Chloroplasts are spheroidal, from 4000 to 6000 nm long and approximately 2000 nm wide.

In addition, chloroplasts also possess a distinct third membrane called the **thylakoid membrane** that is folded into stacks of disc-like vesicular structures called **grana** (singular = **granum**). The internal compartment created by the formation of grana is referred to as the **thylakoid lumen** (or *space*). The thylakoid membranes that interconnect the grana are called **stromal lamellae**. Adjacent

layers of membrane that fit closely together within each granum are said to be appressed. The stromal lamellae are nonappressed.

The pigments and proteins responsible for the light-dependent reactions of photosynthesis are found within thylakoid membrane (**Figure 13.4**). Most of these molecules are organized into the working units of photosynthesis.

1. **Photosystem I.** PSI (**Figure 13.5**) is a large membrane-spanning multisubunit complex that energizes and transfers the electrons that eventually are donated to $NADP^+$. Two special chlorophyll a molecules that reside within the reaction center perform the essential role of PSI, the donation of energized electrons to a series of electron carriers within thylakoid membrane. These chlorophyll a molecules, referred to as a *special pair*, are located in the core complex of PSI, the PsaAB dimer. The special pair in PSI is sometimes referred to as P700 because it has a maximum absorption of light at 700 nm. In addition to the special pair, the PsaAB dimer contains a series of single electron carriers: A_0, A_1, and F_x. A_0 is a specific chlorophyll a molecule that accepts an energized electron from P700 and transfers it to A_1, which has been identified as phylloquinone (vitamin K_1). The electron is then transferred from A_1 to a series of iron–sulfur clusters (F_x, F_A, and F_B). Ultimately, the electron is donated to $NADP^+$ to form NADPH. PSI also contains a large number of chlorophyll a molecules other than the special pair, as well as chlorophyll b and carotenoids that act as antenna pigments. **Antenna pigments** absorb light energy and transfer it to the reaction center. Additional antenna pigment molecules in a peripheral light-harvesting complex (LHCI) associated with PSI also contribute to efficient absorption of light energy. This phenomenon is described more fully in Section 13.2. Most PSI complexes are located in nonappressed thylakoid membrane, that is, membrane that is directly exposed to the stroma.

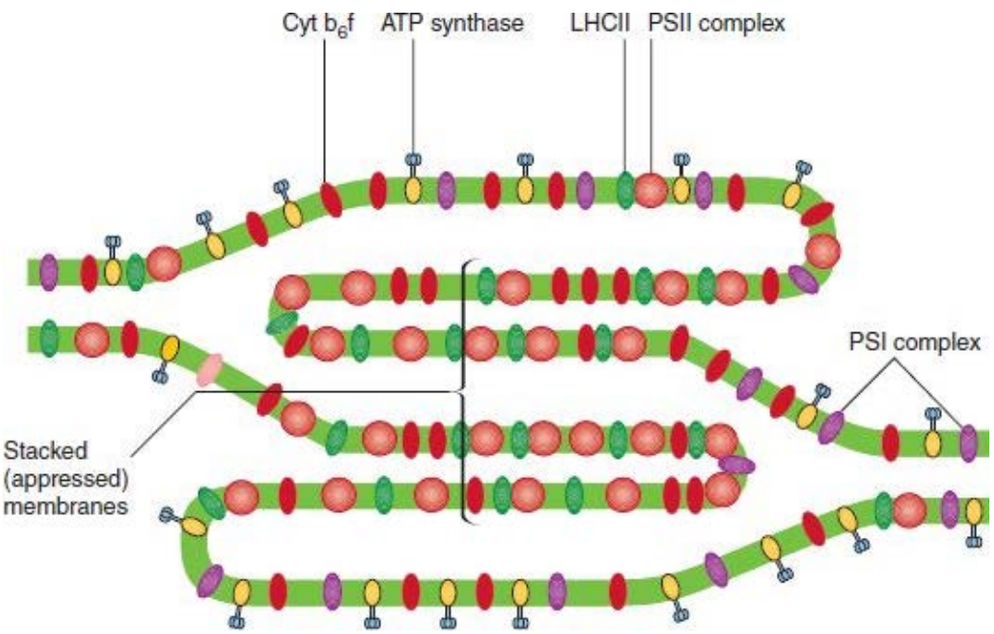


FIGURE 13.4

The Working Units of Photosynthesis

PSI complexes are most abundant in the unstacked stromal lamellae. In contrast, PSII complexes are located primarily in the stacked regions of thylakoid membrane. Cytochrome b_6f is found in both areas of thylakoid membrane. The ATP synthase is found only in thylakoid membrane that is directly in contact with the stroma.

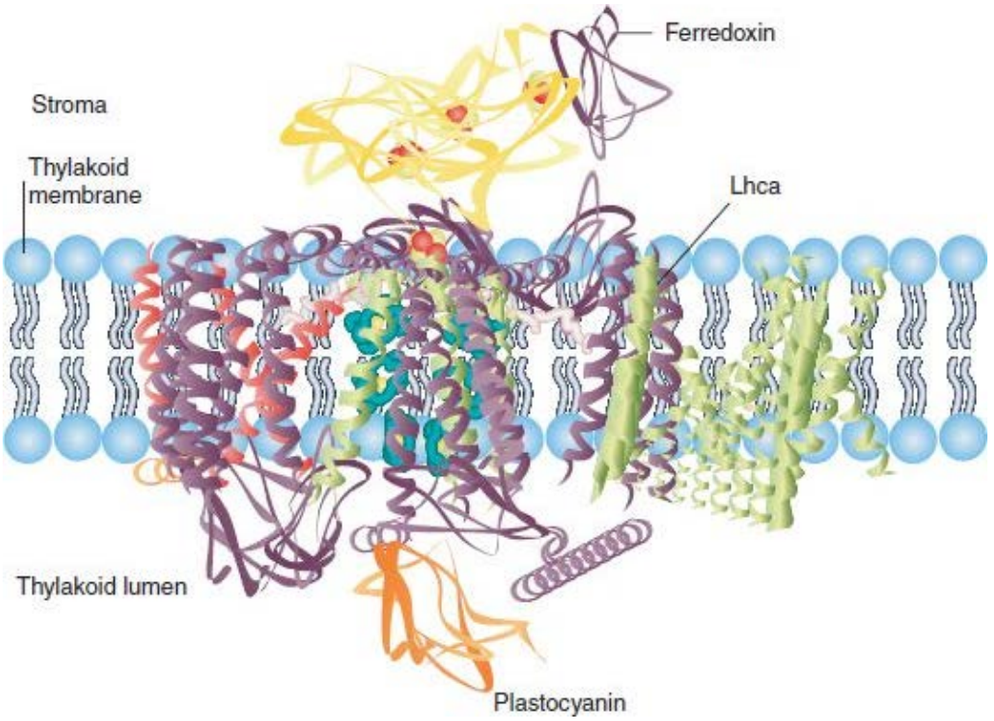


FIGURE 13.5

Structure of Photosystem I in Plants

PSI is a multisubunit protein complex composed of a reaction center formed largely by a PsaA and PsaB heterodimer and the peripheral light-harvesting complex (LHCI), shown here as green ribbon structures. Two other electron transport proteins are also shown: plastocyanin (orange ribbon structure) and ferredoxin (dark pink ribbon structure). The yellow ribbon structure is a subunit of PSI. Only one of four Lhca subunits is shown.

2. **Photosystem II.** The function of PSII is to oxidize water molecules and donate energized electrons to electron carriers that eventually reduce PSI. PSII is a large membrane-spanning protein–pigment complex located in appressed grana membrane (**Figure 13.6**). The most active form of PSII is a dimer. The PSII reaction center is composed of two polypeptide subunits known as D_1 (33 kDa) and D_2 (31 kDa) (the D_1/D_2 dimer), two core subunits, CP47 and CP43, and cytochrome b_{559}. The D_1/D_2 complex binds a special pair of chlorophyll a molecules (referred to as P680) that absorbs light at 680 nm. Once it absorbs light energy, P680 transfers an excited electron to pheophytin a (Pheo), thereby creating two charged species: one positively charged ($P680^+$) and one negatively charged ($Pheo^-$). This event is referred to as a *charge separation*. The energized electron is then transferred to a series of other electron acceptors and eventually to *plastoquinone* (PQ or Q), a molecule similar to ubiquinone. The electron donated by the reaction center is replaced by the *oxygen-evolving complex* (OEC), also known as the water-splitting complex, and a tyrosine residue, often referred to as Y_Z, located on D_1. The water-splitting site is a cube-like Mn_4CaO_5 cluster surrounded by amino acid side chains of D_1 and CP43 that form direct ligands to the metals. Several hundred antenna pigment molecules are also associated with the reaction center. The preponderance of accessory pigment molecules and several proteins belong to a detachable unit referred to as *light-harvesting complex II (LHCII)*. LHCII is a trimer of light-harvesting proteins, each of which binds 12 to 14 chlorophyll a and b molecules as well as several carotenoid molecules. In plants, LHCII trimers are tightly packed in appressed grana membranes.

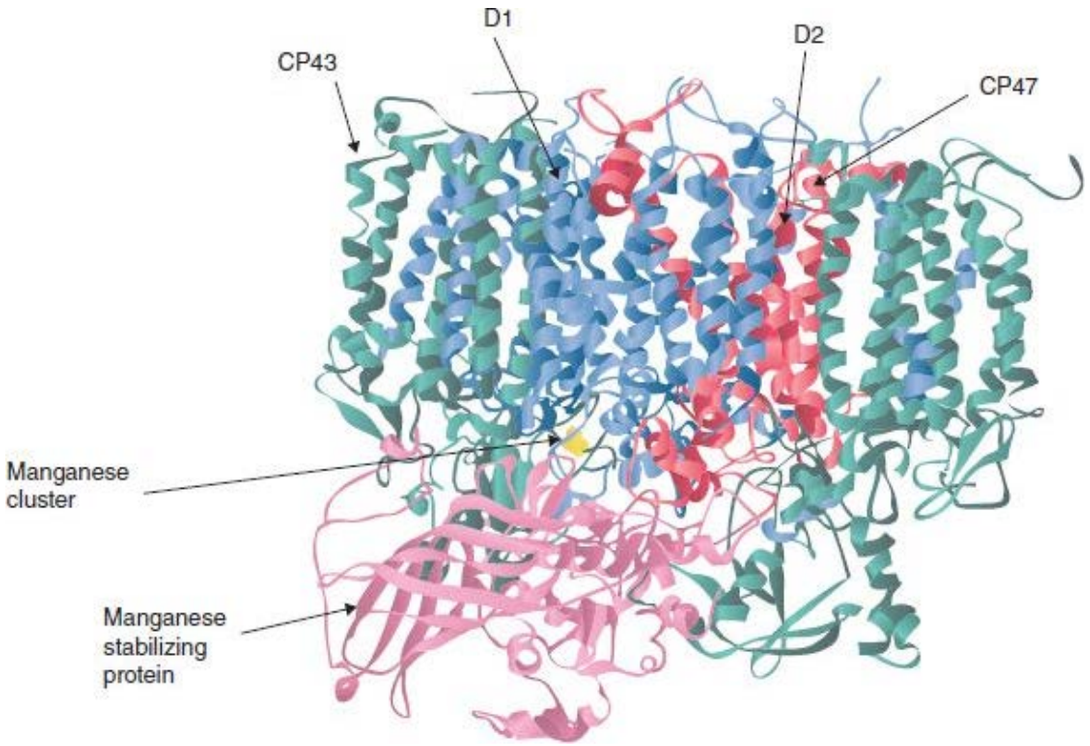


FIGURE 13.6

Photosystem II Monomer Structure

The PSII monomer is composed of about 20 subunits. The most important of these are the reaction pair subunits D_1 and D_2 and CP43 and CP47, antenna subunits that bind chlorophyll (not shown). Manganese stabilizing protein is a peripheral membrane protein component of PSII that maintains the efficiency of oxygen production by stabilizing the manganese cluster.

3. **Cytochrome b_6f complex**. Cytochrome b_6f complex, found throughout the thylakoid membrane, is similar in structure and function to the cytochrome bc_1 complex (**Figure 10.9**) in mitochondrial inner membrane. The cytochrome b_6f complex plays a critical role in the transfer of photoexcited electrons from PSII to PSI. An iron–sulfur protein in the complex, referred to as the Rieske Fe-S protein, accepts electrons from the membrane-soluble electron carrier PQ and donates them to a small water-soluble, copper-containing protein called plastocyanin (PC). The mechanism (**Figure 13.7**) that transports electrons from PQH_2 through the cytochrome b_6f complex is similar to the Q cycle in mitochondria.
4. **ATP synthase**. The chloroplast ATP synthase (**Figure 13.8**), also referred to as *CF_oCF_1ATP synthase*, is structurally and functionally similar to the mitochondrial ATP synthase. The CF_o component is a membrane-spanning protein complex that contains a proton-conducting channel. The CF_1 head piece, which projects into the stroma, possesses an ATP-synthesizing activity. A transmembrane proton gradient produced during light-driven electron transport causes the rotation of the CF_o proton channel complex, which in turn drives ADP phosphorylation. The synthesis of each ATP molecule requires pumping approximately four protons across the membrane into the thylakoid space. Thylakoid membrane that is directly in contact with the stroma contains the ATP synthase.

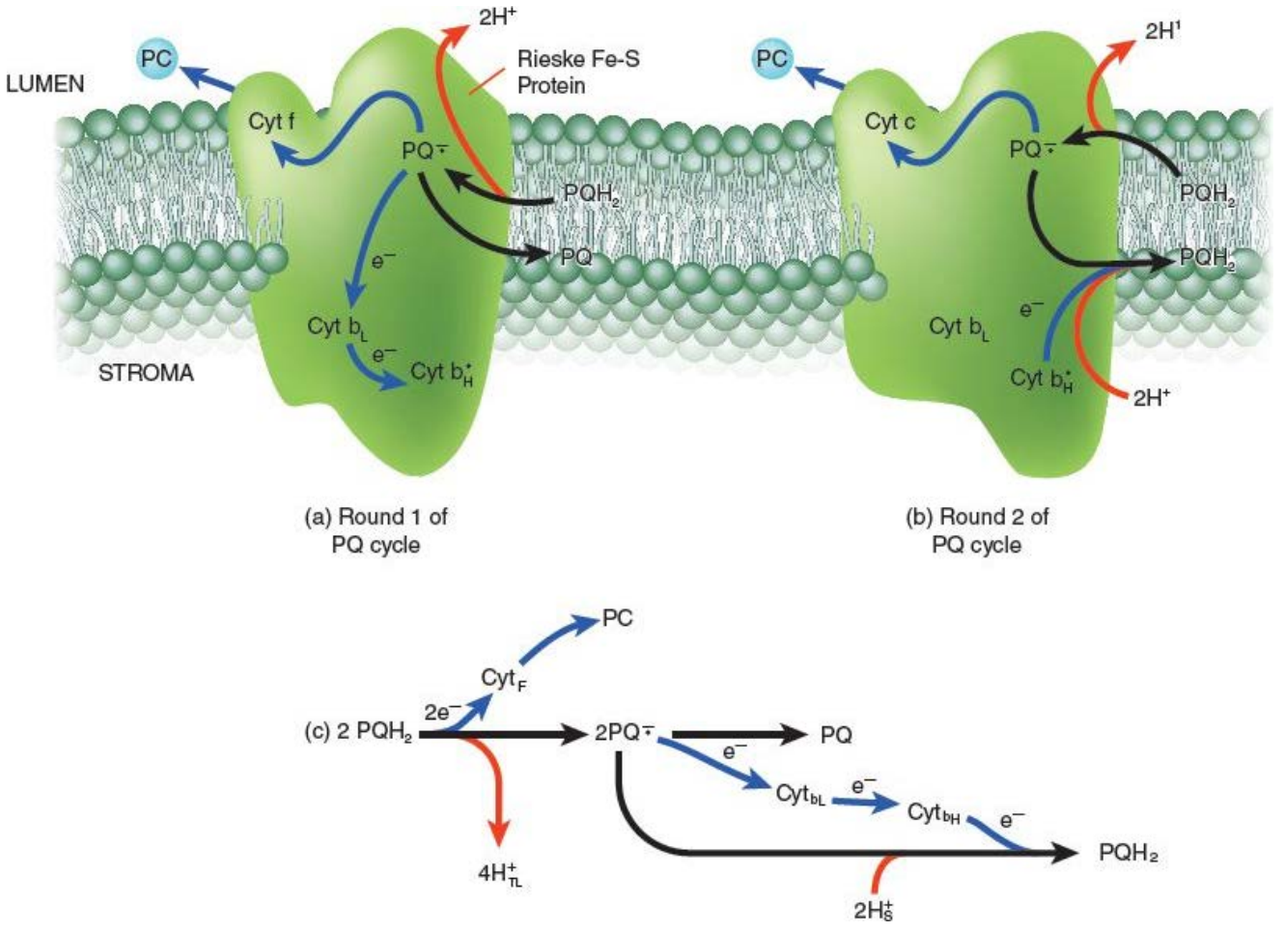


FIGURE 13.7

Electron Transport through Cytochrome b_6f Complex

Two molecules of plastoquinol (PQH_2) are oxidized sequentially to supply two electrons (e^-) to plastocyanin (PC). PC, a water-soluble copper-containing protein, then transfers each electron to the reaction center (P700) of PSI. The first electron from each PQH_2 is transferred to the Rieske iron–sulfur protein and then to cyt f and PC as two protons each are transferred to the thylakoid lumen (TL). (a) An electron from one PQ^- is transferred to cyt b_L and then to cyt b_H. PQ is released into the thylakoid membrane. (b) The second PQ^- picks up the electron from cyt b_H and 2 protons from the stroma (S) to form PQH_2. (c) The summary reaction shows that 2 PQH_2 molecules enter cyt b_6f and PQ and PQH_2 are released from the protein complex. (Black arrows = reaction arrows; red arrows = proton transfer; blue arrows = electron flow.)

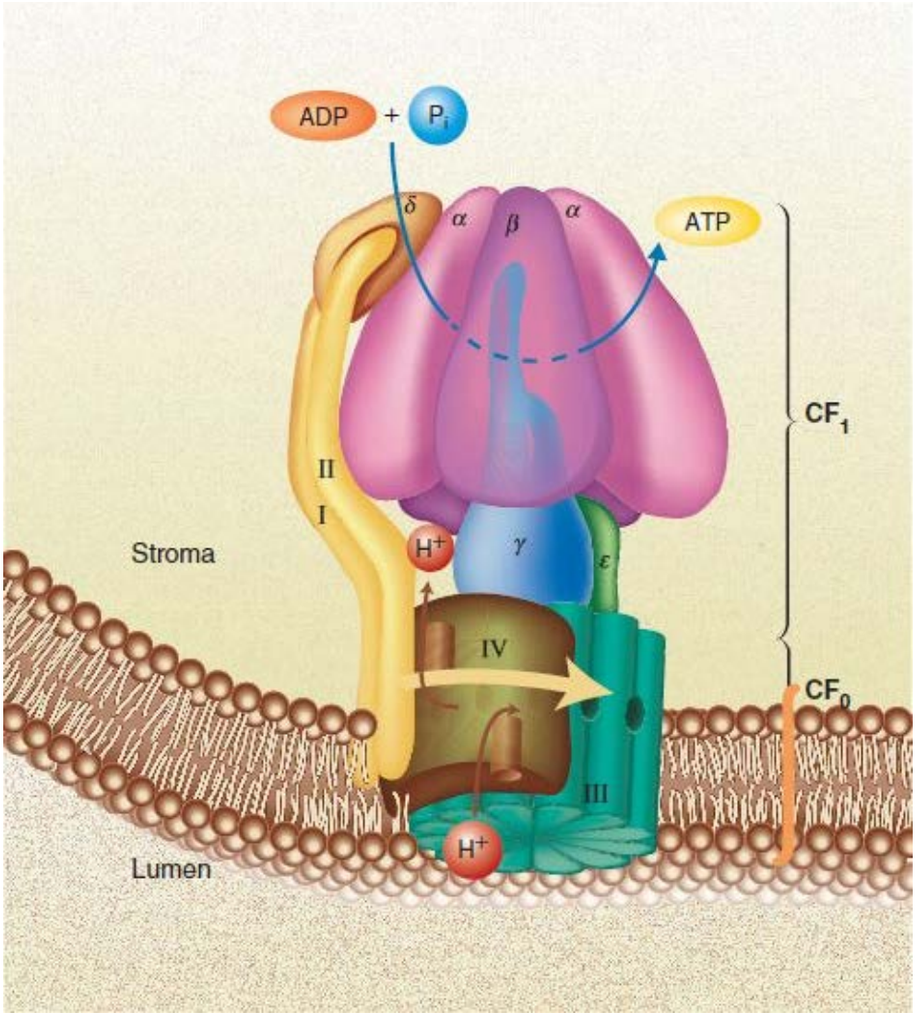


FIGURE 13.8

Diagrammatic View of the Chloroplast ATP Synthase

The ATP synthase is composed of two components: an integral membrane protein complex (CF_o) that contains a proton pore and an extrinsic protein complex (CF_1) that synthesizes ATP. CF_o contains four different types of subunits: I, II, III, and IV. The subunits I and II correspond to the b mitochondrial subunits. The proton pore is composed of 12–14 copies of subunit III, corresponding to mitochondrial c subunits. Subunit IV corresponds to the mitochondrial a subunit. CF_1 consists of five different subunits: α, β, γ, δ, and ε, which are similar to the subunit composition of mitochondrial F_1 complex (Chapter 10).

KEY CONCEPTS

- In chloroplasts a double membrane encloses an inner space called the stroma. The stroma contains the enzymes that catalyze the light-independent reactions of photosynthesis. The third membrane forms into flattened sacs called thylakoids.
- Thylakoid membrane contains the pigments and proteins of the light-dependent reactions.

QUESTION 13.1

Draw a sketch of a chloroplast and indicate where each of the following is located: CF_oCF_1, P700, P680, and Calvin cycle reactions. Describe the function of each.

13.2 LIGHT

The sun emits energy in the form of electromagnetic radiation, which propagates through space as waves, some of which impinge on Earth. Visible light, the energy source that drives photosynthesis, occupies a small part of the electromagnetic radiation spectrum (**Figure 13.9**). Of the approximately 178,000 TW of solar energy that reach Earth per year, only a fraction (100 TW) are captured by photosynthetic organisms. Many of the properties of light are explained by its wave behavior (**Figure 13.10**). Energy waves are described by the following terms:

1. **Wavelength**. Wavelength (λ) is the distance from the crest of one wave to the crest of the next wave.
2. **Amplitude**. Amplitude (a) is the height of a wave. The intensity of electromagnetic radiation (e.g., the brightness of light) is proportional to a^2.
3. **Frequency**. Frequency (v) is the number of waves that pass a point in space per second.

For each type of radiation, the wavelength multiplied by the frequency equals the velocity (c) of the radiation:

$$\lambda v = c$$

This equation rearranges to

$$\lambda = c/v$$

The wavelength therefore depends on both the frequency and the velocity of the wave.

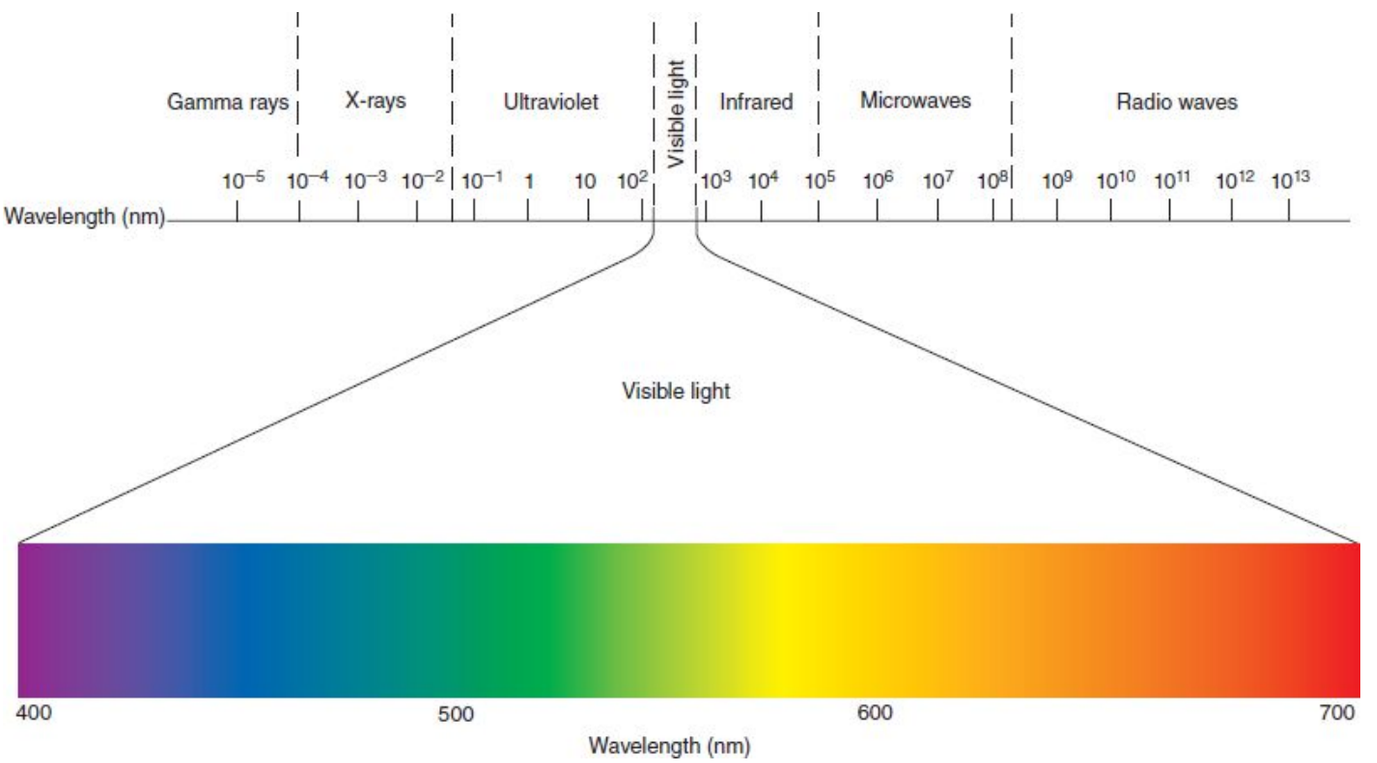


FIGURE 13.9

The Electromagnetic Spectrum

Gamma rays, which have short wavelengths, have high energy. At the other end of the spectrum, the radio waves (long wavelengths) have low energy. Visible light is the portion of the spectrum to which the visual pigments in the retina of the eye are sensitive. Pigment molecules in chloroplasts are also sensitive to portions of the visible

spectrum. UV light near the visible spectrum is subdivided into UVA (400–315 nm), UVB (315–280 nm), and UVC (280–100 nm).

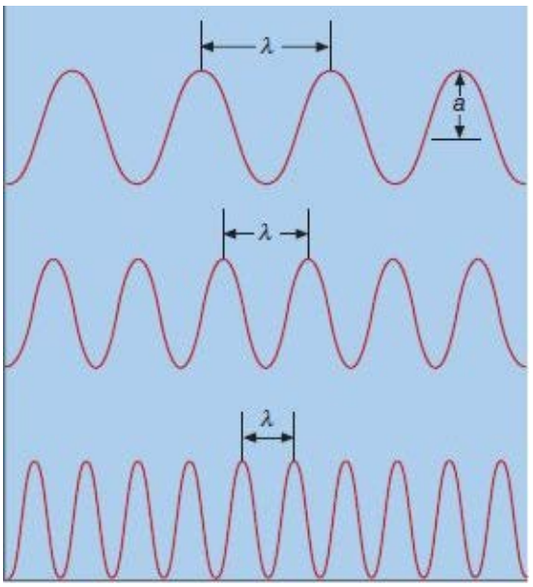


FIGURE 13.10

Properties of Waves

A wavelength (λ) is the distance between two consecutive peaks in a wave. The amplitude (a) or height of a wave is related to the intensity of electromagnetic radiation. Frequency is the number of waves that pass a point in space per second. Radiation with the shortest wavelength has the highest frequency.

The wavelengths of visible light range from 400 nm (violet light) to 700 nm (red light). In comparison, highly energetic X-rays and γ-rays have wavelengths that are 10^4 to 10^7 times shorter. On the other end of the spectrum are low-energy radio waves; these have wavelengths on the order of meters to kilometers.

QUESTION 13.2

Why do green light waves have less energy than blue light waves?

In addition to behaving like a wave, visible light and other types of electromagnetic radiation exhibit the properties of particles such as mass and acceleration. Einstein's observation that energy has mass, or $E = mc^2$, applies to the photon. When light interacts with matter, it does so in discrete packets of energy called photons. The energy (E) of a photon is proportional to the frequency of the radiation:

$$E = h\nu$$

where h is Planck's constant (6.63×10^{-34} J • s).

According to quantum theory, radiant energy can be absorbed or emitted only in specific quantities called quanta. When a molecule absorbs a quantum of energy, an electron is promoted from its ground state orbital (lowest energy level) to a higher-energy state. For absorption to occur, the

energy difference between the two energy states must exactly equal the energy of the absorbed photon. Complex molecules often absorb at several wavelengths. For example, chlorophyll produces an absorption spectrum with broad and multiple peaks (blue-violet region and red region). Both of these facts suggest that chlorophyll absorbs photons of many different energies with varying probabilities. The wavelengths that are not absorbed are visible to us, and so a chlorophyll solution (or a leaf) appears green. Carotenoids, observable in autumnal leaves in which chlorophyll molecules have been degraded, absorb blue light and therefore appear yellow or red.

Molecules that absorb electromagnetic energy have structural components called chromophores. Electrons in **chromophores** move easily to higher-energy levels when energy is absorbed. Visible chromophores typically possess extended chains of conjugated double bonds and aromatic rings. For example, the anthocyanins, a class of water-soluble pigments, contain chromophores that protect plants from light-induced damage by absorbing blue-green and UV light. Chromophores undergo electronic transitions wherein an electron moves from a ground state occupied orbital to a higher-energy unoccupied orbital (**Figure 13.11**). Molecules with a small number of conjugated double bonds or isolated double bonds absorb energy in the UV portion of the electromagnetic spectrum. The extensive conjugation of the chromophores of photosynthetic and accessory pigments allows electronic transitions to occur at longer wavelengths (lower energy) across the visible spectrum. The π electrons of these systems require less energy to make the transition into a higher-energy orbital.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in the Lab article on photosynthetic studies for a discussion of the absorption spectrum of chlorophyll and several other pigment molecules.

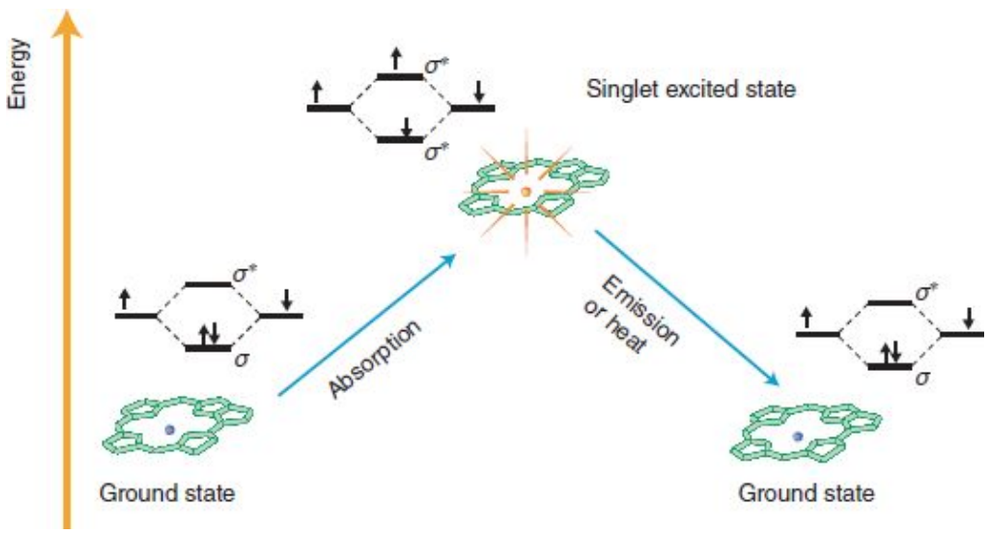


FIGURE 13.11

Excitation of a Chromophore

If a chromophore molecule (e.g., chlorophyll) absorbs a photon of visible light, its energy increases as a result of the movement of an electron to a higher molecular orbital. An excited molecule can return to its ground state by releasing energy as emission (fluorescence) or heat. The excited molecule can also donate its excitation energy to a neighbor molecule (excitation energy acceptor) or give away its electron to an electron acceptor.

Once excited, an electron can return to its ground state in several ways:

1. **Fluorescence**. In **fluorescence**, a molecule's excited state decays as it emits a photon. Because

the excited electron loses some energy initially by relaxing to a lower vibrational (energy) state, a transition resulting in the emission of a photon has lower energy than the photon originally absorbed. Fluorescent decay can occur as quickly as 10^{-15} s. (Although various chlorophylls absorb light energy throughout the visible spectrum, they emit only photons with low energy at or beyond the red end of the visible spectrum.)

2. **Resonance energy transfer**. In **resonance energy transfer**, the excitation energy is transferred to a neighboring chromophore through interaction between adjacent molecular orbitals. A chromophore whose absorption spectrum overlaps the emission spectrum of the target chromophore can absorb the quantum of energy released when that chromophore returns to its ground state.
3. **Oxidation-reduction**. An excited electron is transferred to a neighboring molecule. An excited electron occupies a normally unoccupied orbital and is bound less tightly than when it occupies a ground state orbital. A molecule with an excited electron is a strong reducing agent. It returns to its ground state by reducing another molecule.
4. **Radiationless decay**. The excited molecule decays to its ground state by converting the excitation energy into heat.

Of all these responses to energy absorption, the most important in photosynthesis are resonance energy transfer and oxidation-reduction. Resonance energy transfer plays a critical role in the harvesting of light energy by accessory pigment molecules (**Figure 13.12a**). Eventually, the energy absorbed and transferred by light-harvesting complexes reaches the special chlorophyll a molecules P700 and P680, in the reaction centers of PSI and PSII, respectively. When these molecules are excited, they are referred to as P700* and P680*. Both P700* and 680* are electron donors because they can easily lose an electron to a specific acceptor molecule. P700* passes an electron to the electron acceptor A_0, a chlorophyll a molecule, and P680* transfers its electron to a molecule of pheophytin a. The electron hole left in the oxidized P700 and P680 is filled by an electron from a donor molecule. Plastocyanin and water play this role in PSI and PSII, respectively. Fluorescence also plays a role in photosynthesis when light absorption exceeds the capacity of the photosystems to transfer energy (**Figure 13.12b**). Then photons are reemitted by a protective mechanism.

KEY CONCEPTS

- The light energy absorbed by chromophores causes electrons to move to higher-energy levels.
- In photosynthesis, it is energy absorption that drives electron flow.

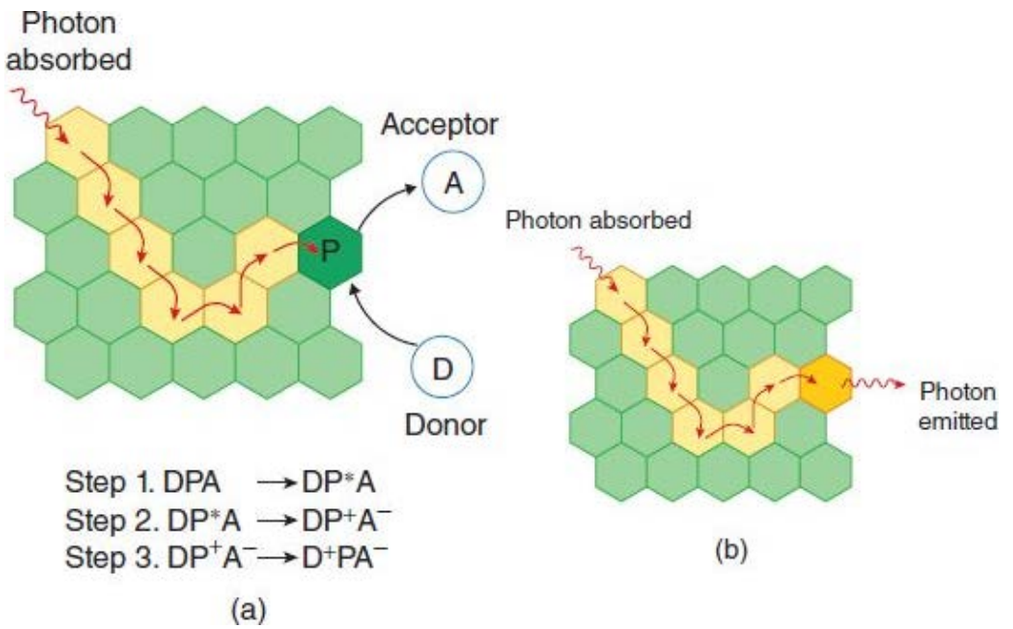


FIGURE 13.12

Energy Transfer in Photosystems

A photon absorbed by a chlorophyll molecule in the light-harvesting antenna promotes it to a singlet excited state. The excited chlorophyll molecule donates its energy to neighboring molecules by resonance energy transfer. The excitation randomly migrates through the antenna molecules (yellow hexagons) until it is trapped by the reaction center (dark green hexagon, P). (a) The excitation trap in the reaction center is a special chlorophyll molecule called the primary donor (P700 or P680), whose lowest excited state is lower than the antenna molecules. A molecule of the primary donor can be excited either by direct absorption of light energy or by transfer of the excitation from the nearest antenna molecule (step 1). The molecule of the primary donor in its excited state (P*) initiates the electron transfer in the reaction center by reducing the acceptor molecule (A) (step 2). The oxidized primary donor (P^+) extracts an electron from the nearest electron donor (D) (step 3). The electron from the reduced acceptor molecule (A^-) is transferred further along the chain of electron transfer carriers in the reaction center. (b) Under high-stress conditions such as high light intensity, reduced numbers of available reaction centers result in the loss of excitations through fluorescence.

QUESTION 13.3

Explain the observation that the different absorption spectra of antenna pigment molecules are different from those of P680 and P700.

13.3 LIGHT REACTIONS

In photosynthesis, the **light reactions** are a mechanism by which electrons are energized and subsequently used in ATP and NADPH synthesis. Species that evolve O_2 require both PSI and PSII. Species that live without oxygen can use either PSI- or PSII-like complexes. Working in series, the two photosystems couple the light-driven oxidation of water molecules to the reduction of $NADP^+$. The overall reaction is

$$2\,NADP^+ + 2H_2O \rightleftharpoons 2\,NADPH + O_2 + 2H^+ \qquad (1)$$

The standard reduction potentials for the half-reactions are

$$O_2 + 4e^- + 4H^+ \rightarrow 2H_2O \qquad E^{o\prime} = +0.816\ V$$

(2)

and

$$NADP^+ + H^+ + 2e^- \rightarrow NADPH \qquad E^{\circ\prime} = -0.320\ V \tag{3}$$

Therefore, the coupled process has a standard redox potential of −1.136 V. The minimum free energy change for this process (calculated from $\Delta G^{\circ\prime} = -nF\,\Delta E^{\circ\prime}$; see Section 9.1) is approximately 438 kJ (104.7 kcal) per mole of O_2 generated. In comparison, a mole of photons of 700-nm light provides approximately 170 kJ (40.6 kcal). Experimental observations have revealed that the absorption of eight or more photons (i.e., two photons per electron) is required for each O_2 generated. Consequently, a total of 1360 kJ (325 kcal) (i.e., eight times 170 kJ) is absorbed for each mole of O_2 produced. This energy is more than sufficient to account for reducing $NADP^+$ and to establish the proton gradient for ATP synthesis.

The process of light-driven photosynthesis begins with the excitation of PSII by light energy. One electron at a time is transferred to a chain of electron carriers that connects the two photosystems. As electrons are transferred from PSII to PSI, protons are pumped across the thylakoid membrane from the stroma into the thylakoid space. ATP is synthesized as protons flow back into the stroma through the ATP synthase. When P700 absorbs an additional photon, it releases an energized electron. This electron is immediately replaced by an electron provided by PSII. The newly energized PSI electron is passed through a series of iron–sulfur proteins and a flavoprotein to $NADP^+$, the final electron acceptor. This entire sequence, referred to as the **Z scheme**, is outlined in **Figure 13.13**.

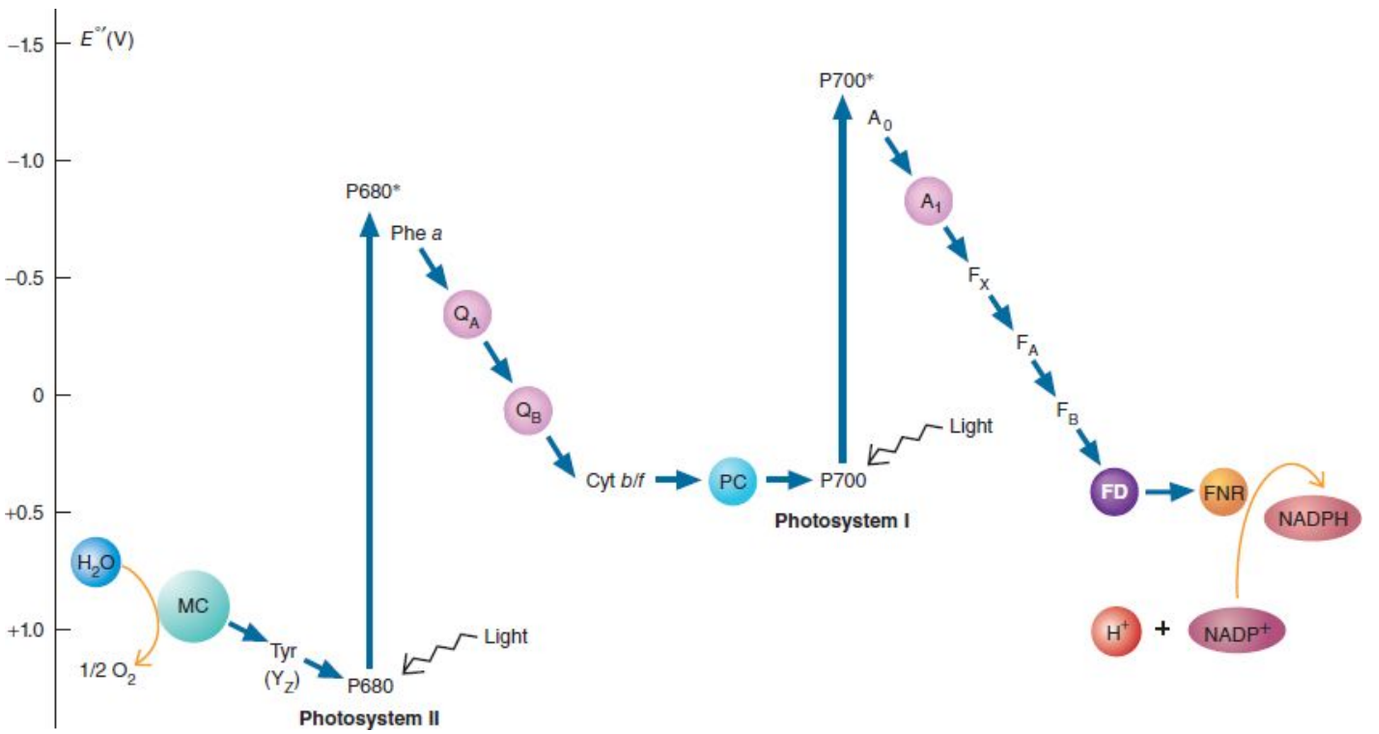


FIGURE 13.13

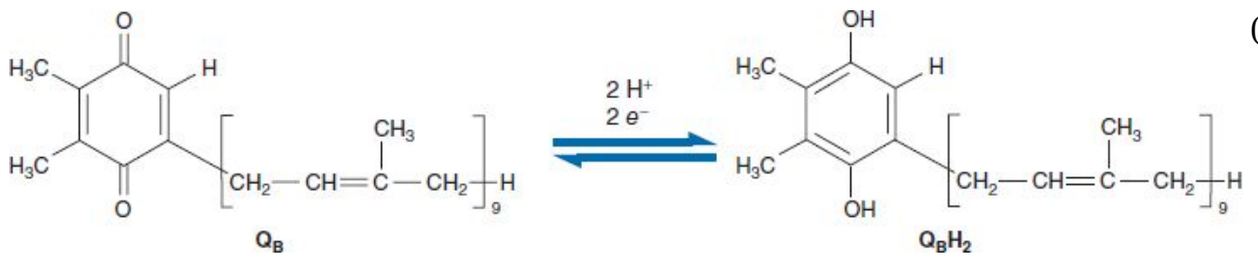
The Z Scheme

The flow of electrons from photosystem II to photosystem I drives the transport of protons into the thylakoid lumen. The precise mechanism of electron transfer through the iron–sulfur proteins Fe_A and Fe_B is not understood. The $E^{\circ\prime}$ values are approximate. (MC = manganese cluster, tyr (Y_Z) = tyr^{161} in D_1, PC = plastocyanin, A_0 = chlorophyll a, A_1 = phylloquinone, F_X, F_A, and F_B = a series of iron–sulfur clusters, FD =

ferredoxin, FNR = ferredoxin–$NADP^+$ oxidoreductase)

Photosystem II and Water Oxidation

When LHCII absorbs a photon, its energy is transferred to P680 in PSII; the newly energized electron is ejected and subsequently donated to *pheophytin a* (**Figure 13.14**), a molecule similar to chlorophyll in its structure. Reduced pheophytin a passes this electron to a chain of two electron carriers Q_A and Q_B (**Figure 13.14**). Although both molecules are plastoquinones, they perform different functions in PSII. Q_A, tightly bound to the protein, is a single electron carrier, and it never binds protons. Q_A transfers its electron to Q_B, which is loosely bound to the protein and can be doubly reduced receiving its electrons from Q_A one at a time upon binding two stromal protons. Reaction of a double reduction of Q_B is shown below:



(4)

The reduced Q_B (plastoquinol, Q_BH_2) is then released to the membrane pool of plastoquinones that donate electrons to cytochrome b_6f complex forming a transmembrane proton gradient.

An OEC, composed in part of the Mn_4CaO_5 cluster on the lumenal side of PSII and the tyrosine residue located on D_1, is responsible for the transfer of electrons from H_2O to oxidized P680 ($P680^+$). Recall that the excited state of P680 (P680*) reduces pheophytin a, thereby resulting in the formation of $P680^+$. The very high redox potential of this ion (+1.25 V) enables it to oxidize the tyrosine residue Y_z in D_1, which in turn extracts an electron from the manganese cluster. Tyrosine is effective in electron transfer because the tyrosyl radical formed is resonance stabilized.

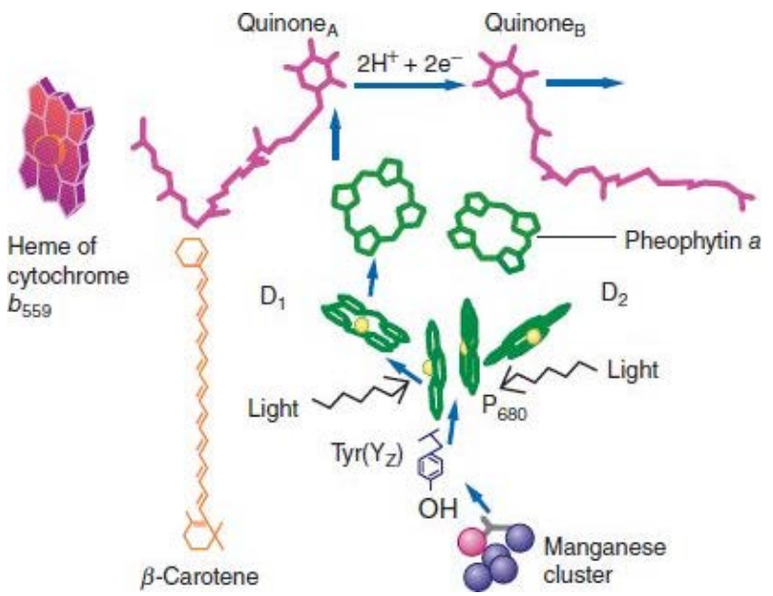


FIGURE 13.14

Photosystem II Electron Transport

Arrows indicate the electron transport pathway in PSII. A photon of light energizes an electron in P680. The

electron donated by P680 to pheophytin a is then transferred to the quinones. Note that the manganese cluster and the Tyr(Y_Z) side chain replace the electron in oxidized P680. Cytochrome b_{559} plays a role in PSII photodamage protection.

3D animation of pheophytin alpha

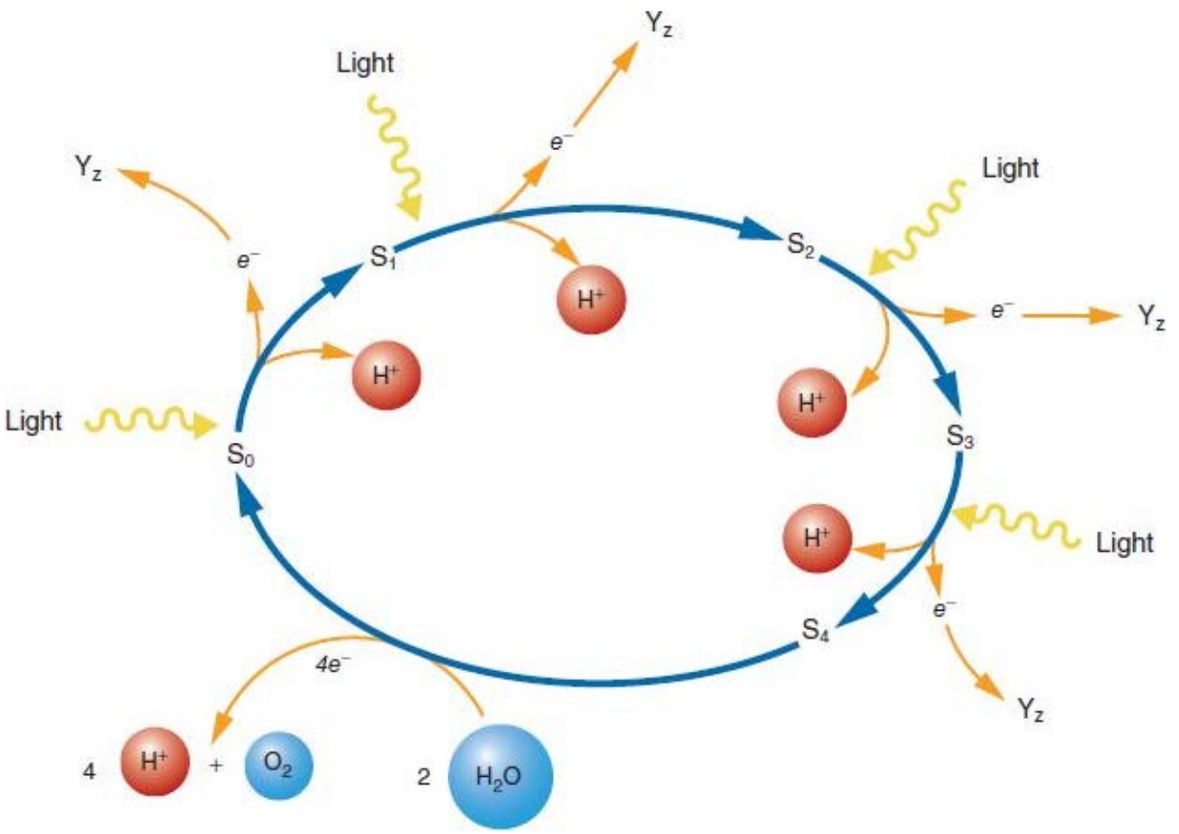


FIGURE 13.15

The Water-Oxidizing Clock

The absorption of four photons causes the abstraction of four electrons and four protons from two water molecules to yield O_2. The O_2-evolving complex has five oxidation states (S_0, S_1, S_2, S_3, and S_4), which represent different oxidation states of the Mn cluster. The sequential absorption of each of four photons drives the removal of an electron from a water molecule. Each electron is donated first to Y_Z, the tyrosine residue in D_1, and then to $P680^+$. Four protons created by the oxidation of the two water molecules are released into the thylakoid lumen.

The evolution of one O_2 requires splitting two H_2O molecules, which releases four protons and four electrons. Experimental evidence indicates that H_2O is converted to O_2 by a mechanism referred to as the *water-oxidizing clock* (**Figure 13.15**). The O_2-evolving complex has five oxidation states: S_0, S_1, S_2, S_3, and S_4. In the OEC, S_0 is the most reduced state, and S_4 is the most oxidized state. It is now believed that the Mn_4CaO_5 cluster, near the PSII reaction center, is responsible for these transitions. Oxygen–oxygen bond formation is the rate-limiting step in water oxidation. The oxygen-evolving complex also abstracts protons from H_2O as it cycles through the oxidation states. The protons are released into the thylakoid lumen, where they contribute to the pH gradient that drives ATP synthesis.

QUESTION 13.4

Excessive amounts of light can depress photosynthesis. Recent research indicates that PSII is extremely vulnerable to light damage. Plants often survive this damage because they possess efficient repair systems. Cells delete and resynthesize damaged components and recycle undamaged ones. For example, the D_1 polypeptide, apparently the most vulnerable component of PSII, is rapidly replaced after it is damaged. Review the role of PSII and suggest the proximate cause of light-induced damage of D_1. [*Hint*: The D_1/D_2 dimer binds two molecules of β-carotene.]

Photosystem I and NADPH Synthesis

The absorption of a photon by P700 leads to the release of an energized electron that is passed through a series of electron carriers (**Figure 13.16**). The first electron carrier is a chlorophyll a molecule (A_0). As the electron is donated sequentially to phylloquinone (A_1) and to several iron–sulfur proteins (the last of which is ferredoxin), it is moved from the lumenal surface of the thylakoid membrane to its stromal surface. Ferredoxin, a mobile, water-soluble protein, then donates each electron to a flavoprotein called ferredoxin–NADP oxidoreductase (FNR). The flavoprotein uses a total of two electrons and a stromal proton to reduce $NADP^+$ to NADPH. The transfer of electrons from ferredoxin to $NADP^+$ is referred to as the *noncyclic electron transport pathway* (**Figure 13.17**). In this pathway, the absorption of eight photons yields an ATP/NADPH ratio of 3:2. In *cyclic electron transport* (**Figure 13.18**), reduced ferredoxin donates its electrons to plastoquinone, which then passes them to the cyt b_f complex, plastocyanin, and eventually P700. In this process, which typically occurs when a chloroplast has a high $NADPH/NADP^+$ ratio, no NADPH is produced. Instead, electron transport results in the transfer of additional protons across the thylakoid membrane. As a result, additional molecules of ATP are synthesized.

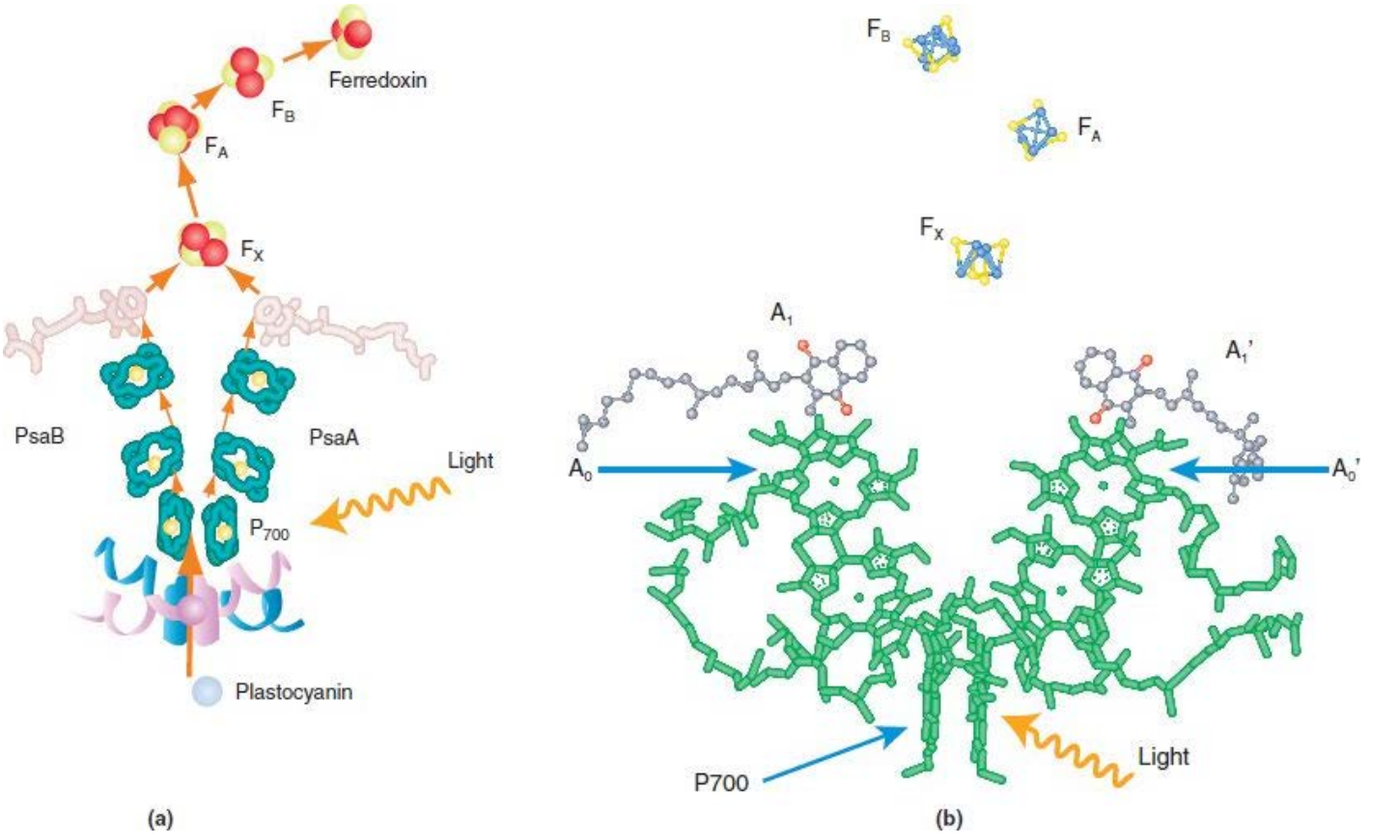


FIGURE 13.16

Two Views of Electron Transport in Photosystem I

(a) Electron carriers in the reaction center are bound to the PsaA and PsaB protein subunits of PSI and are in the same positions as **Figure 13.5**. (b) Excited P700 transfers an electron to the acceptor A_0, a chlorophyll a molecule, which then passes the electron to an electron carrier A_1. A_1 is a molecule of plastoquinone that transfers the electron to a series of iron–sulfur clusters (F_X, F_A, and F_B) and finally to ferredoxin. Note that there are two potential branches of the electron transfer from the reaction center toward F_X. Oxidized P700 obtains its missing electron from the nearby plastocyanin in a process that is facilitated by two tryptophan residues in the two polypeptide segments shown in (a) as pink and blue ribbon structures.

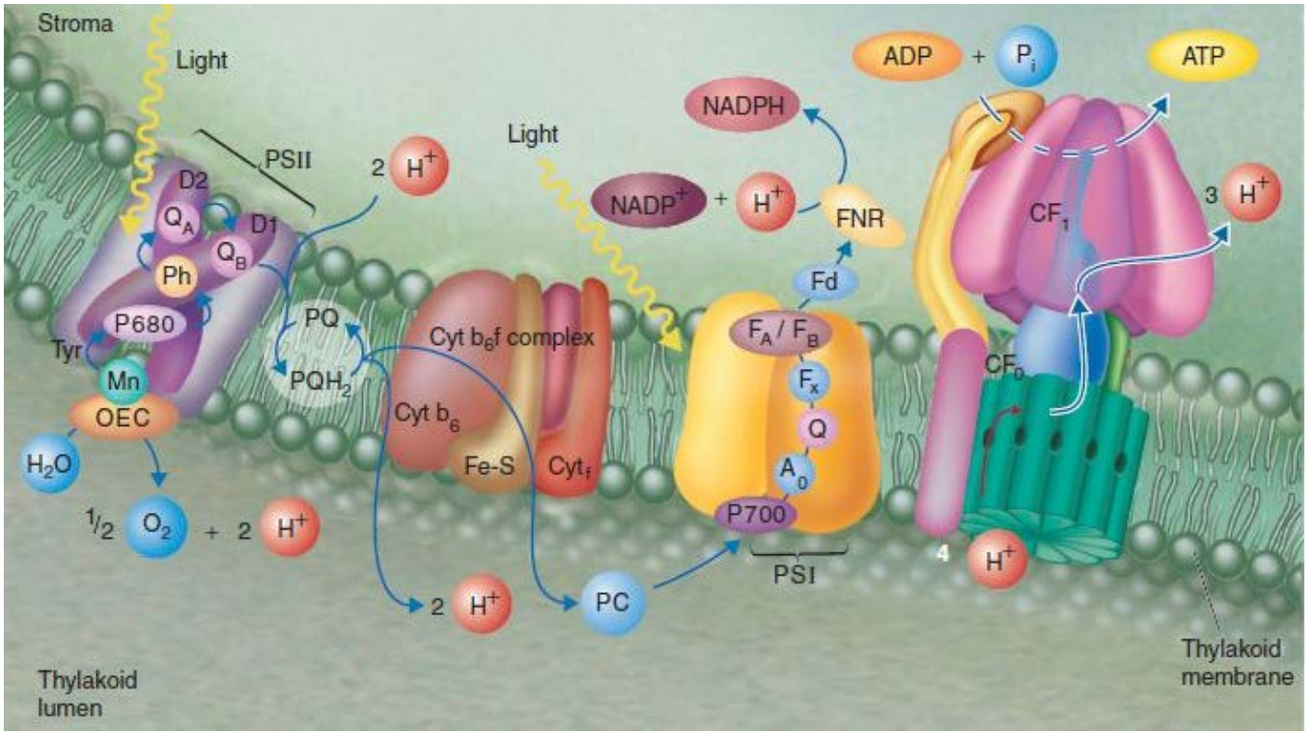


FIGURE 13.17

Membrane Organization of the Light Reactions in Chloroplasts: The Noncyclic Electron Transport Chain and the ATP Synthase Complex

In the noncyclic (or linear) electron transfer pathway, two electrons originating from water molecules are energized by PSII and then donated to Q_B (PQ) along with two protons from the stroma to yield Q_BH_2 (PQH_2). Once the reduced plastoquinone (plastoquinol) is released from PSII into the thylakoid membrane, it donates two electrons, one at a time, to cytochrome b_6f (**Figure 13.7**). Cytochrome b_6f transfers each electron to PC, which then catalyzes the transfer of electrons to PSI. Once donated electrons are reenergized by light within PSI, they pass through an electron transport chain and are eventually donated to $NADP^+$ in combination with one proton from the stroma to yield NADPH. For every two electrons removed from a water molecule and donated to $NADP^+$ (blue arrows), two H^+ are pumped from the stroma into the thylakoid lumen. Two additional H^+ are generated within the lumen by the oxygen-evolving complex. The flow of protons through the proton pore in CF_0 drives the synthesis of ATP in CF_1. (OEC = oxygen-evolving complex, Ph = pheophytin; Fd = ferredoxin; FNR = the flavoprotein ferredoxin–NADP oxidoreductase; PC = plastocyanin; Q = phylloquinone)

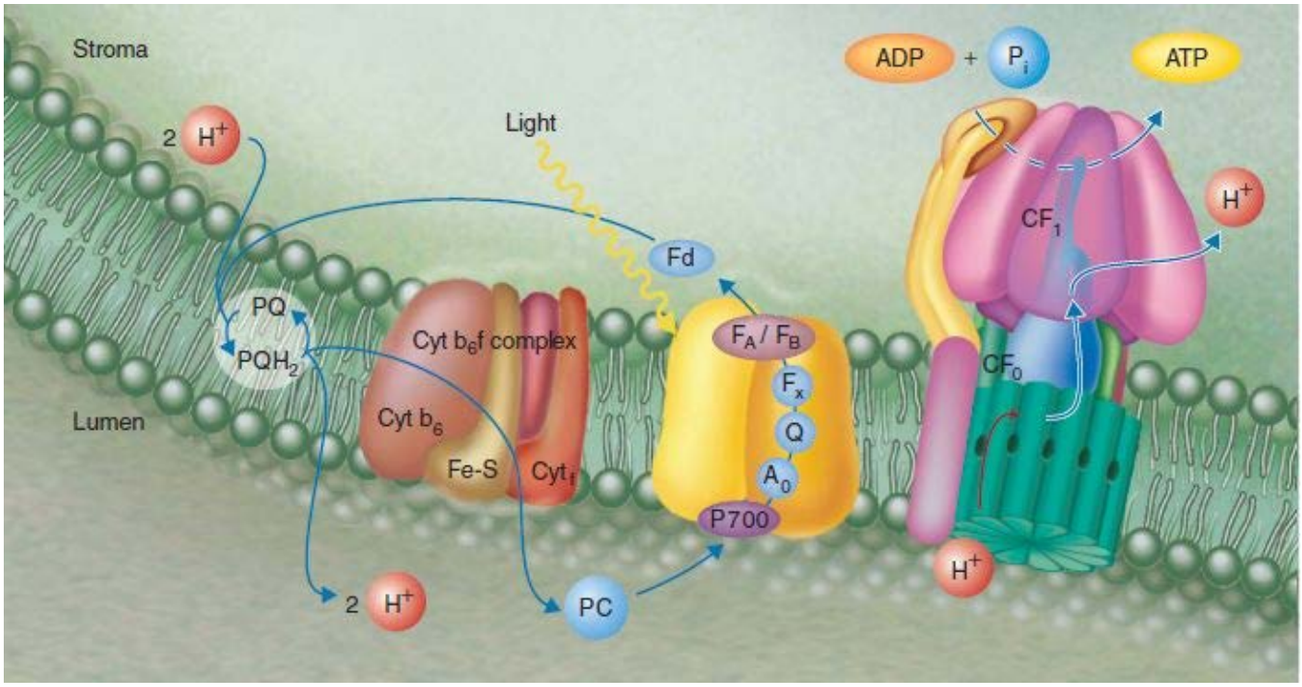


FIGURE 13.18

The Cyclic Electron Transport Pathway

The cyclic electron transport pathway involves the recycling of electrons from ferredoxin to PQ. As a result, two protons are pumped from the stroma into the lumen for each electron transported. The proton flow drives ATP synthesis. No NADPH is produced.

The Calvin cycle (p. 510) requires an ATP/NADPH ratio of 3:2. However, ATP is also used for processes other than carbohydrate synthesis. Consequently, both noncyclic and cyclic photophosphorylation pathways are required for sufficient ATP synthesis during photosynthesis.

KEY CONCEPTS

- Eukaryotic photosynthesizing cells possess two photosystems, PSI and PSII, which are connected in series in a mechanism referred to as the Z scheme.
- The water-oxidizing clock component of PSII generates O_2.
- The protons are used in the synthesis of ATP in a chemiosmotic mechanism.
- PSI is responsible for the synthesis of NADPH.

WORKED PROBLEM 13.1

Calculate $\Delta G°'$ for the four-electron oxidation of H_2O by $NADP^+$ in the light reactions.

SOLUTION

The overall reaction is

$$2\ H_2O + 2\ NADP^+ \rightarrow O_2 + 2\ NADPH + 2\ H^+$$

The reduction potentials ($\Delta E°'$) for the two half-reactions are

$$1/2\ O_2 + 2\ H^+ + 2e^- \rightarrow H_2O\ (\Delta E^{\circ\prime} = +0.82\ V)$$
$$NADP^+ + H^+ + 2e^- \rightarrow NADPH + H^+\ (\Delta E^{\circ\prime} = -0.32\ V)$$

$\Delta G^{\circ\prime}$ is calculated using the equation $\Delta G = -nF\Delta E^{\circ\prime}$.
Substituting the $\Delta E^{\circ\prime}$ values for the two half reactions

$$\begin{aligned}\Delta G^{\circ\prime} &= -4\ (96.5\ kJ/V \cdot mol)\ [-0.32\ V - (0.82\ V)] \\ &= (386\ kJ/V \cdot mol)\ (-1.14\ V) \\ &= -440\ kJ/mol\end{aligned}$$

■

QUESTION 13.5

Describe the role of each of the following molecules in photosynthesis:

a. plastocyanin
b. β-carotene
c. ferredoxin
d. plastoquinone
e. pheophytin a
f. lutein

Photophosphorylation

During photosynthesis, light energy captured by an organism's photosystems is transduced into ATP phosphate bond energy. This conversion is referred to as **photophosphorylation**. It is apparent from the preceding discussions that there are many similarities between mitochondrial and chloroplast ATP synthesis. For example, many of the same molecules and terms that are encountered in aerobic respiration (Chapter 10) are also relevant to discussions of photosynthesis. Although there are a variety of differences between aerobic respiration and photosynthesis, the essential difference between the two processes is the conversion of light energy into redox energy by chloroplasts. (Recall that mitochondria produce redox energy by extracting high-energy electrons from food molecules.) Another critical difference involves the permeability characteristics of mitochondrial inner membrane and thylakoid membrane. In contrast to the inner membrane of mitochondria, the thylakoid membrane is permeable to Mg^{2+} and Cl^-. Therefore, the movement of Mg^{2+} and Cl^- across the thylakoid membrane dissipates the electrical potential as protons are transported across the membrane during the light reaction. The electrochemical gradient across the thylakoid membrane that drives ATP synthesis therefore consists mainly of a proton gradient that may be as great as 3.5 pH units.

Experimental measurements of H^+/ATP ratios indicate that the movement across the thylakoid membrane of about 12 protons in noncyclic photophosphorylation yields three molecules of ATP. The synthesis of these ATPs is made possible by the absorption of eight photons, one for each of the electrons from two water molecules. Proton transport occurs as these electrons are transported down the noncyclic electron transport system (**Figure 13.17**). (The measured difference between the proton/ATP ratio of chloroplasts ($4H^+$/ATP) and mitochondria ($3H^+$/ATP) is explained in part by a structural difference in the ATP synthase proton channels of the two organelles. More protons are required for a 360° turn of the chloroplast C_0 proton channel complex because it consists of a larger number of subunits than does the mitochondrial proton channel complex.) In cyclic

photophosphorylation, the pumping of eight protons by the cyt b_6f complex, as the result of the absorption of four photons, yields two molecules of ATP.

QUESTION 13.6

A variety of herbicides kill plants by inhibiting photosynthetic electron transport. Atrazine, a triazine herbicide, blocks electron transport between Q_A and Q_B in PSII. The compound 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) also blocks electron flow between the two molecules of plastoquinone. Paraquat is a member of a family of compounds called bipyridylium herbicides. Paraquat is reduced by PSI but is easily reoxidized by O_2 in a process that produces superoxide and hydroxyl radicals. Plants die because their cell membranes are destroyed by radicals. Of the herbicides just discussed, determine which, if any, are most likely to be toxic to humans and other animals. What specific damage may occur?

13.4 THE LIGHT-INDEPENDENT REACTIONS

Within the chloroplast stroma, CO_2 is incorporated into carbohydrate by eukaryotic photosynthesizing organisms in a process commonly known as the **Calvin cycle**. Because the reactions of the Calvin cycle can occur without light if sufficient ATP and NADPH are supplied, they have often been called the *dark reactions*. The term is somewhat misleading, however. The Calvin cycle reactions typically occur only when the plant is illuminated because ATP and NADPH are produced by the light reactions. Therefore, **light-independent reactions** is a more appropriate term. Because of the types of reaction that occur in the Calvin cycle, it is also referred to as the *reductive pentose phosphate cycle* (RPP cycle) and the *photosynthetic carbon reduction cycle* (PCR cycle).

The Calvin Cycle

The net equation for the Calvin cycle (**Figure 13.19**) is

$$3\,CO_2 + 6\,NADPH + 9\,ATP \rightarrow \text{glyceraldehyde-3-phosphate} + 6\,NADP^+ + 9\,ADP + 8\,P_i \quad (5)$$

For every three molecules of CO_2 that are incorporated into carbohydrate molecules, there is a net gain of one molecule of glyceraldehyde-3-phosphate. The fixation of six CO_2 into glucose occurs at the expense of 12 NADPH and 18 ATP. The reactions of the cycle can be divided into three phases: carbon fixation, reduction, and regeneration.

CARBON FIXATION **Carbon fixation**, the mechanism by which inorganic CO_2 is incorporated into organic molecules, consists of a single reaction. Ribulose-1,5-bisphosphate carboxylase (Rubisco) is a Mg^{2+}-requiring enzyme that catalyzes the carboxylation of ribulose-1,5-bisphosphate to form two molecules of glycerate-3-phosphate (reaction 1 in **Figure 13.20**). Plants that produce glycerate-3-phosphate as the first stable product of photosynthesis are referred to as **C3 plants**. (An alternative photosynthetic process is described on pages 515–17.) Rubisco, composed of eight large (L 54 kDa) subunits and eight small (S 14 kDa) subunits, is the pacemaker enzyme of the Calvin cycle. Each L subunit contains an active site that binds substrate. The catalytic activity of the L subunits is enhanced by the S subunits. Because the CO_2 fixation reaction is extremely slow, plants compensate by producing a large number of copies of the enzyme, which often constitutes approximately half of

a leaf's soluble protein. For this reason, rubisco is often described as the world's most abundant enzyme.

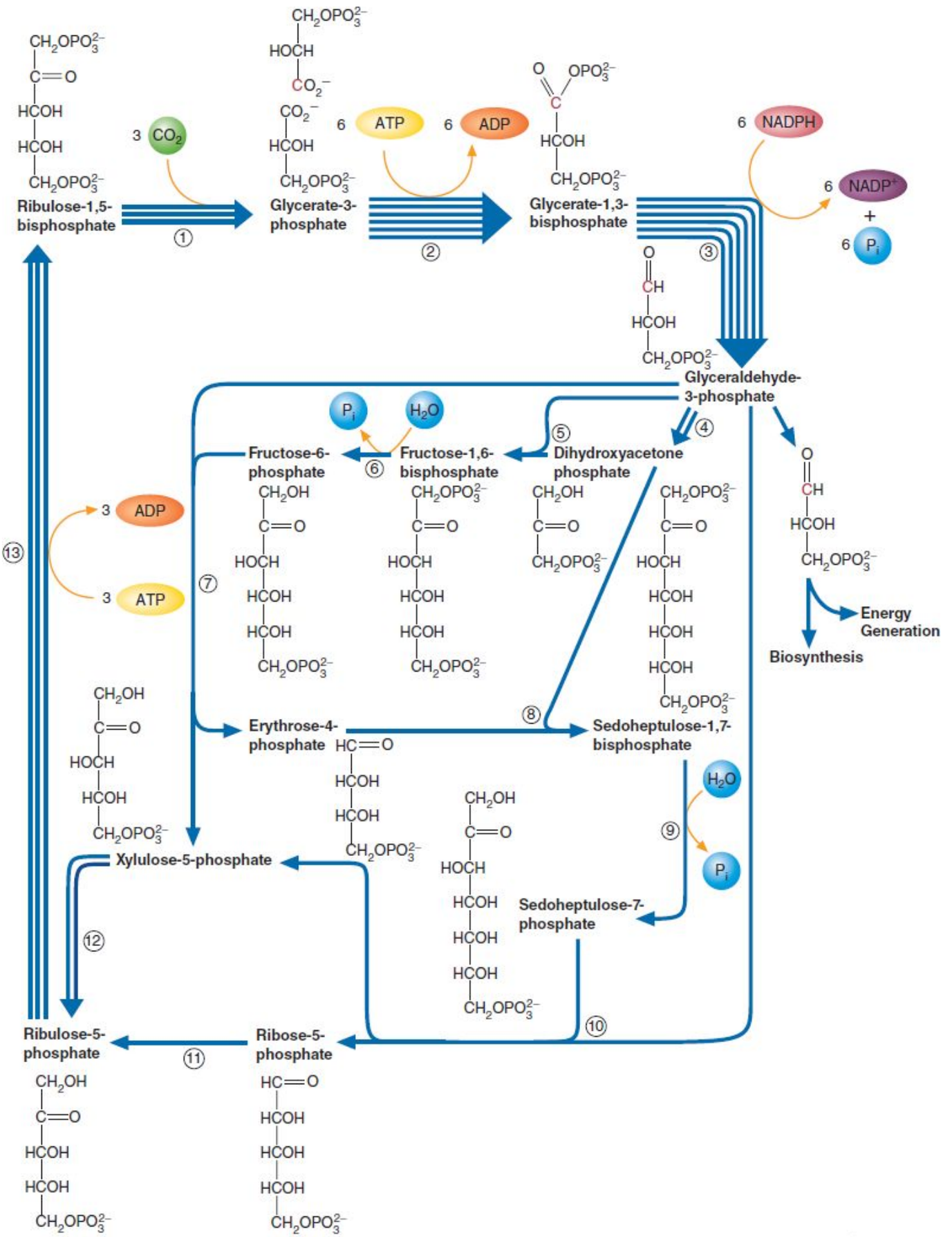


FIGURE 13.19

The Calvin Cycle

Rubisco catalyzes the carboxylation of ribulose-1,5-bisphosphate to yield glycerate-3-phoshate (reaction 1). For every three ribulose-1,5-bisphosphate molecules that are carboxylated in the Calvin cycle, there is a net gain of one glyceraldehyde-3-phosphate product. The other five glyceraldehyde-3-phosphate molecules are used to regenerate three ribulose-1,5-bisphosphate molecules (reactions 4-13). For reactions 4 through 13, the number of arrows drawn at each step indicates the number of molecules proceeding through that step for every three CO_2 molecules that enter the cycle. Enzymes: 1 = ribulose bisphosphate carboxylase (rubisco), 2 = phosphoglycerate kinase, 3 = $NADP^+$-glyceraldehyde-3-phosphate dehydrogenase, 4 = triosephosphate isomerase, 5 = aldolase, 6 = fructose-1,6-bisphosphatase, 7 = transketolase, 8 = aldolase, 9 = sedoheptulose-1,7-bisphosphatase, 10 = transketolase, 11 = ribose-5-phosphate isomerase, 12 = ribulose-5-phosphate epimerase, 13 = ribulose-5-phosphate kinase.

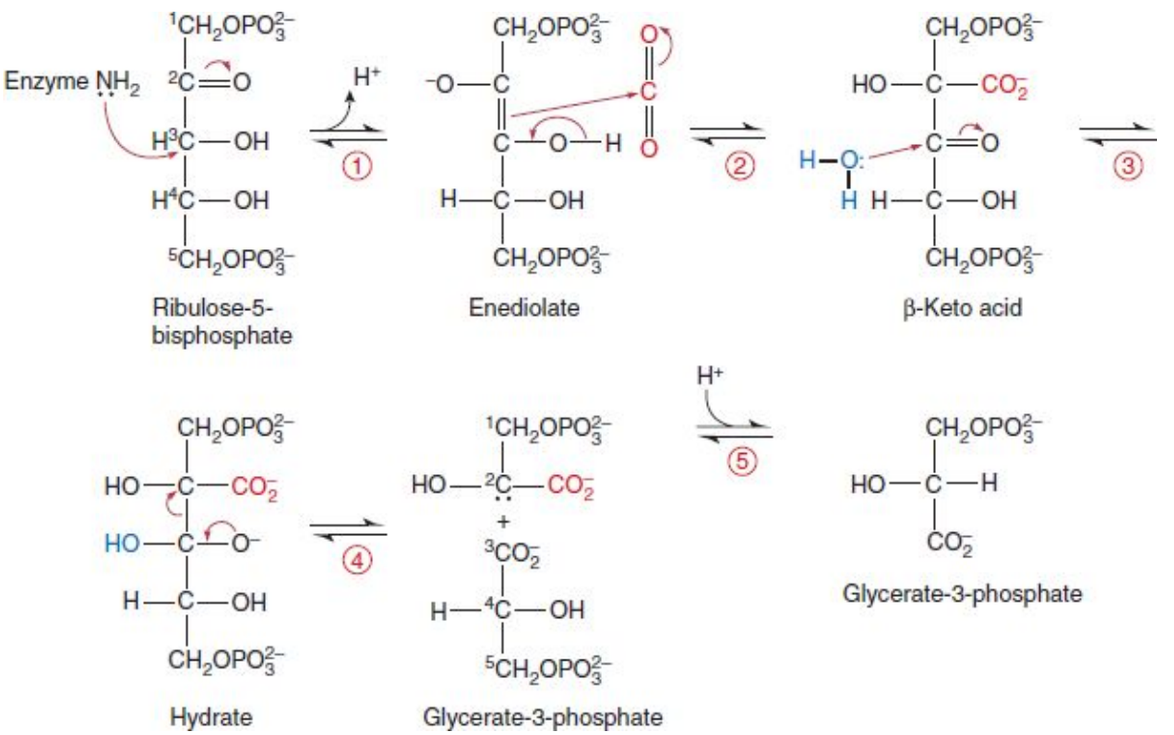


FIGURE 13.20

The Rubisco Carboxylation Mechanism

(1) The C-3 proton, made more acidic by its proximity to Mg^{2+}, is removed by a lysine side chain to yield an enediolate. (2) The enediolate attacks a CO_2 polarized by Mg^{2+} to form a six-carbon β-keto acid. (3) A water molecule then attacks the carbonyl carbon of the β-keto acid to yield a hydrated intermediate. (4) The hydrated intermediate is rapidly cleaved into two three-carbon products: a glycerate-3-phosphate anion and glycerate-3-phosphate. (5) The protonation of the anion by another lysine side chain (not shown) yields a second molecule of glycerate-3-phosphate.

REDUCTION In the reduction phase of the Calvin cycle, glycerate-3-phosphate is reduced to glyceraldehyde-3-phosphate. In the first of two reactions, six molecules of glycerate-3-phosphate are phosphorylated at the expense of six ATP molecules to form six glycerate-1,3-bisphosphate molecules (reaction 2). These latter molecules are then reduced by $NADP^+$-glyceraldehyde-3-phosphate dehydrogenase to form six molecules of glyceraldehyde-3-phosphate (reaction 3). These reactions are similar to reactions encountered in gluconeogenesis (pp. 301–09). Unlike the dehydrogenase in gluconeogenesis, the Calvin cycle enzyme uses NADPH as a reducing agent.

REGENERATION Several regenerative phase reactions are similar to those of other biochemical

pathways. Two reactions each are catalyzed by aldolase (glycolysis) and transketolase (pentose phosphate pathway). Fructose-1,6-bisphosphatase is a gluconeogenic enzyme. As noted previously, the net production of fixed carbon in the Calvin cycle is one molecule of glyceraldehyde-3-phosphate. The other five glyceraldehyde-3-phosphate molecules are processed in the remainder of the Calvin cycle reactions to regenerate three molecules of ribulose-1,5-bisphosphate.

The regeneration of ribulose-1,5-bisphosphate begins with reactions involving glyceraldehyde-3-phosphate. In reaction 4, two molecules of glyceraldehyde-3-phosphate are isomerized to form two dihydroxyacetone phosphate. Aldolase catalyzes the condensation of one of these dihydroxyacetone molecules with a third glyceraldehyde-3-phosphate to form fructose-1,6-bisphosphate (reaction 5). The latter molecule is then hydrolyzed by fructose-1,6-bisphosphatase (reaction 6) to fructose-6-phosphate. Fructose-6-phosphate subsequently combines with a fourth molecule of glyceraldehyde-3-phosphate in a transketolase-catalyzed reaction to form xylulose-5-phosphate and erythrose-4-phosphate (reaction 7). In reaction 8, aldolase catalyzes the condensation of erythrose-4-phosphate with the second molecule of DHAP to form sedoheptulose-1,7-bisphosphate, which is then hydrolyzed to form sedoheptulose-7-phosphate (reaction 9). Transketolase catalyzes the reaction of the fifth molecule of glyceraldehyde-3-phosphate with sedoheptulose-7-phosphate to form ribose-5-phosphate and a second molecule of xylulose-5-phosphate (reaction 10). Ribose-5-phosphate and both molecules of xylulose-5-phosphate are separately isomerized (reactions 11 and 12) to ribulose-5-phosphate. In the last step, three molecules of ribulose-5-phosphate are phosphorylated at the expense of three ATP molecules by ribulose-5-phosphate kinase (reaction 13) to form three molecules of ribulose-1,5-bisphosphate. The remaining molecule of glyceraldehyde-3-phosphate is either used within the chloroplast in starch synthesis or exported to the cytoplasm, where it may serve in the synthesis of sucrose or other metabolites.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on starch and sucrose metabolism.

QUESTION 13.7

When plant cells are illuminated, their cytoplasmic ATP/ADP and NADH/NAD^+ ratios rise significantly. The following shuttle mechanism is believed to contribute to the transfer of ATP and reducing equivalents from the chloroplast into the cytoplasm. Once DHAP has been transported from the stroma into the cytoplasm, it is converted to glyceraldehyde-3-phosphate and then to glycerate-1,3-bisphosphate. (This reaction is the reverse of the reaction in which glyceraldehyde-3-phosphate is formed during carbon fixation.) In the cytoplasmic reaction, the reducing equivalents are donated to NAD^+ to form NADH. In a later reaction, glycerate-1,3-bisphosphate is converted to glycerate-3-phosphate with the concomitant production of one molecule of ATP. Glycerate-3-phosphate is then transported back into the chloroplast, where it is reconverted to glyceraldehyde-3-phosphate.

This shuttle somewhat depresses mitochondrial respiration processes. Review the regulation of aerobic respiration (Chapter 9) and suggest how photosynthesis suppresses this aspect of mitochondrial function.

WORKED PROBLEM 13.2

Glyceraldehyde-3-phosphate, the first product of the Calvin cycle, is used to synthesize the energy storage molecules starch and sucrose. Outline the pathway by which two glyceraldehyde-3-phosphate molecules are incorporated into starch and calculate the cost in ATP molecules. Note that the precursor molecule in starch synthesis is ADP-glucose. Compare the cost of incorporating a glucose molecule into starch with that of its degradation into CO_2 and H_2O. Assume that the net production of glucose catabolism is 30 ATP.

SOLUTION

One molecule each of glyceraldehyde-3-phosphate and its isomer DHAP are converted by aldolase into fructose-1,6-bisphosphate. The latter molecule is converted via the gluconeogenic enzyme fructose-bisphosphate phosphatase into fructose-6-phosphate, which is isomerized to form glucose-6-phosphate. Glucose-1-phosphate, formed from glucose-6-phosphate by phosphoglucomutase, is then converted into ADP-glucose by ADP-glucose pyrophosphorylase. Therefore, only one ATP is used to synthesize ADP-glucose from two glyceraldehyde-3-phosphate. The cost of the incorporation of these two molecules into starch compared with the energy released by glucose is 1/30 or 3.3% of the total number of ATPs that can be generated by glucose catabolism. ■

Photorespiration

Photorespiration (**Figure 13.21**) is perhaps the most curious feature of photosynthesis. In this light-dependent process, oxygen is consumed and CO_2 is liberated by plant cells that are actively engaged in photosynthesis. Photorespiration is a multistep mechanism initiated by ribulose bisphosphate carboxylase, which also possesses an oxygenase activity. (For this reason, the name *ribulose-1,5-bisphosphate carboxylase-oxygenase*, or *rubisco*, is sometimes used.) Both CO_2 and O_2 compete for rubisco's active site.

In the oxygenation reaction, ribulose-1,5-bisphosphate is converted to glycolate-2-phosphate (**Figure 13.22**) and glycerate-3-phosphate. Glycolate-2-phosphate is hydrolyzed to form glycolate, which is then oxidized by O_2 to form glyoxylate and H_2O_2. Glyoxylate is converted through a series of reactions (outlined in **Figure 13.21**) to glycerate-3-phosphate. Glycerate-3-phosphate then enters the Calvin cycle, where it is converted to ribulose-1,5-bisphosphate. Photorespiration is a wasteful process. It loses fixed carbon (as CO_2), and consumes both ATP and NADH.

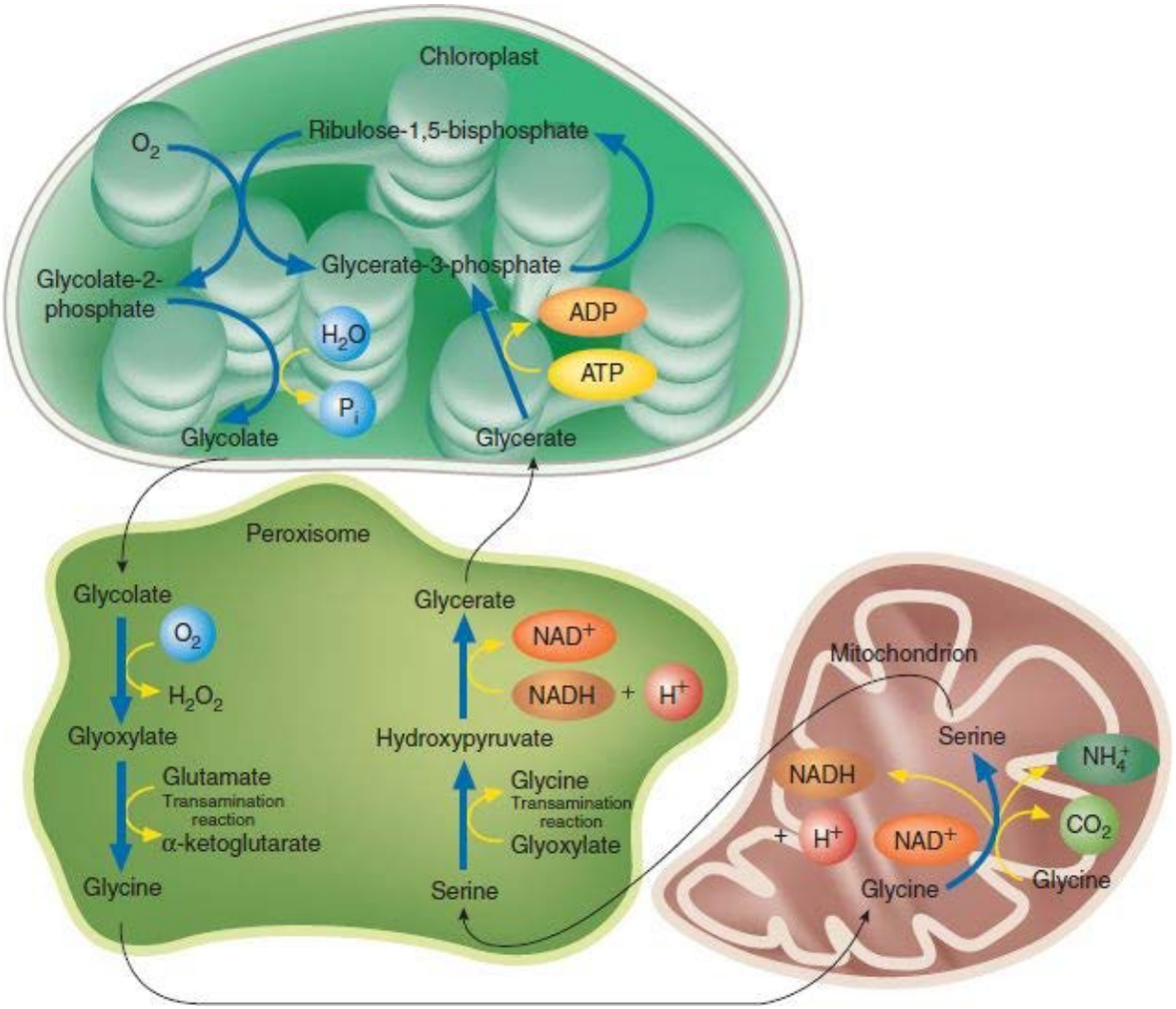
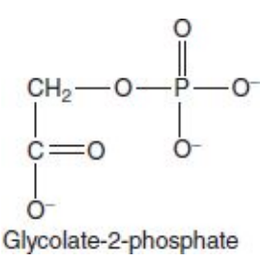


FIGURE 13.21

Photorespiration

Photorespiration is a wasteful process that occurs because rubisco can catalyze, under certain conditions, the oxidation of ribulose-1,5-bisphosphate by O_2 to form glycolate-2-phosphate. Enzymes in three organelles (chloroplasts, peroxisomes, and mitochondria) catalyze this multistep pathway, which is a mechanism for salvaging fixed carbon from glycolate-2-phosphate. After glycolate-2-phosphate is hydrolyzed in chloroplast stroma, the product glycolate is transferred to a peroxisome, where it reacts with O_2 to form glyoxylate and H_2O_2. Glyoxylate then undergoes a transamination reaction to form glycine, which is then transferred out of the peroxisome and into a mitochondrion. Within the mitochondrial matrix two glycine molecules are converted into serine, CO_2, and NH_3. The NAD^+ produced in this reaction is reduced to form NADH in a series of reactions catalyzed by enzymes in the glycine decarboxylase complex. Serine then returns to the peroxisome, where it undergoes a transamination reaction involving glyoxylate to form hydroxypyruvate. Hydroxypyruvate is subsequently reduced by NADH to yield glycerate. Once glycerate enters a chloroplast, it reacts with ATP to yield glycerate-3-phosphate, the Calvin cycle intermediate.



Glycolate-2-phosphate

FIGURE 13.22

Structure of Glycolate-2-Phosphate

The rate of photorespiration depends on several parameters, including the concentrations of CO_2 and O_2 to which photosynthesizing cells are exposed. Photorespiration is depressed by CO_2 concentrations above 0.2%. (Because photorespiration and photosynthesis occur concurrently, CO_2 is released during CO_2 fixation. When the rates of CO_2 release and fixation are equal, the *CO_2 compensation point* has been reached. The lower the CO_2 compensation point, the less photorespiration takes place. Many C3 plants have CO_2 compensation points between 0.02% and 0.03% of CO_2 in the air near photosynthesizing cells.) In contrast, high O_2 concentrations and high temperatures promote photorespiration. Consequently, this process is favored when plants are exposed to high temperatures and any condition that causes low CO_2 and/or high O_2 concentrations. For example, photorespiration is a serious problem for C3 plants in hot, dry environments. To conserve water, these plants close their stomata, thus reducing the CO_2 concentration within leaf tissue. (*Stomata* are pores on the surface of leaves. When they are open, CO_2, O_2, and H_2O vapor can readily diffuse down the concentration gradients between the leaf's interior and the external environment.) As photosynthesis continues, O_2 levels increase. Depending on the severity of the circumstances, 30 to 50% of a plant's yield of fixed carbon may be lost. This effect can be serious because several C3 plants (e.g., soybeans and oats) are major food crops.

Photorespiration is an artifact of the evolutionary history of photosynthesis. In the early atmosphere in which the first photosystem evolved, oxygen levels were very low. Thus, over the long time period before oxygen levels became a problem, there was no selection pressure to improve the capacity of the rubisco active site to distinguish between CO_2 and O_2. Selection pressure could occur only when oxygen levels increased significantly. It is noteworthy that selection for CO_2 over O_2 is higher in modern green plants than in bacteria. The pathway that evolved to convert glycolate-2-phosphate to glycerate-3-phosphate, though costly in ATP and NADH consumption, is viewed as a salvage operation that recovers previously fixed and partly reduced carbon. C4 plants, which have developed an elaborate mechanism to suppress photorespiration, are described next.

KEY CONCEPTS

- The Calvin cycle is a series of light-independent reactions in which CO_2 is incorporated into organic molecules.
- The Calvin cycle reactions occur in three phases: carbon fixation, reduction, and regeneration.
- Photorespiration is a wasteful process in which photosynthesizing cells evolve CO_2.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on crassulacean metabolism.

Alternatives to C3 Metabolism

In addition to C3 photosynthesis, which is used by most plants, there are two other mechanisms for fixing CO_2: C4 metabolism and crassulacean acid metabolism. Both improve the efficiency of photosynthesis in climates where temperatures are high and water is scarce. Crassulacean acid metabolism, a plant adaptation that conserves water in regions with high light intensity and drought, is described in an online reading, Crassulacean Acid Metabolism.

C4 METABOLISM C4 plants include sugarcane and maize (corn); they thrive in the tropics and can successfully tolerate conditions of drought and high temperatures. The name **C4 plants** indicates the

prominent role of a four-carbon molecule (oxaloacetate [OAA]) in a biochemical pathway that avoids photorespiration. This pathway is called **C4 metabolism**, the *C4 pathway*, or the *Hatch–Slack pathway* (after its discoverers).

The leaves of C4 plants possess two types of photosynthesizing cells: mesophyll cells and bundle sheath cells. (In C3 plants, photosynthesis occurs in mesophyll cells.) Most mesophyll cells in both plant types are positioned so that they are in direct contact with air when the leaf's stomata are open. In C4 plants, CO_2 is captured in specialized mesophyll cells where it is converted into bicarbonate and then incorporated into oxaloacetate (**Figure 13.23**). PEP carboxylase catalyzes this reaction, which is an indirect means of carbon fixation. Since PEP carboxylase has a lower K_m for CO_2 (i.e., a higher affinity) than rubisco and O_2 is a poor substrate, C4 plants are more effective at capturing CO_2 than are C3 plants. Once formed, OAA is reduced to malate, which then diffuses into bundle sheath cells. As the name implies, bundle sheath cells form a layer around vascular bundles, which contain phloem and xylem vessels. Unlike C3 plants, the bundle sheath cells of most C4 plants possess chloroplasts.

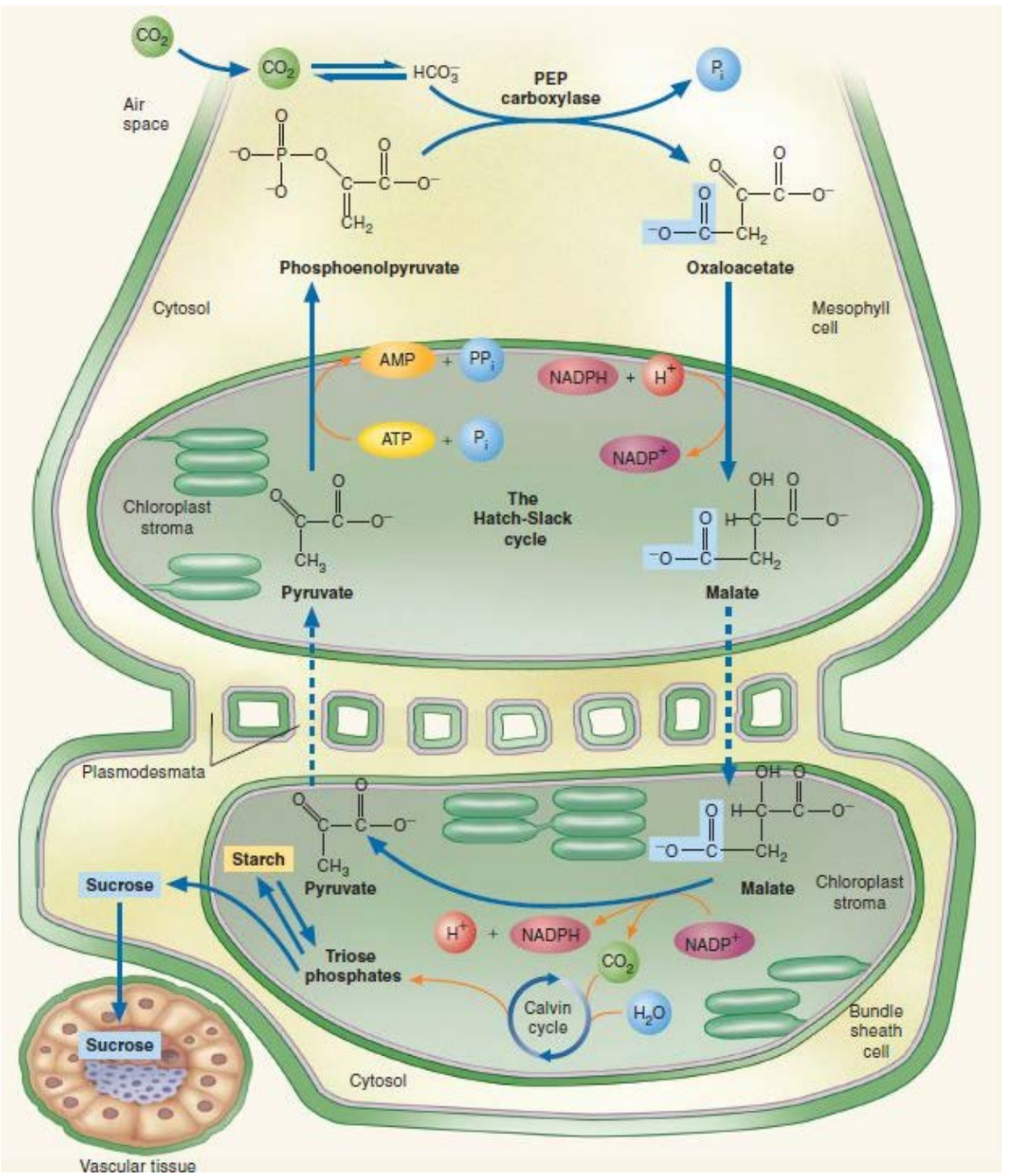


FIGURE 13.23

C4 Metabolism

In the C4 pathway, mesophyll cells, which are in direct contact with the air space in the leaf, take up CO_2 and use it to synthesize oxaloacetate, which is then reduced to malate. (Some C4 plants synthesize aspartate instead of malate.) Malate then diffuses to bundle sheath cells, where it is reconverted to pyruvate. The CO_2 released in this reaction is used in the Calvin cycle, eventually yielding triose phosphate molecules. Triose phosphate is subsequently converted to starch or sucrose. Pyruvate returns to the mesophyll cell where it is converted to phosphoenolpyruvate. Starch is synthesized from glucose-1-phosphate, which is subsequently converted to ADP-glucose by ADP-glucose pyrophosphorylase. ADP-glucose molecules are then incorporated into a preexisting polysaccharide chain by starch synthase. Sucrose-6-phosphate is synthesized from UDP-glucose and fructose-6-phosphate by sucrose phosphate synthase. Sucrose phosphatase catalyzes the hydrolysis of sucrose-6-phosphate to form sucrose and P_i. Note that the term triose phosphates refers to glyceraldehyde-3-phosphate and dihydroxyacetone phosphate.

In the bundle sheath cells, malate is decarboxylated to pyruvate in a reaction that reduces $NADP^+$ to NADPH. The pyruvate product of this latter reaction diffuses back to a mesophyll cell, where it can be reconverted to PEP. Although this reaction is driven by the hydrolysis of one molecule of ATP, there is a net cost of two ATP molecules. An additional ATP molecule is required to convert the AMP product to ADP so that it can be rephosphorylated during photosynthesis. This circuitous process delivers CO_2 and NADPH to the chloroplasts of bundle sheath cells, where rubisco and the other enzymes of the Calvin cycle use them to synthesize triose phosphates. The concentrations of CO_2 available to rubisco in the bundle sheath cells of C4 plants are significantly higher (10–20 times as great) than in C3 plants. C4 plants also use water more efficiently than C3 plants because they can close stomata when ambient temperature is high, thereby reducing transpiration.

13.5 REGULATION OF PHOTOSYNTHESIS

Plants must adapt to a wide variety of environmental conditions. The regulation of photosynthesis is, therefore, complicated. Although the control of most photosynthetic processes is far from being completely understood, several control features are well established. Most of these processes are directly or indirectly controlled by light. A brief description of general light-related effects is followed by a discussion of how the activity of rubisco—the key regulatory enzyme in photosynthesis—is controlled.

Light Control of Photosynthesis

Investigations of photosynthesis are complicated by several factors. The most prominent of these is that the photosynthetic rate depends on temperature and cellular CO_2 concentration, as well as on light. Nevertheless, numerous investigations have firmly established light as an important regulator of most aspects of photosynthesis.

Many effects of light on plants are mediated by changes in the activities of key enzymes. Because plant cells possess enzymes that operate in several competing pathways (i.e., glycolysis, pentose phosphate pathway, and the Calvin cycle), careful metabolic regulation is critical. Light assists in this regulation by activating certain photosynthetic enzymes and deactivating several enzymes in degradative pathways. Among the light-activated enzymes are ribulose-l,5-bisphosphate carboxylase, $NADP^+$-glyceraldehyde-3-phosphate dehydrogenase, fructose-1,6-bisphosphatase, sedoheptulose-1,7-bisphosphatase, and ribulose-5-phosphate kinase. Light-inactivated enzymes include phosphofructokinase and glucose-6-phosphate dehydrogenase.

Light affects enzymes by indirect mechanisms. Among the best researched are the following.

1. **pH**. During the light reactions, protons are pumped across the thylakoid membrane from the stroma into the thylakoid lumen. As the pH of the stroma increases from 7 to approximately 8, the activities of several enzymes are affected. For example, the pH optimum of ribulose-1,5-bisphosphate carboxylase is 8.

2. **Mg^{2+}**. Several photosynthetic enzymes (e.g., fructose-1,6-bisphosphatase) are activated by Mg^{2+}. Light induces an increase in the stromal Mg^{2+} concentration from 1 to 3 mM to about 3 to 6 mM. (Recall that Mg^{2+} moves across the thylakoid membrane into the stroma during the light reactions.)

3. **The ferredoxin–thioredoxin system**. Thioredoxins (p. 389) are small proteins that transfer electrons from reduced ferredoxin to certain enzymes (**Figure 13.24**). When exposed to light, PSI reduces ferredoxin, an electron donor in PSI. Reduced ferredoxin then reduces ferredoxin–thioredoxin reductase (FTR), an iron-sulfur protein that mediates the transfer of electrons between ferredoxin and thioredoxin. Reduced thioredoxins alter the activities of several enzymes. For example, the Calvin cycle enzymes fructose-1,6-bisphosphatase, sedoheptulose-1,7-bisphosphatase, $NADP^+$-glyceraldehyde-3-phosphate dehydrogenase, and ribulose-5-phosphate kinase are activated, and glucose-6-phosphate dehydrogenase, the pentose phosphate pathway enzyme, is inhibited. By depressing the activity of the pentose phosphate pathway, light-driven reduced thioredoxin production prevents carbon flow through an opposing pathway.

4. **Phytochrome**. Phytochrome is a 120 kDa protein that possesses a red light–sensitive chromophore that exists in two forms: P_r and P_{fr}. The inactive blue form, P_r, absorbs red light (670 nm). The absorption of longer wavelengths (i.e., far red light, 720 nm) converts P_r to P_{fr}, the active green form. In the dark, P_{fr} decays back to P_r. Phytochrome activation triggers several signal transduction pathways that mediate hundreds of plant responses to light. In addition to phytochrome's effects on plant processes such as seed germination, it also has specific effects on photosynthetic processes. These include controlling the rate of synthesis of the small subunit of rubisco, regulation of light absorption by LHCIIb, a component of LHCII, and positioning chloroplasts within photosynthesizing cells.

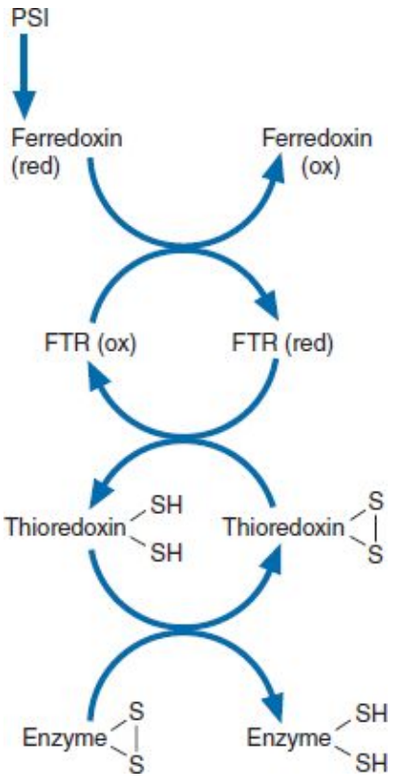


FIGURE 13.24

The Ferredoxin–Thioredoxin System

Using the light energy captured by PSI, energized electrons are donated to ferredoxin. Electrons donated by ferredoxin to ferredoxin–thioredoxin reductase (FTR) are used to reduce the disulfide bridge of thioredoxin. Thioredoxin then reduces the disulfide bridges of susceptible enzymes. Some enzymes are activated by this process, whereas others are inactivated.

Control of Ribulose-1,5-Bisphosphate Carboxylase

Rubisco is regulated by light-stimulated changes in gene expression, ion concentrations, and covalent modification. Rubisco genes, found within the chloroplast (the L subunit) and the nucleus (the S subunit), are transcribed as a result of increased light intensity. Once the newly synthesized S subunit has been transported from the cytoplasm into the chloroplast, both subunits assemble to form the L_8S_8 holoenzyme. When illumination is low, the synthesis of both subunits is rapidly depressed.

The activity of rubisco is also modified by several metabolic signals. When photosynthesis is active, the pH in the stroma increases (protons are being pumped out of the stroma into the thylakoid lumen) and the Mg^{2+} concentration increases (Mg^{2+} moves into the stroma as H^+ moves out). Both changes increase the activity of rubisco. An important consideration in this process is whether the stomata are open or closed (see the discussion of photorespiration on p. 514). Although CO_2 is the preferred substrate for rubisco, under physiological conditions both the carboxylase activity and the oxidase activity compete significantly with each other. If the stomata are closed, as they would be on a hot, dry day, O_2 accumulation in the leaf tissue greatly decreases the proportional participation of the carboxylase activity of rubisco. Recall that C4 plants diminish this competition by trapping the CO_2 in a four-carbon intermediate and delivering the CO_2 via decarboxylation directly to a rubisco molecule that is protected from exposure to O_2.

Rubisco is subject to covalent modification. The active site of the L subunit must be carbamoylated at a specific lysine residue to be active. *Carbamoylation* is the nonenzymatic carboxylation of a free primary amino group, in this case the ε-amino group of a certain lysine in the active site of rubisco (**Figure 13.25**). The rate of carbamoylation is dependent on the CO_2 concentration and an alkaline

pH, which ensures that CO_2 fixation occurs at an appreciable rate only when the CO_2 levels and available energy are high. Ribulose-1,5-bisphosphate can and does bind to the active site in both its modified and unmodified forms, but catalysis can occur only when rubisco is carbamoylated. The level of activation is cooperative and increases as more of the eight subunits are modified.

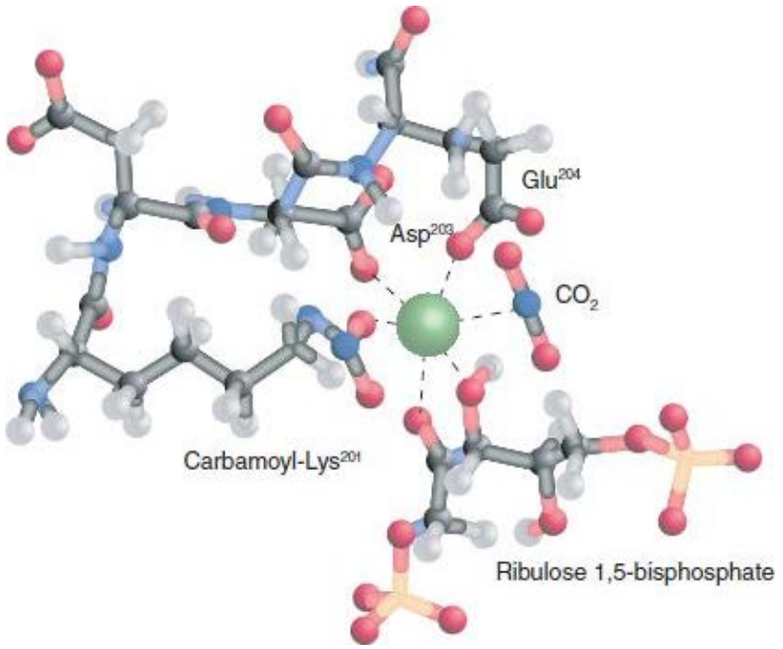


FIGURE 13.25

The Carbamoylated Active Site of Rubisco

Within the active site of rubisco, a magnesium ion orients the substrates ribulose-1,5-bisphosphate and CO_2 with oxygen atoms in aspartate and glutamate side chains and the carbamoyl group of a lysine residue.

An enzyme called rubisco activase (about 5% of soluble protein in leaves) mediates an ATP-dependent remodeling of rubisco that allows the release of ribulose-1,5-bisphosphate from the active site so that carbamoylation can occur, followed by enzyme activation. In the absence of light, photosynthesis is depressed and the ATP required for this activation process is greatly reduced, as is the NADPH required for the Calvin cycle. In some plants the absence of light results in the binding of a competitive inhibitor molecule 2-carboxyarabinitol-1-phosphate (CA1P) (**Figure 13.26**) to the active site of rubisco. In the light, rubisco activase facilitates the release of CA1P from rubisco.

Photosynthesis is also highly sensitive to heat stress. The reasons for this constraint on plant growth include depressed ATP synthesis (impaired electron transport) and reduced rubisco activase activity. Recent evidence indicates that rubisco activase is inactivated at temperatures higher than 32°C, whereas rubisco is stable up to 55°C.

KEY CONCEPTS

Light is the principal regulator of photosynthesis. In photosynthetic processes, light affects the activities of regulatory enzymes such as rubisco by means of indirect mechanisms, which include changes in pH, Mg^{2+} concentration, the ferredoxin–thioredoxin system, and phytochrome. Rubisco is also regulated by covalent modification in that carbamoylation of an active site lysine residue is required for optimal activity.

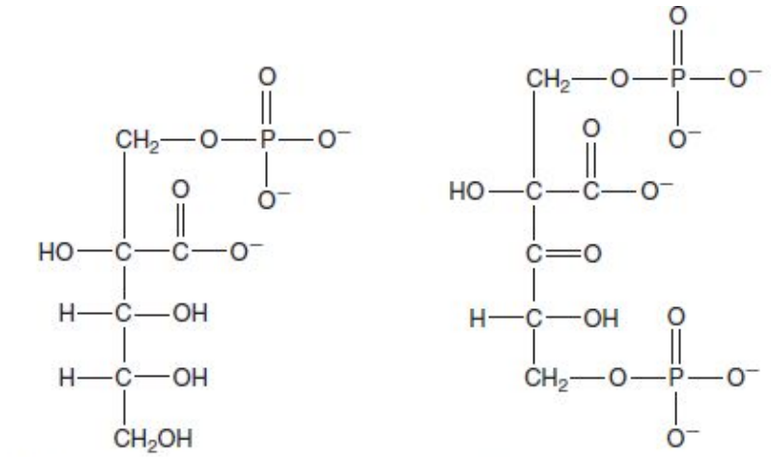


FIGURE 13.26

2-Carboxyarabinitol-1-Phosphate

(a) 2-Carboxyarabinitol-1-phosphate (CA1P) is a competitive inhibitor of rubisco. It is a substrate analogue that has a similar structure to the (b) transition state intermediate of the enzyme's substrate ribulose-1,5-bisphosphate.

Biochemistry IN PERSPECTIVE

The Artificial Leaf: Biomimetic Photosynthesis

Can humans successfully produce solar fuels as a substitute for environmentally damaging fossil fuels? As human use of dwindling fossil fuels increases and Earth's atmospheric concentration of the greenhouse gas CO_2 moves ever higher, scientists and engineers search for alternative sources of energy. Solar energy, radiant energy derived from nuclear fusion reactions within the sun, is an attractive possibility since the annual total power requirement of humans is currently about 16 TW, just a small fraction of the estimated 120,000 TW of solar energy that reaches Earth's land masses every year. Solar power, the use of manufactured photovoltaic cells (composed of monocrystalline silicon wafers) to convert sunlight into electrical energy, is an effort to reduce reliance on fossil fuels. Although solar power is a clean and environmentally safe energy source, current solar technology is expensive. For example, many solar cells are manufactured with rare transition metals such as platinum and iridium to increase energy conversion efficiency. In addition, solar cells must be used in large arrays, called solar panels, to be economically viable, and they require constant maintenance.

Artificial photosynthesis research is an attempt to imitate the essential feature of the natural process, photoinduced electron transfer creating a charge-separated state, which harnesses solar power in the form of a synthesized fuel. The principal components of any successful artificial photosynthetic system are the same as those observed in natural photosynthesis: (1) a chromophore-containing light-harvesting mechanism associated with a reaction center complex that transduces photon energy into electrochemical energy (i.e., a charge-separated state is created when an electron is released from the reaction center complex, thereby creating an electron hole); (2) a water-oxidation catalyst that converts a water molecule to four electrons (which, one at a time, fill the electron hole in

the reaction center complex), four hydrogen ions (H^+), and O_2; and (3) a catalyst that uses the water-derived electrons (reducing equivalents) and hydrogen ions to make fuel molecules, most notably H_2 ($2H^+ + 2e^- \rightarrow H_2$) or methanol (via reduction of CO_2). A model of artificial photosynthesis based on a light-harvesting supermolecular complex is outlined in **Figure 13A**.

Progress has been made in the development and assembly of a variety of artificial photosynthetic components, but the goal of producing a robust, photochemically stable and environmentally friendly means of solar-driven fuel synthesis has not yet been achieved. Of the numerous problems that must be solved, researchers currently consider the water-splitting mechanism the most challenging. The reaction $2\ H_2O \rightarrow 4e^- + 4H^+ + O_2$ is endothermic and, therefore, requires a catalyst to achieve or surpass the electron transfer efficiency of the natural water-oxidizing complex (the manganese cluster and the tyrosine amino acid side chain). Among the best candidates for this role are manganese-containing complexes, dye-sensitized titanium dioxide (TiO_2), and cobalt oxide (CoO). However, these and other water-splitting catalysts so far investigated have electron transfer properties that are too slow. In addition, they lack coordination in the four electron-releasing process of water oxidation with the one-electron charge separation mechanism in the reaction center complex. As a result, water splitting remains a significant barrier to reaching the goal of competitively priced solar-driven energy production and reduction in the use of global warming-linked fossil fuels.

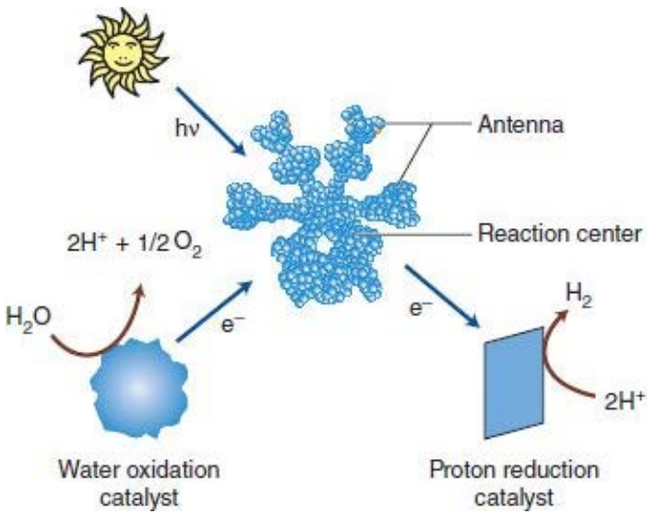


FIGURE 13A

An Artificial Photosynthesis Model

As with natural photosynthesis, all artificial photosynthetic devices must include a photoinduced electron transfer mechanism that creates a charge-separated state. Light energy is harvested by the chromophore components of a supermolecular complex such as the one illustrated in the center of the figure. This complex, referred to as a molecular hexad, is composed of three types of chromophores [bis(phenylethynyl) anthracene, borondipyrromethene, and the centrally located zinc tetraarylporphyrin] that together absorb energy across the visible region of the electromagnetic spectrum. After energy transfer from the other antennas, the zinc porphyrin donates an electron to the attached spherical fullerene to yield a charge-separated state. Despite the success of such molecular complexes in creating a charge-separated state, artificial photosynthesis is not yet feasible. In addition to the current lack of effective water-oxidizing and proton reduction catalysts, current research has yet to provide a mechanism that can coordinate the four-electron water-oxidizing process (yielding O_2), the one-electron charge-separating reaction center process, and the two-electron H_2-generating reaction.

SUMMARY Artificial photosynthesis, the use of solar energy to drive the synthesis of storable fuel such as H_2, is a goal that scientists and engineers are currently

working toward. Although progress in the creation of light-driven charge separation has been made in the laboratory, the goal of cost-effective solar-driven fuel synthesis remains elusive. The seemingly intractable problems encountered in artificial photosynthesis research were solved by living organisms several billions of years ago.

Chapter Summary

1. In plants, photosynthesis takes place in chloroplasts. Chloroplasts possess three membranes. The outer membrane is highly permeable, whereas the inner membrane possesses a variety of carrier molecules that regulate molecular traffic into and out of the chloroplast. A third membrane, called the thylakoid membrane, forms an intricate series of flattened vesicles called grana connected by stromal lamellae that house the photosynthetic machines.
2. Photosynthesis consists of two major phases: the light reactions and the light-independent reactions. During the light reactions, water is oxidized, O_2 is evolved, and the ATP and NADPH required to drive carbon fixation are produced. The major working units of the light reactions are photosystems I and II, the cytochrome b_6f complex, and the ATP synthase. In noncyclic electron transport, electrons from water molecules are transferred from photosystem II to photosystem I to $NADP^+$ with the production of O_2, ATP, and NADPH. Cyclic electron transport involves only PSI and generates additional ATP but no NADPH. During the light-independent reactions, CO_2 is incorporated into organic molecules. The first stable product of carbon fixation is glycerate-3-phosphate. The Calvin cycle is composed of three phases: carbon fixation, reduction, and regeneration.
3. Most of the carbon incorporated during the Calvin cycle is used initially to synthesize starch and sucrose, both of which are important energy sources. Sucrose is also important because it is used to translocate fixed carbon throughout the plant.
4. Photorespiration is an apparently wasteful process whereby O_2 is consumed and CO_2 is released from plants. Its role in plant metabolism is not understood. C4 plants, which must tolerate hot, dry environments, have developed biochemical and anatomical mechanisms for suppressing photorespiration.
5. Light is an important regulator of most aspects of photosynthesis. Many of the effects of light are mediated by changes in the activities of key enzymes. The mechanisms by which light effects these changes include changes in pH, Mg^{2+} concentration, the ferredoxin–thioredoxin system, and phytochrome. The most important enzyme in photosynthesis is ribulose-l,5-bisphosphate carboxylase. Its activity is highly regulated. Light activates the synthesis of both of the enzyme's subunits. Allosteric effectors as well as covalent modification affect rubisco's activity. Carbamoylation of the enzyme's active site lysine residue is required for activation.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on water to help you prepare for exams.

Take your learning further by visiting the

Suggested Readings

Alric J. 2015. The plastoquinone pool, poised for cyclic electron flow? Frontiers Plant Sci 6:540–3.

Cardona T, et al. 2012. Charge separation in photosystem II: a comparative and evolutionary overview. Biochim Biophys Acta Bioenergetics 1817(1):26–43.

Lane N. 2009. Photosynthesis. In Life ascending: the great inventions of evolution. London (UK): W. W. Norton.

Mirkovic T, et al. 2017. Light absorption and energy transfer in the antenna complexes of photosynthetic organisms. Chem Rev 117(2):249–93.

Nelson N, Ben-Shem A. 2004. The complex architecture of oxygenic photosynthesis. Nat Rev Mol Cell Biol 5(12):971–82.

Romero E, et al. 2017. Quantum-coherent dynamics in photosynthetic charge separation revealed by wavelet analysis. Sci Reports 7:2890–7.

Van Eerden FJ, et al. 2017. Exchange pathways of plastoquinone and plastoquinol in the photosystem II complex. Nat Communications 8:15214–21.

Key Words

antenna pigment, *497*
carbamoylation, *518*
C3 plant, *510*
C4 metabolism, *515*
C4 plant, *515*
Calvin cycle, *510*
carbon fixation, *510*
carotenoid, *494*
chlorophyll, *494*
chromophore, *501*
fluorescence, *502*
granum, grana, *495*
light-harvesting, *494*
light-harvesting antenna, *493*
light-independent reaction, *510*
light reaction, *503*
photophosphorylation, *509*
photorespiration, *514*
photosystem, *493*
reaction center, *493*
resonance energy transfer, *502*
stroma, *495*
stromal lamella, *495*
thylakoid membrane, *495*
thylakoid lumen, *495*
Z scheme, *505*

Review Questions

SECTION 13.1

Comprehension Questions

1. Define the following terms:
    a. photosynthesis
    b. photosystem
    c. light harvesting antenna
    d. reaction center
    e. chlorophyll
2. Define the following terms:
    a. carotenoid
    b. stroma
    c. β-carotene
    d. lutein
    e. thylakoid membrane
3. Define the following terms:
    a. granum
    b. stromal lamellae
    c. thylakoid lumen
    d. PSI
    e. PSII
4. Define the following terms:
    a. OEC
    b. PQ
    c. Y_Z
    d. pheophytin a
    e. PC
5. Define the following terms:
    a. Rieske protein
    b. light-harvesting complex
    c. Psa dimer
    d. D_1/D_2 dimer
    e. A_0
6. Define the following terms:
    a. cytochrome b_6f complex
    b. CF_o
    c. CF_1
    d. LHCII
    e. Mn_4CaO_5

Fill in the Blanks

7. The photosynthetic mechanisms referred to as _____ are membrane-bound protein complexes found in chloroplasts.
8. Unlike mitochondria, chloroplasts contain a distinct third membrane called the _____ membrane.

9. Chloroplasts are descendants of primitive _____.
10. The water-splitting site in PSII is a cube-like _____ cluster.
11. The oxygen-evolving complex is also known as the _____.
12. The function of the _____ complex in photosynthesis is similar to the Q cycle in mitochondria.

Short-Answer Questions

13. List five ways in which chloroplasts resemble mitochondria.
14. List four consequences of global warming.
15. List three criteria for sustainable biofuel production.
16. What properties make O_2 superior to sulfur, sulfate, and nitrate as a metabolic waste product?
17. What roles do the chloroplast inner membrane have?
18. List the three primary photosynthetic pigments and describe the role each plays in photosynthesis.
19. What is the function of the phytol chain in chlorophyll molecules?
20. What is a photosynthetic special pair, and how does it function?
21. The chloroplast has a highly organized structure. How does this structure make photosynthesis possible?

Critical-Thinking Questions

22. The burning of fossil fuels releases CO_2 into the atmosphere and is detrimental to Earth's ecosystems. Explain why the use of biofuels, which also release CO_2, is an improvement over fossil fuels.
23. H_2S oxidation was one of the earliest mechanisms of photosynthesis on the ancient Earth. This mechanism requires less energy input than H_2O oxidation. Suggest a reason why this major shift in electron source occurred.
24. It has been suggested that chloroplasts, like mitochondria, evolved from independent living organisms. What features of the chloroplast suggest that this is true?
25. How would biofuels generated by photosynthetic algae be an improvement over current methods?

SECTION 13.2

Comprehension Questions

26. Define the following terms:

    a. wavelength
    b. chromophore
    c. Planck's constant
    d. fluorescence
    e. resonance energy transfer

27. Define the following terms:

    a. radiationless decay
    b. P700
    c. P680
    d. $c = \lambda v$
    e. UV

28. Define the following terms:
    a. LCHI
    b. blue light
    c. red light
    d. wave frequency
    e. light-harvesting

Fill in the Blanks

29. Molecules that absorb electromagnetic energy have structural components called _____.
30. _____ is a process in which an excited molecule decays to its ground state by converting excess energy to heat.
31. In _____ the excited molecule decays to its ground state by emitting a photon.
32. _____ is the term used to describe the height of a wave.
33. The yellow of the leaves in autumn is caused by the presence of _____ molecules.

Short-Answer Questions

34. Using the action spectrum for photosynthesis, determine what wavelengths of light appear to be optimal for photosynthesis. [*Hint*: Refer to the Biochemistry in the Lab essay Photosynthetic Studies on the companion website, www.oup.com/us/mckee.]
35. What types of responses to energy absorption are important in photosynthesis?
36. What are P680 and P700 and what photosystems contain them?
37. What is the general function of each photosystem used by plants?
38. Excited molecules can return to the ground state by several means. Describe each briefly. Which of these processes are important in photosynthesis? Describe how they function in a living organism.

Critical-Thinking Questions

39. The statement has been made that the more extensively conjugated a chromophore, the less energy a photon needs to excite it. What is conjugation in organic molecules, and how does it contribute to this phenomenon?
40. Suggest a reason why photosynthetic pigments readily absorb in the blue region of the visible spectrum but do so with low probability in the ultraviolet.

SECTION 13.3

Comprehension Questions

41. Define the following terms:
    a. light-dependent reactions
    b. Z-scheme
    c. water-oxidizing clock
    d. noncyclic electron transport
    e. cyclic electron transport
42. Define the following terms:
    a. F_A

b. QBH_2
c. plastoquinol
d. OEC
e. photophosphorylation

Fill in the Blanks

43. In photosynthesis, _____ pigments absorb light energy and transfer it to the reaction center.
44. The products of the light reactions of photosynthesis are _____ and NADPH.
45. During photosynthesis, ATP is synthesized as protons flow across the thylakoid membrane into the _____.
46. When P700 absorbs a photon, it releases an energized electron, which is immediately replaced by an electron provided by _____.

Short-Answer Questions

47. What is the final electron acceptor in photosynthesis when the $NADPH/NADP^+$ ratio is low? Does your answer change if the $NADPH/NADP^+$ ratio is high?
48. Why is the oxygen-evolving system referred to as a clock?
49. What is a special pair, and how does it function?
50. Describe the Z scheme of photosynthesis. How are the products of this reaction used to fix carbon dioxide?
51. What effect would you expect dinitrophenol to have on photosynthesis?

Critical-Thinking Questions

52. Both oxidative phosphorylation and photophosphorylation trap energy in high-energy bonds. How are these processes different? How are they the same?
53. Although both organelles originated as free-living prokaryotes, mitochondria are significantly smaller than chloroplasts. Suggest a reason for this discrepancy. [*Hint*: Consider the energy sources for both organelles.]
54. Photosynthesizing organisms of the deep ocean capture long-wave radiation. Determine how many quanta of light at 1000 nm are required to provide as much energy as plants that absorb at 700 nm. [*Hint*: Recall that $E = hc/\lambda$, where h is Planck's constant (1.58×10^{-37} kcal/s), c is the speed of light (3×10^8 m/s), and λ is wavelength.]
55. Increasing the intensity of the incident light but not its energy increases the rate of photosynthesis. Explain.

SECTIONS 13.4 and 13.5

Comprehension Questions

56. Define the following terms:

a. Calvin cycle
b. carbon fixation
c. C3 metabolism
d. C4 metabolism

e. rubisco

57. Define the following terms:

a. photorespiration
b. CA1P
c. reductive pentose phosphate cycle
d. phytochrome
e. stomata

Fill in the Blanks

58. Both carbon dioxide and _____ compete for the active site of rubisco.
59. The Hatch–Slack pathway is an alternative name for _____.
60. For every three molecules of CO_2 that are incorporated into carbohydrate during the Calvin cycle there is a net gain of one molecule of _____.
61. In C4 plants, CO_2 is first converted to the _____ ion, which is then incorporated into _____.
62. Sucrose-6-phosphate is synthesized from _____ and fructose-6-phosphate by sucrose phosphate synthase.
63. In C4 metabolism _____ carries the newly incorporated carbon and oxygen atoms from mesophyll cells to bundle sheath cells.

Short-Answer Questions

64. Plants actively engaged in photosynthesis also evolve carbon dioxide. Explain.
65. The products of photosynthesis are carbohydrate molecules and O_2. From which of the substrate molecules (CO_2 or H_2O) are the oxygen atoms derived?
66. Where does carbon dioxide fixation take place in a chloroplast with reference to the light-independent reactions?
67. Why does CO_2 depress photorespiration?
68. Describe the function of thioredoxin in photosynthesis regulation.

Critical-Thinking Questions

69. Generally, increasing the concentration of carbon dioxide increases the rate of photosynthesis. What conditions could prevent this effect?
70. Which plants expend more energy per molecule of glucose produced, C3 or C4?
71. Triazine herbicides are effective C3 plant toxins that are used to suppress weed growth. They apparently bind to a plastocyanin-binding protein. Suggest a mechanism whereby triazines undermine C3 growth.
72. Explain the following observation. When a photosynthetic system is exposed to a brief flash of light, no oxygen is evolved. Only after several bursts of light is oxygen evolved.

MCAT Study Questions

73. Which of the following is not a product of the light reactions in photosynthesis?

a. O_2
b. NADPH
c. H_2O

d. ATP

74. In photosynthesis, NADPH is the product of

a. Photosystem II
b. Photosystem I
c. the manganese cluster enzyme complex
d. P680

75. The initial carbohydrate product of the Calvin cycle is

a. glycerate-3-phosphate
b. ribulose-1,5-bisphosphate
c. ribulose-5-phosphate
d. fructose-6-phosphate

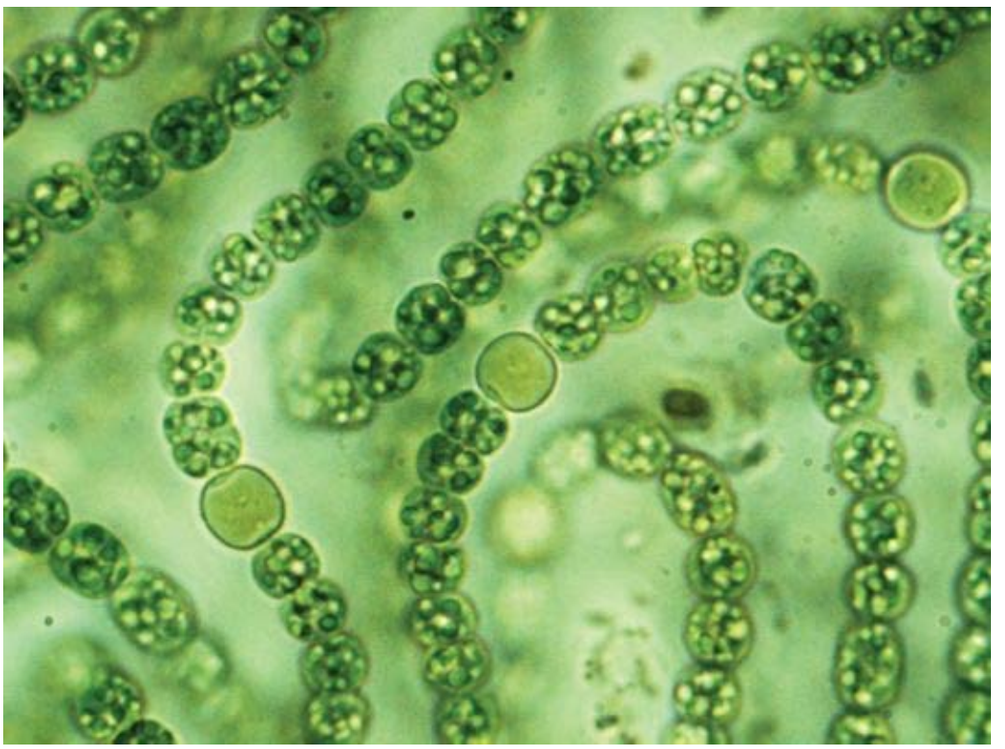
76. Which of the following wavelengths in the electromagnetic spectrum has the most energy?

a. microwaves
b. visible light
c. ultraviolet
d. radio waves

77. During photosynthesis, ATP is formed and released initially into the

a. thylakoid lumen
b. thylakoid lamellae
c. grana
d. stroma

CHAPTER 14 Nitrogen Metabolism I: Synthesis



Nitrogen Fixation by Cyanobacteria Nitrogen-fixing organisms such as cyanobacteria convert atmospheric nitrogen (N_2) to ammonia (NH_3), a biologically useful form. Nitrogen fixation takes place in heterocysts: large, specialized cells with several thick cell walls that exclude O_2, the waste product released by the more numerous photosynthesizing vegetative cells.

OUTLINE

NITROGEN AND THE GULF OF MEXICO DEAD ZONE

14.1 NITROGEN FIXATION
The Nitrogen Fixation Reaction
Nitrogen Assimilation

14.2 AMINO ACID BIOSYNTHESIS
Amino Acid Metabolism Overview
Reactions of Amino Groups
Biosynthesis of the Amino Acids

14.3 BIOSYNTHETIC REACTIONS INVOLVING AMINO ACIDS

One-Carbon Metabolism
Glutathione
Neurotransmitters
Nucleotides
Heme

Biochemistry in Perspective

Gasotransmitters

AVAILABLE ONLINE

Biochemistry in Perspective
The Amine Neurotransmitters
Biochemistry in Perspective
The Catecholamines
Biochemistry in Perspective
Parkinson's Disease and Dopamine
Biochemistry in Perspective
Heme and Chlorophyll Biosynthesis
Biochemistry in Perspective
The Essential Amino Acids
Biochemistry in Perspective
The Nucleotides: IMP Biosynthesis
Biochemistry in Perspective
Lead Poisoning

Nitrogen and the Gulf of Mexico Dead Zone

Every summer for at least three decades, there has been an expanding phytoplankton overgrowth in the coastal water of the Gulf of Mexico near Louisiana and Texas. This overgrowth, referred to as a phytoplankton "bloom," triggers the formation of a zone in which no aerobic organisms can exist. In 2017, this *dead zone* (**Figure 14.1**) covered almost 23,000 km^2 (about 8800 mi^2), an area one and one-half times the size of Massachusetts. The dead zone not only causes habitat destruction, but also poses an obvious threat to both human health and a multibillion-dollar commercial and recreational fishing industry.

Phytoplankton (e.g., cyanobacteria and diatoms) are microorganisms that produce more than half of Earth's oxygen. Consumed by zooplankton (microscopic animals) and fish, phytoplankton serve as the foundation of aquatic food webs. Their growth requires light, CO_2 and nutrient availability. Ordinarily, phytoplankton growth is limited only by temperature and low levels of nutrients, especially those of nitrogen, and, to a lesser extent, phosphorus. Small-scale, transient blooms, which occur in ocean water when air and surface water temperatures are high, are typically linked to upwelling. In this process, caused by high winds, cold nutrient-dense ocean water moves from the bottom to the surface.

The Gulf of Mexico dead zone, one of several hundred now observed on Earth, has been definitively linked to excess nitrogen discharged into the Gulf by the Mississippi River (e.g., 165,000 metric tons of nitrate fertilizer per year). In addition to agricultural fertilizers (ammonia and nitrate), the major causes of the massive Gulf phytoplankton bloom include agricultural animal waste runoff, untreated or undertreated sewage release from water treatment plants, and septic tank runoff into the rivers and streams in the Mississippi River watershed that drain water from about 40% of the continental United States. The concentration of phytoplankton is so high (thousands of cells per milliliter) in the Gulf phytoplankton bloom

that, depending on the dominant species, the water changes color to green, brown, or red.

Phytoplankton blooms trigger dead zone formation because once a bloom ends, dead phytoplankton and the zooplankton that fed on them drift to the bottom, where decomposition is carried out by aerobic bacteria. So much O_2 is used in this decaying process that the water becomes *hypoxic* (dissolved oxygen <2 mg/L). (In a healthy ocean environment, the concentration of dissolved oxygen is about 5 mg/L.) Soon, fish and other oxygen-requiring organisms (e.g., shrimp, clams, and oysters) suffocate. Once the bottom sediments become *anoxic* (no oxygen), foul-smelling substances [e.g., hydrogen sulfide (H_2S)] released by anaerobic bacteria begin accumulating.

The Gulf of Mexico dead zone recovery will be difficult and expensive because of the size of the Mississippi River watershed. In addition to substantial improvements in the efficient and targeted use of fertilizers, recovery will require significantly upgraded water treatment facilities throughout the Mississippi River watershed. It will also depend on the restoration of wetlands, ecosystems that can reduce nutrient pollution before runoff water reaches the river.

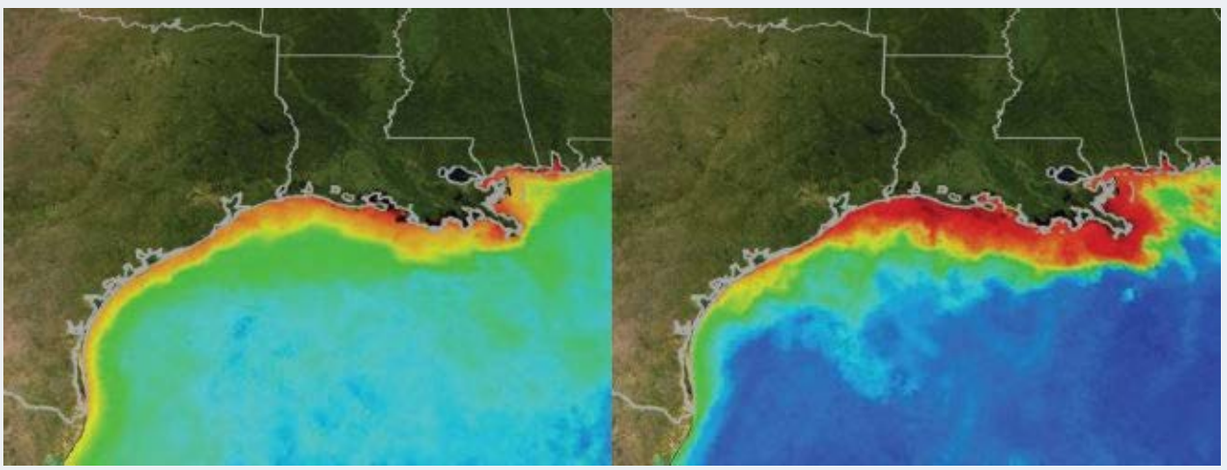


FIGURE 14.1

Satellite Images of the Northern Gulf of Mexico in Winter and Summer

In these images, red and orange represent high concentrations of phytoplankton and sediment that originated in the Mississippi River. Note that the size of the dead zone off the coasts of Louisiana and Texas is larger and extends farther from land in summer (*right*) than in winter (*left*).

Overview

NITROGEN IS FOUND IN AN ASTONISHINGLY VAST ARRAY OF BIOMOLECULES. THESE INCLUDE THE AMINO ACIDS AND THE NITROGENOUS BASES THAT are used in the synthesis of proteins and the nucleic acids, respectively. Other essential nitrogen-containing biomolecules include the porphyrins (e.g., heme and chlorophyll), certain membrane lipids, and a diverse group of metabolically important biomolecules that are synthesized in smaller amounts (e.g., several neurotransmitters and glutathione). This chapter traces nitrogen from nitrogen fixation, the process that converts inert N_2 to biologically useful ammonia (NH_3), through the synthesis of the major nitrogen-containing biomolecules.

he nitrogen cycle is the biogeochemical cycle in which nitrogen atoms flow through the

The biosphere. Several biochemical processes convert nitrogen from one form to another. Nitrogen fixation, the incorporation of nitrogen into organic molecules, begins with the fixation (reduction) of N_2 by prokaryotic microorganisms to form ammonia (NH_3). Plants such as corn depend on absorbing NH_3 and NO_3^- (nitrate), the oxidation product of NH_3, either synthesized by soil bacteria or provided by artificial fertilizers. Nitrogen supply is often the limiting factor in plant growth and development because the amount of fixed nitrogen available to plants is usually small.

Whether plants acquire NH_3 by nitrogen fixation, by absorption from the soil, or by reduction of absorbed NO_3^-, it is assimilated by conversion into the amide group of glutamine. This "organic nitrogen" is then transferred to other carbon-containing molecules to produce the amino acids used by the plant to synthesize nitrogenous molecules (e.g., proteins, nucleotides, and heme). Organic nitrogen, primarily in the form of amino acids, then flows throughout the ecosystem as plants are consumed by animals and decomposing microorganisms. When organisms die, organic nitrogen is *mineralized*, meaning that it is converted through the actions of numerous types of microbes into NH_3, NO_3^-, NO_2^- (nitrite), and eventually N_2.

This chapter begins with a discussion of nitrogen fixation and a description of the essential features of amino acid biosynthesis, followed by descriptions of the biosynthesis of select nitrogen-containing molecules. A special emphasis is placed on the anabolic pathways of the nucleotides. Chapter 15 traces the flow of nitrogen atoms through several catabolic pathways to the nitrogenous waste products excreted by animals.

14.1 NITROGEN FIXATION

Several circumstances limit the amount of usable nitrogen available in the biosphere. The most notable are the limited number of species that can convert the chemically stable molecule N_2 (bond energy = 940 kJ/mol) into NH_3, a more chemically reactive molecule, and the high-energy requirements of this process, referred to as **nitrogen fixation**. Among the most prominent nitrogen-fixing species are free-living bacteria (e.g., *Azotobacter vinelandii* and *Clostridium pasteurianum*), the cyanobacteria (e.g., *Nostoc muscorum* and *Anabaena azollae*), and symbiotic bacteria (e.g., several species of *Rhizobium*). Symbiotic organisms form *mutualistic*—that is, mutually beneficial—relationships with host plants or animals. *Rhizobium* species, for example, infect the roots of leguminous plants such as soybeans and alfalfa.

Nitrogen fixation requires a large energy input because reduction of N_2 to form NH_3 involves the breaking of the nonpolar triple bond of atmospheric dinitrogen gas. In commercial nitrogen fixation, NH_3 is the product of the Haber–Bosch reaction (p. P-17), where H_2 and N_2 are heated at 400–650°C under 200–400 atm pressure in the presence of an iron catalyst. Unlike the Haber–Bosch process, nitrogen-fixing species convert N_2 to NH_3 at ambient temperature and atmospheric pressure. The energy requirements for the biological process, however, are also high, with a minimum of 16 ATP required to reduce one N_2 to form two ammonia molecules. The overall reaction of nitrogen fixation is as follows:

$$N_2 + 8e^- + 16\,ATP + 10H^+ \rightarrow 2NH_4^+ + 16\,ADP + 16\,P_i + H_2$$

The Nitrogen Fixation Reaction

All species that can fix nitrogen possess the *nitrogenase complex*. Its structure, similar in all species so far investigated, consists of two proteins called dinitrogenase and dinitrogenase reductase (**Figure 14.2**). Dinitrogenase (240 kDa), also referred to as MoFe *protein*, is an $\alpha_2\beta_2$ heterotetramer. Each $\alpha\beta$

dimer is a catalytic unit that contains two unique metal prosthetic groups: a *P cluster* [8Fe-7S] and a *molybdenum–iron cofactor* (MoFe cofactor or M cluster), which contains a carbide atom (a carbon atom at the center of a metal cluster). The FeMo cofactor is linked to the tricarboxylic acid homocitrate (7Fe-9S-Mo-C-homocitrate). The MoFe protein catalyzes the reaction $N_2 + 8H^+ + 8e^- \rightarrow 2NH_3 + H_2$. (Note that NH_3 is the initial product. Under cellular pH conditions it will be in equilibrium with NH_4^+.) Dinitrogenase reductase (60 kDa) (also referred to as *Fe protein*) is a dimer containing identical subunits, each of which has a Mg^{2+}ATP-binding site. A 4Fe–4S cluster is bound at the interface of the two subunits, 15 Å from the Mg^{2+}ATP-binding site and close to the docking site for the dinitrogenase tetramer (MoFe protein). The Fe protein transfers electrons, ultimately derived from NAD(P)H, one at a time to the Mo–Fe protein. Both proteins in the nitrogenase complex are irreversibly inactivated by O_2.

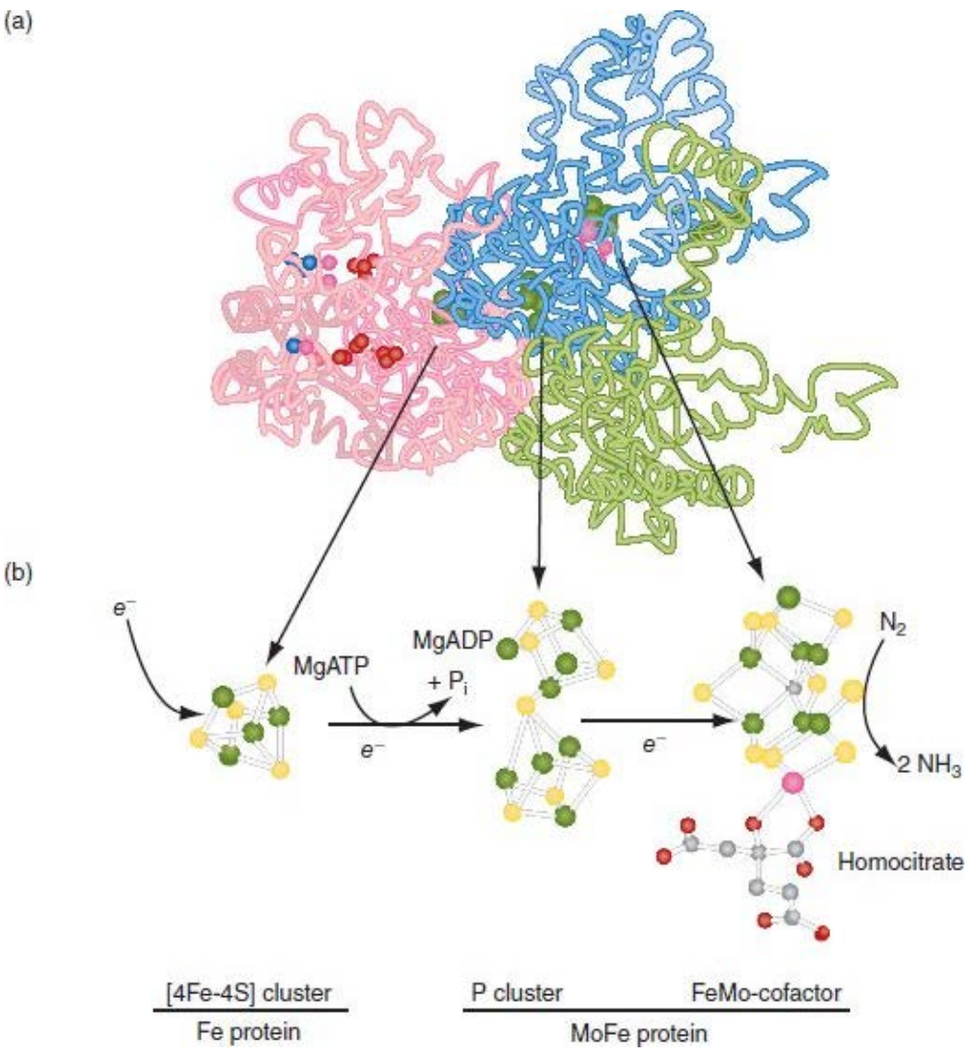


FIGURE 14.2

Nitrogenase Complex Structure

(a) [*left*] Fe protein dimer (pink and red); [*right*] an $\alpha\beta$ dimer of the MoFe protein, with the α subunit in blue and the β subunit in green. (b) The metal clusters of the Fe protein [4Fe–4S] and the P cluster and FeMo cofactor of the FeMo protein are illustrated as ball-and-stick models. Homocitrate, a component of the MoFe cofactor, is a derivative of citrate that contains an additional methylene group. Atom colors: carbon, gray; oxygen, red; phosphorus, dark green; sulfur, yellow; iron, green; molybdenum, pink.

3D animation of nitrogenase

The first step in nitrogen fixation (**Figure 14.3**) is the transfer of electrons from NAD(P)H to ferredoxin, a powerful reducing agent that in turn donates the electrons to the Fe protein FeS cluster, one at a time. Each electron transfer begins with the reduction of the oxidized [4Fe–4S] cluster in the Fe protein with two bound ADP [$Fe^{ox}(ADP)_2$] to yield [$Fe^{red}(ADP)_2$]. This reduction triggers the replacement of ADP with ATP in both subunits. The product [$Fe^{red}(ATP)_2$] then docks with the MoFe protein to form the active complex [$Fe^{red}(ATP)_2$:MoFe]. The docking event is followed by the transfer of an electron from the P cluster of the MoFe protein to the M cluster. An electron from the Fe protein is then transferred to the oxidized P cluster to yield [$Fe^{ox}(ATP)_2$:$MoFe^{red}$]. ATP hydrolysis and P_i release cause a conformational change that allows the dissociation of [$Fe^{ox}(ADP)_2$] from the reduced MoFe protein. The electron transfer process is then repeated until eight electrons have been delivered to the MoFe cofactor. (In its oxidized electron-deficient form, the Fe protein has a higher affinity for Mg^{2+}ATP.) The reduction potentials for the Fe protein and the P cluster have been recorded in the –400-mV and –300-mV ranges, respectively. The transfer of the first two electrons leads to the reduction of H^+ (the $MoFeH_2$ state of the enzyme). This stage of the process occurs with or without the presence of N_2. The incoming N_2 exchanges with H_2 in the active site (between the molybdenum ion and four coordinating Fe centers) to form a stable intermediate ($MoFeN_2$). Subsequently, six electrons and six protons are transferred to the active site to form the diimine (HN=NH, two electrons added), then the hydrazine ($H_2N—NH_2$, four electrons added), and, finally, two NH_3 products (six electrons added). Two electrons are then used to reduce $2H^+$ to H_2 (essential to the catalytic process). In this reaction sequence, six electrons are used to reduce N_2 in three steps to yield two NH_3 molecules.

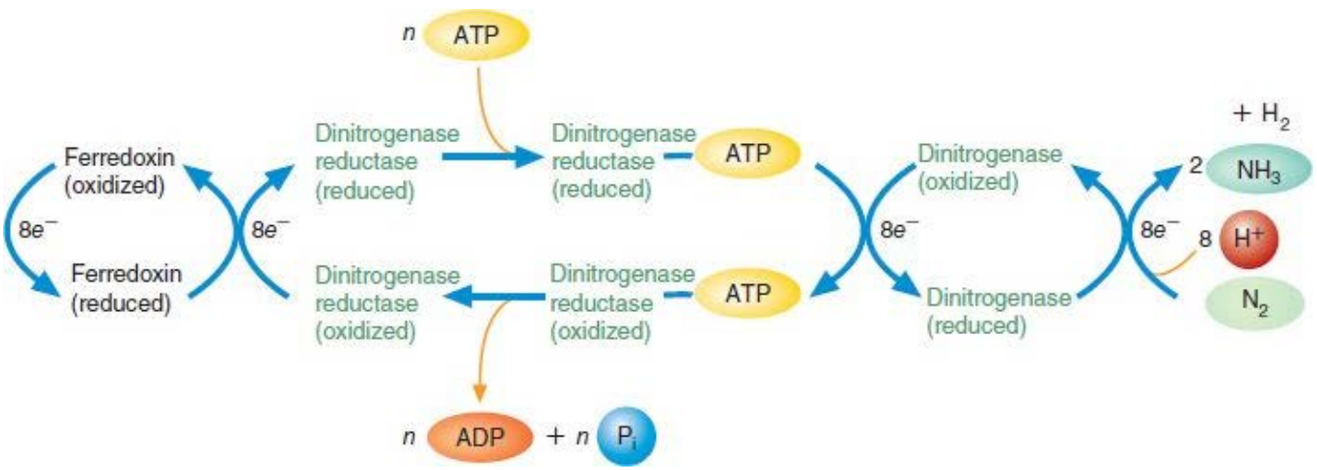


FIGURE 14.3

Schematic Diagram of the Nitrogenase Complex Illustrating the Flow of Electrons and Energy in Enzymatic Nitrogen Fixation

Several pathways can provide reducing equivalents for the reduction of ferredoxin. In cyanobacteria exposed to light, ferredoxin is reduced by PSI (cyclic photophosphorylation, which supplies both reduced ferredoxin and ATP). The high energy of activation of nitrogen fixation is overcome by a large number of ATP molecules (about 16 ATP per N_2 molecule). Both the binding of ATP to dinitrogenase reductase (Fe protein) and its subsequent hydrolysis cause conformational changes in the protein that facilitate the transfer of electrons to dinitrogenase (P

cluster/MoFe cofactor).

In addition to the large quantities of ATP that are required to drive the reduction of N_2 (a minimum of 16 ATP), a substantial number of copies of nitrogenase complex proteins must be synthesized because of the enzyme's slow turnover time: about 6 NH_3 produced per second per molecule of enzyme. (Nitrogenase complex protein may constitute as much as 20% of cellular protein in diazotrophs, the bacteria that fix nitrogen.) As a result, the regulation of nitrogen fixation by diazotrophs, which must respond to numerous environmental variations (e.g., fixed nitrogen and oxygen levels and carbon source availability), is both intricate and stringent. The major form of regulation is transcriptional control of the approximately 20 nitrogen fixation (*nif*) genes. In addition to coding for dinitrogenase reductase (Fe protein) and the α and β subunits of dinitrogenase (MoFe protein), *nif* genes code for a variety of enzymes that synthesize components of the nitrogen fixation process, such as metal clusters, homocitrate, or ferredoxin, and regulatory proteins that mediate responses to environmental cues.

WORKED PROBLEM 14.1

How expensive is nitrogen fixation? Determine the number of ATPs required to reduce a nitrogen atom to NH_4^+. Assume that 4 NADH supply the electrons for the reduction of N_2 and one NADH is the equivalent of 2.5 ATPs.

SOLUTION

Refer to the overall equation for nitrogen fixation on p. 529. The reduction of a molecule of N_2 to 2 NH_4^+ requires 8 electrons and 16 ATP. Once the number of ATP equivalents (4 × 2.5 ATP = 10) for 4 NADHs (4 × 2 electrons = 8) is calculated, the number of ATPs required for the reduction of N_2 rises to 26. The minimum number of ATPs required for the reduction of one nitrogen atom is half of this number, or 13 per NH_4^+ formed. ■

Nitrogen Assimilation

Nitrogen assimilation is the incorporation of inorganic nitrogen compounds into organic molecules. In plants, nitrogen assimilation begins in the roots, whether by transfer of NH_4^+ (ammonium ions) from symbiotic bacteria in the root nodules of leguminous plants or by absorption of NH_4^+ or NO_3^- (nitrate) from soil. Nitrate is produced by nitrifying soil bacteria. Organisms such as *Nitrosomonas* oxidize NH_4^+ to form NO_2^- (nitrite), which is then further oxidized to NO_3^- by bacteria such as *Nitrobacter*.

The assimilation of inorganic nitrogen into biomolecules in plants is effected by the incorporation of ammonium nitrogen into amino acids. Glutamine synthetase, the most important enzyme in nitrogen assimilation, catalyzes the ATP-dependent reaction of glutamate with NH_4^+ to form glutamine. In the next step, glutamine reacts with α-ketoglutarate to form glutamate. (The synthesis of other amino acids via the transfer of the amino group of glutamate is described in Section 14.2.) When nitrate is the nitrogen source, it must first be converted to NH_4^+ in a two-step process. After nitrate has been converted to nitrite by nitrate reductase, ammonia is produced by the reduction of nitrite, catalyzed by nitrite reductase.

QUESTION 14.1

The nitrogenase complex can reduce molecules other than N_2. Provide the structures for the products for each of the following substrates (real and hypothetical) that contain triple bonds: hydrogen cyanide, dinitrogen, and acetylene.

QUESTION 14.2

A red heme-containing protein called leghemoglobin is found in the nitrogen-fixing root nodules of leguminous plants. The protein component is produced by the plant, whereas bacterial cells produce the precursor of heme. Can you deduce the function of leghemoglobin? [*Hint*: Leghemoglobin is a member of the globin superfamily.]

14.2 AMINO ACID BIOSYNTHESIS

Living organisms differ in their capacity to synthesize the amino acids required for protein synthesis. Although plants and many microorganisms can produce all their amino acids from readily available precursors, other organisms must obtain some preformed amino acids from their environment. Animals can synthesize only about half the amino acids they require. These **nonessential amino acids (NAAs)** are synthesized from readily available molecules such as glycerate-3-phosphate (a glycolytic intermediate) and oxaloacetate (a citric acid cycle intermediate). The amino acids that must be provided in the diet are referred to as **essential amino acids (EAAs)**. Mammalian tissues can synthesize NAAs (**Table 14.1**) by relatively simple reaction pathways. In contrast, EAAs must be obtained from the diet because mammals lack the complicated reaction pathways required for their synthesis.

TABLE 14.1 The Essential And Nonessential Amino Acids In Humans

Essential	Nonessential
Isoleucine	Alanine
Leucine	Arginine*
Lysine	Asparagine
Methionine	Aspartate
Phenylalanine	Cysteine*
Threonine	Glutamate
Tryptophan	Glutamine
Valine	Glycine
Histidine	Proline
	Serine
	Tyrosine*

* Referred to as the *semi-essential amino acids*, these molecules are essential for infants and children up to five years of age. The pathways that produce the semi-essential amino acids are not fully functional in young children.

Amino Acid Metabolism Overview

Amino acids serve a number of functions. In addition to their most important role in the synthesis of proteins, amino acids are the principal source of the nitrogen atoms required in various synthetic reaction pathways. The non-nitrogen components of amino acids (referred to as carbon skeletons) are a source of energy, as well as precursors in several reaction pathways. Therefore, an adequate intake of amino acids, in the form of dietary protein, is essential for an animal's proper growth and development.

DIETARY AMINO ACIDS Dietary protein sources differ widely in their proportions of the EAA. In general, complete proteins (those containing sufficient quantities of EAA) are of animal origin (e.g., meat, milk, and eggs). Plant proteins often lack one or more EAA. For example, gliadin (wheat protein) has insufficient amounts of lysine, and zein (corn protein) is low in both lysine and tryptophan. Because plant proteins differ in their amino acid compositions, plant foods can provide a high-quality source of essential amino acids only if they are eaten in appropriate combinations. One such combination consists of beans (low in methionine) and cereal grains (low in lysine). Following the digestion of dietary protein in the body's digestive tract, free amino acids are transported across intestinal enterocytes and into the blood. Most diets do not provide amino acids in the proportions that the body requires. Their concentrations, therefore, must be adjusted by metabolic mechanisms. The amino acids released to the blood in the intestine already show some changes in their relative concentrations. The intestinal mucosa is a very active and constantly replaced tissue that sustains its structure and function by means of incoming nutrients other than glucose. The amino acid glutamine, for example, is a primary energy source for enterocytes. The blood from the GI tract goes first to the liver, another active tissue. The liver synthesizes the serum proteins, among others, and draws amino acids from the blood for this purpose. It also preferentially uses amino acids (especially alanine and serine) to synthesize glucose for export.

BRANCHED CHAIN AMINO ACIDS AND THE AMINO ACID POOL The blood that leaves the liver to nourish the rest of the body has a much higher concentration of **branched-chain amino acids (BCAAs)**—leucine, isoleucine, and valine—than does the blood leaving the GI tract. The BCAA are EAAs and provide critical hydrophobic side chains in protein structure (e.g., leucine zipper motifs in DNA-binding proteins). BCAA also represent a major transport form of amino nitrogen from the liver to other tissues, where they are used in the synthesis of the NAAs required for protein synthesis, as well as various amino acid derivatives.

Amino acid metabolism is a complex series of reactions in which the amino acid molecules required for syntheses of proteins and metabolites are continuously being synthesized and degraded. Depending on current metabolic requirements, certain amino acids are synthesized or interconverted and then transported to tissues, where they are used. The amino acid molecules that are immediately available for use in metabolic processes are referred to as the **amino acid pool**. In animals, amino acids in the pool are derived from the breakdown of both dietary and tissue proteins.

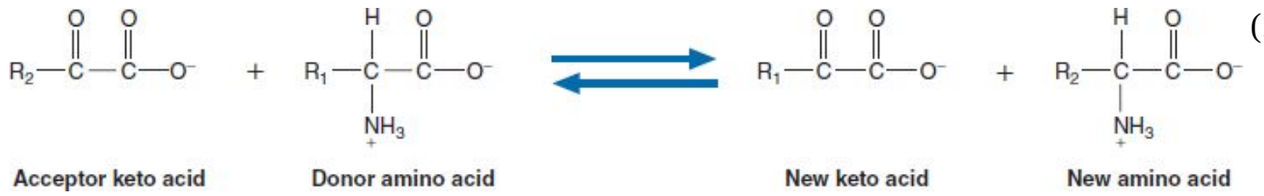
AMINO ACID TRANSPORT Transport of amino acids into cells is mediated by specific membrane-bound transport proteins, several of which have been identified in mammalian cells. They differ in their specificity for the types of amino acid transported and in whether the transport process is linked to the movement of Na^+ across the plasma membrane. (Recall that the gradient created by the active transport of Na^+ can move molecules across the membrane. Na^+-dependent amino acid transport is similar to that observed in the glucose transport process illustrated in **Figure 11.30**.) For example, several Na^+-dependent transport systems have been identified in the lumenal plasma membrane of enterocytes. Na^+-independent transport systems are responsible for transporting amino acids across the portion of enterocyte plasma membrane in contact with blood vessels. The γ-glutamyl cycle (see

Figure 14.20 later in this chapter) is believed to assist in transporting some amino acids into cells in specific tissues (i.e., brain, intestine, and kidney).

Reactions of Amino Groups

Once amino acid molecules have entered cells, the amino groups are available for numerous synthetic reactions. This metabolic flexibility is effected by transamination reactions and reactions in which NH_4^+ or the amide nitrogen of glutamine is used to supply the amino group or the amide nitrogen of certain amino acids. Both reaction types are discussed next.

TRANSAMINATION **Transamination** reactions dominate amino acid metabolism. In these reactions, catalyzed by a group of enzymes referred to as the *aminotransferases* or *transaminases*, α-amino groups are transferred from an α-amino acid to an α-keto acid:



(1)

Transamination reactions, which are readily reversible, play important roles in both the synthesis and degradation of the amino acids.

Eukaryotic cells possess a large variety of aminotransferases. Found within both the cytoplasm and the mitochondria, these enzymes possess two types of specificity: (1) the type of α-amino acid that donates the α-amino group and (2) the α-keto acid that accepts the α-amino group. Although the aminotransferases vary widely in the type of amino acid they bind, most of them use glutamate as the amino group donor. Because glutamate is produced when α-ketoglutarate (a citric acid cycle intermediate) accepts an amino group, these two molecules, the *α-ketoglutarate/glutamate pair*, have a strategically important role in both amino acid metabolism and metabolism in general. Two other such pairs have important functions in metabolism. In addition to its role in transamination reactions, the *OAA/aspartate pair* is involved in the disposal of nitrogen in the urea cycle (Chapter 15). One of the most important functions of the *pyruvate/alanine pair* is in the glucose–alanine cycle (**Figure 8.13**). Since α-ketoglutarate and OAA are citric acid cycle intermediates, transamination reactions often represent an important mechanism for meeting cellular energy requirements. Recall, for example, that in the glucose–alanine cycle transamination reactions are used to recycle the carbon skeleton of pyruvate between muscle and liver.

Transamination reactions require the coenzyme pyridoxal-5′-phosphate (PLP), which is derived from pyridoxine (vitamin B_6). PLP is also required in numerous other reactions of amino acids. Examples include racemizations, decarboxylations, and several side chain modifications. (**Racemizations** are reactions in which mixtures of L- and D-amino acids are formed.) The structures of the vitamin and its coenzyme form are illustrated in **Figure 14.4**.

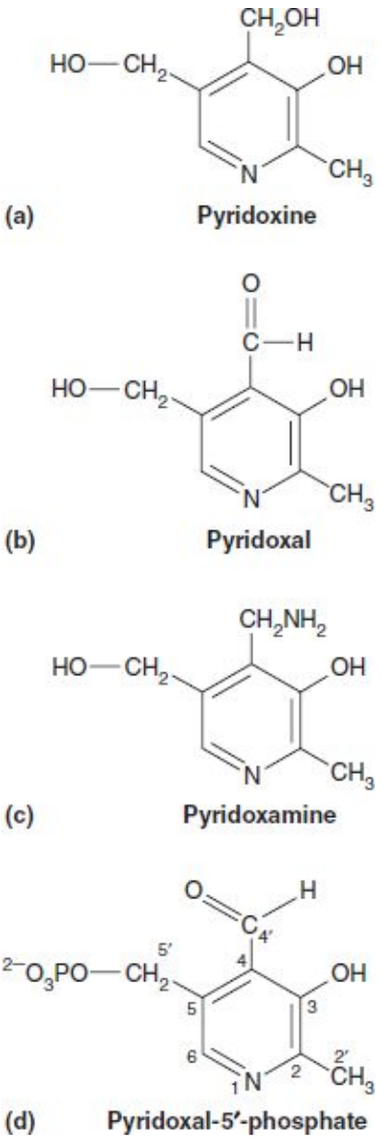


FIGURE 14.4

Vitamin B_6

Vitamin B_6 includes (a) pyridoxine, (b) pyridoxal, and (c) pyridoxamine. (Pyridoxine is found in leafy green vegetables. Pyridoxal and pyridoxamine are found in animal foods such as fish, poultry, and red meat.) The biologically active form of vitamin B_6 is (d) pyridoxal-5′-phosphate.

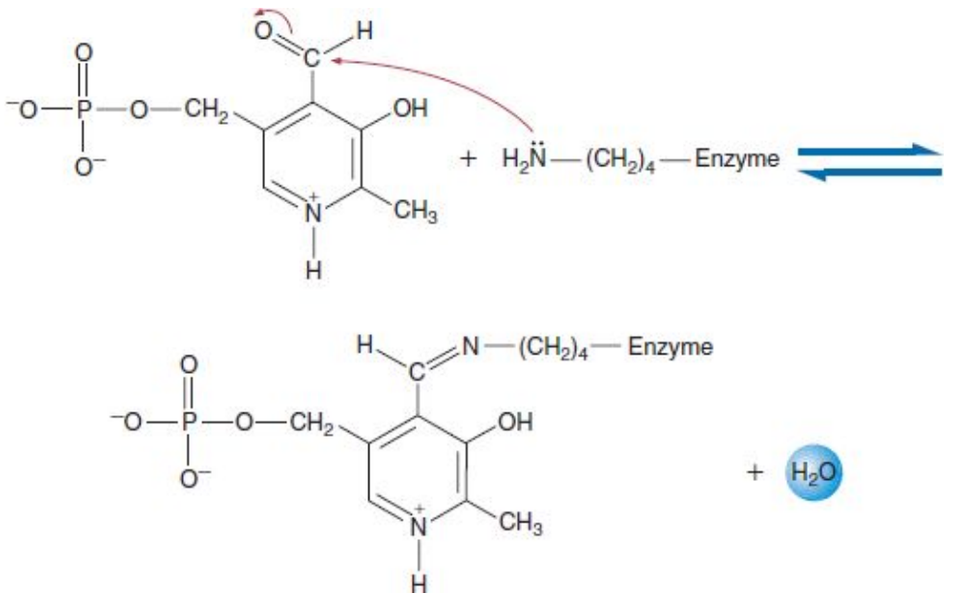
3D animation of pyridoxine

3D animation of pyridoxal

3D animation of pyridoxamine

3D animation of Pyridoxal-5'-phosphate

PLP is bound in the transaminase active site as a Schiff base (R′—CH==N—R, an aldimine) formed by condensation of the aldehyde group of PLP and the ε-amino group of a lysine residue.



Additional stabilizing forces include ionic interactions between amino acid side chains and PLP's pyridinium ring and phosphate group. The positively charged pyridinium ring also functions as an electron sink, stabilizing negatively charged reaction intermediates.

Amino acid substrates become bound to PLP via the α-amino group in an imine exchange reaction. Then one of three bonds in the substrate is selectively broken in the active sites in each type of PLP-dependent enzyme (**Figure 14.5**).

THE TRANSAMINATION MECHANISM The transamination reaction begins with formation of a Schiff base between PLP and the α-amino group of an α-amino acid (**Figure 14.6**). When the α-hydrogen atom is removed by a general base in the enzyme active site, a resonance-stabilized intermediate forms. With the donation of a proton from a general acid and a subsequent hydrolysis, the newly formed α-keto acid is released from the enzyme. A second α-keto acid then enters the active site and is converted into an α-amino acid in a reversal of the reaction process that has just been described. Transamination reactions are examples of a reaction mechanism referred to as a double displacement, or *ping-pong reaction* (p. 225). The mechanism is so named because the first substrate must leave the active site before the second one can enter.

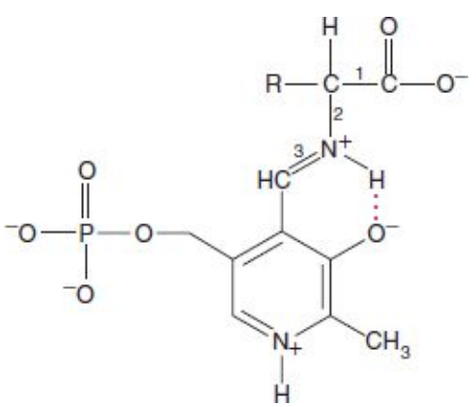
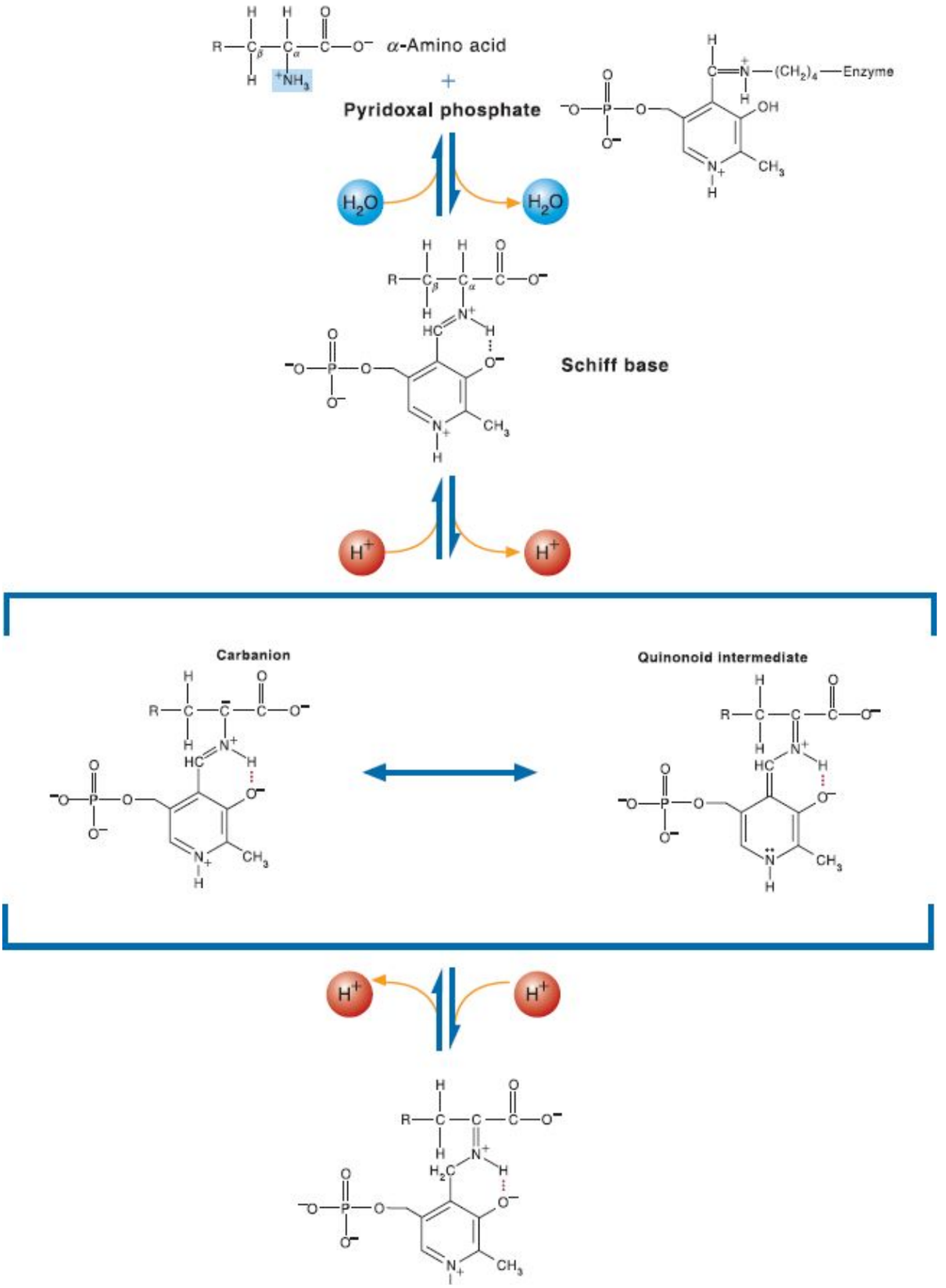


FIGURE 14.5

Intermediate Schiff Base Formed between Pyridoxine and an Amino Acid

When an amino acid binds to pyridoxine in the active site of a PLP-dependent enzyme, one of three bonds will break. This selectivity depends on the presence or absence of a nearby base catalyst and the orientation of the amino acid in the active site. If an initial deprotonation of the α-carbon of the amino group donor occurs, then transamination (bond 2 broken) or racemization or elimination (bond 3 broken) may occur. If the initial deprotonation does not occur, then decarboxylation results (bond 1 broken).



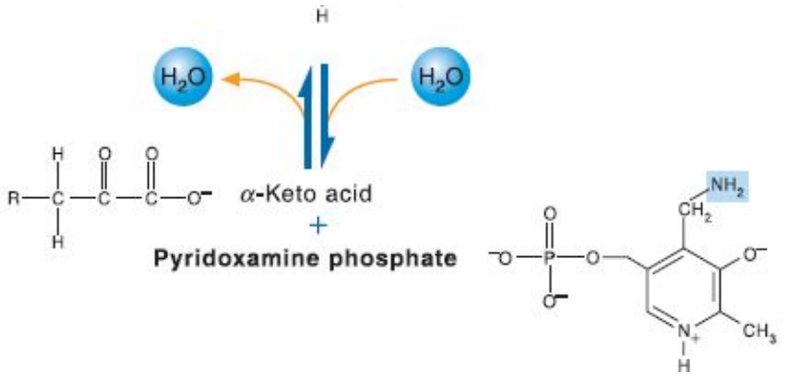


FIGURE 14.6

The Transamination Mechanism

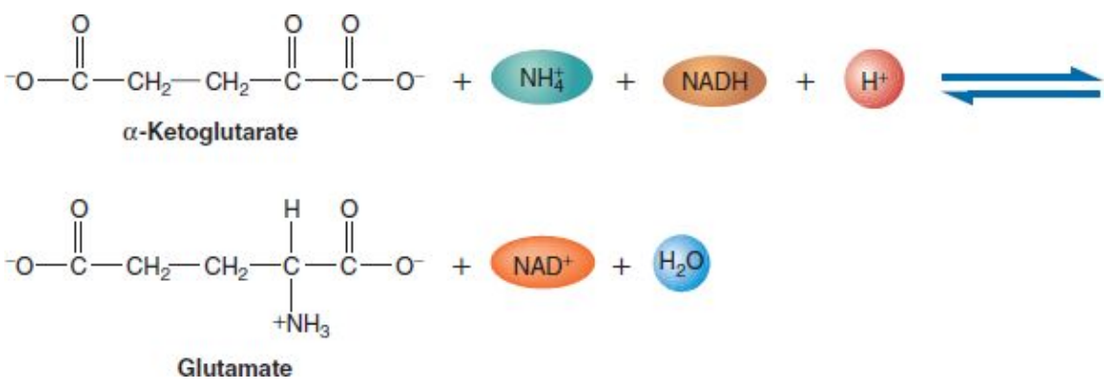
The donor amino acid forms a Schiff base with the coenzyme pyridoxal phosphate within the enzyme's active site. A proton is lost, and a carbanion forms that is resonance-stabilized by interconversion to a quinonoid intermediate. After an enzyme-catalyzed proton transfer and a hydrolysis, the α-keto product is released. A second α-keto acid then enters the active site. This acceptor α-keto acid is converted to an α-amino acid product as the mechanism just described is reversed. Note that the chirality of the donor amino acid is preserved in the α-amino acid product. Within the active site, the orientation of the quinonoid intermediate allows the proton to be added in a manner that confers on the resulting product, the Schiff base, an L-configuration.

Transamination reactions are reversible. It is, therefore, theoretically possible for all amino acids to be synthesized by transamination. However, experimental evidence indicates that there is no net synthesis of an amino acid if its α-keto acid precursor is not independently synthesized by the organism. For example, alanine, aspartate, and glutamate are nonessential for animals because their α-keto acid precursors (i.e., pyruvate, OAA, and α-ketoglutarate) are readily available metabolic intermediates. Because the reaction pathways for synthesizing molecules such as phenylpyruvate, α-keto-β-hydroxybutyrate, and imidazole pyruvate do not occur in animal cells, phenylalanine, threonine, and histidine must be provided in the diet. (Reaction pathways that synthesize amino acids from metabolic intermediates, and not by transamination, are referred to as de novo pathways.)

DIRECT INCORPORATION OF AMMONIUM IONS INTO ORGANIC MOLECULES

Ammonium ions are incorporated into amino acids and eventually other metabolites through two principal means: (1) reductive amination of α-keto acids and (2) formation of the amides of aspartic and glutamic acid with subsequent transfer of the amide nitrogen to form other amino acids.

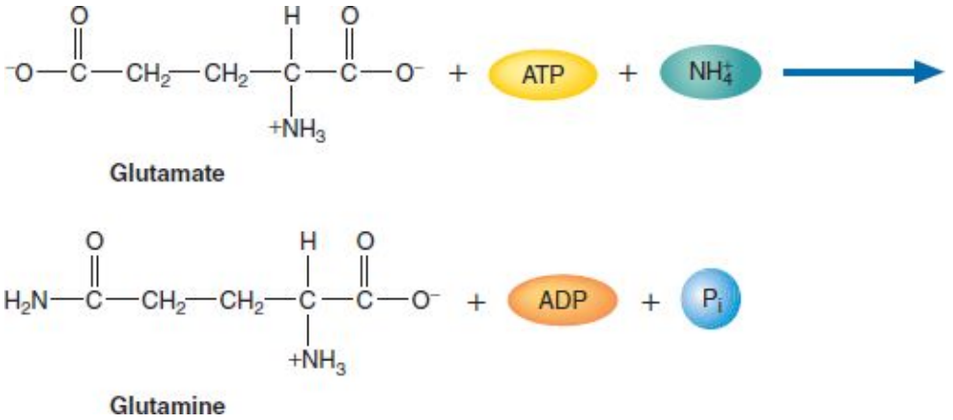
Glutamate dehydrogenase, an enzyme found in both the mitochondria and cytoplasm of eukaryotic cells, as well as in some bacterial cells, catalyzes the direct amination of α-ketoglutarate:



(2)

The primary function of this enzyme in eukaryotes appears to be catabolic (i.e., a means of producing NH_4^+ in preparation for nitrogen excretion). However, the reaction is reversible. When excess ammonia is present, the reaction is driven toward glutamate synthesis.

Ammonium ions are also incorporated into cell metabolites by the formation of glutamine, the amide of glutamate:



(3)

The brain, a rich source of the enzyme glutamine synthetase, is especially sensitive to the toxic effects of NH_4^+. Brain cells convert NH_4^+ and glutamate to glutamine, a neutral, nontoxic molecule. Glutamine is then transported to the liver, where the amide nitrogen is released as NH_4^+ (p. 576). Ammonium ion is then disposed of by incorporation into urea, the principal nitrogenous waste product in mammals (pp. 576–80).

KEY CONCEPTS

- In transamination reactions, amino groups are transferred from one carbon skeleton to another.
- In reductive amination, amino acids are synthesized by the incorporation of free NH_4^+ or the amide nitrogen of glutamine or asparagine into α-keto acids.
- Ammonium ions are also incorporated into cellular metabolites by the amination of glutamate to form glutamine.

Biosynthesis of the Amino Acids

Amino acid synthetic pathways are remarkably diverse, but they have one common feature. The carbon skeleton of each amino acid is derived from a commonly available metabolic intermediate. Thus, in animals, NAA molecules are derivatives of glycerate-3-phosphate, pyruvate, α-ketoglutarate, or OAA. Tyrosine, synthesized from the EAA phenylalanine, is an exception to this rule.

On the basis of the similarities in their synthetic pathways, the amino acids can be grouped into six families: glutamate, serine, aspartate, pyruvate, the aromatics, and histidine (**Figure 14.7**). The amino acids in each family are ultimately derived from one precursor molecule. In the discussions of amino acid synthesis that follow, the intimate relationship between amino acid metabolism and several other metabolic pathways is apparent. Amino acid biosynthesis is outlined in **Figure 14.8**.

THE GLUTAMATE FAMILY The glutamate family includes—in addition to glutamate—glutamine, proline, and arginine. As described, α-ketoglutarate may be converted to glutamate by reductive amination and by transamination reactions involving a number of amino acids. Although the relative contribution of these reactions to glutamate synthesis varies with cell type and metabolic circumstances, transamination appears to play a major role in the synthesis of most glutamate molecules in eukaryotic cells.

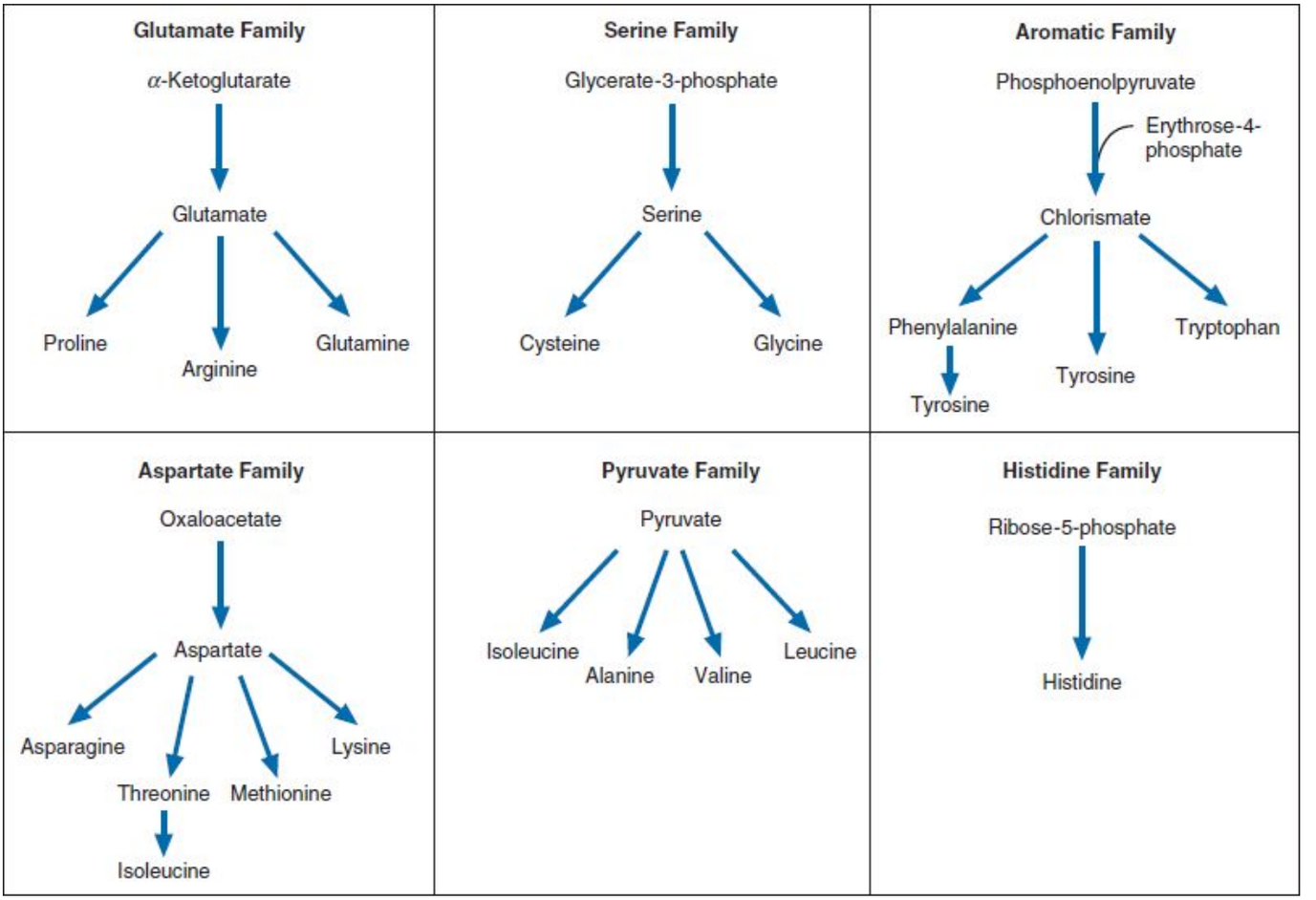


FIGURE 14.7

The Amino Acid Biosynthetic Families

Each family of amino acids is derived from a common precursor molecule.

3D animation of glutamate

3D animation of pyruvate

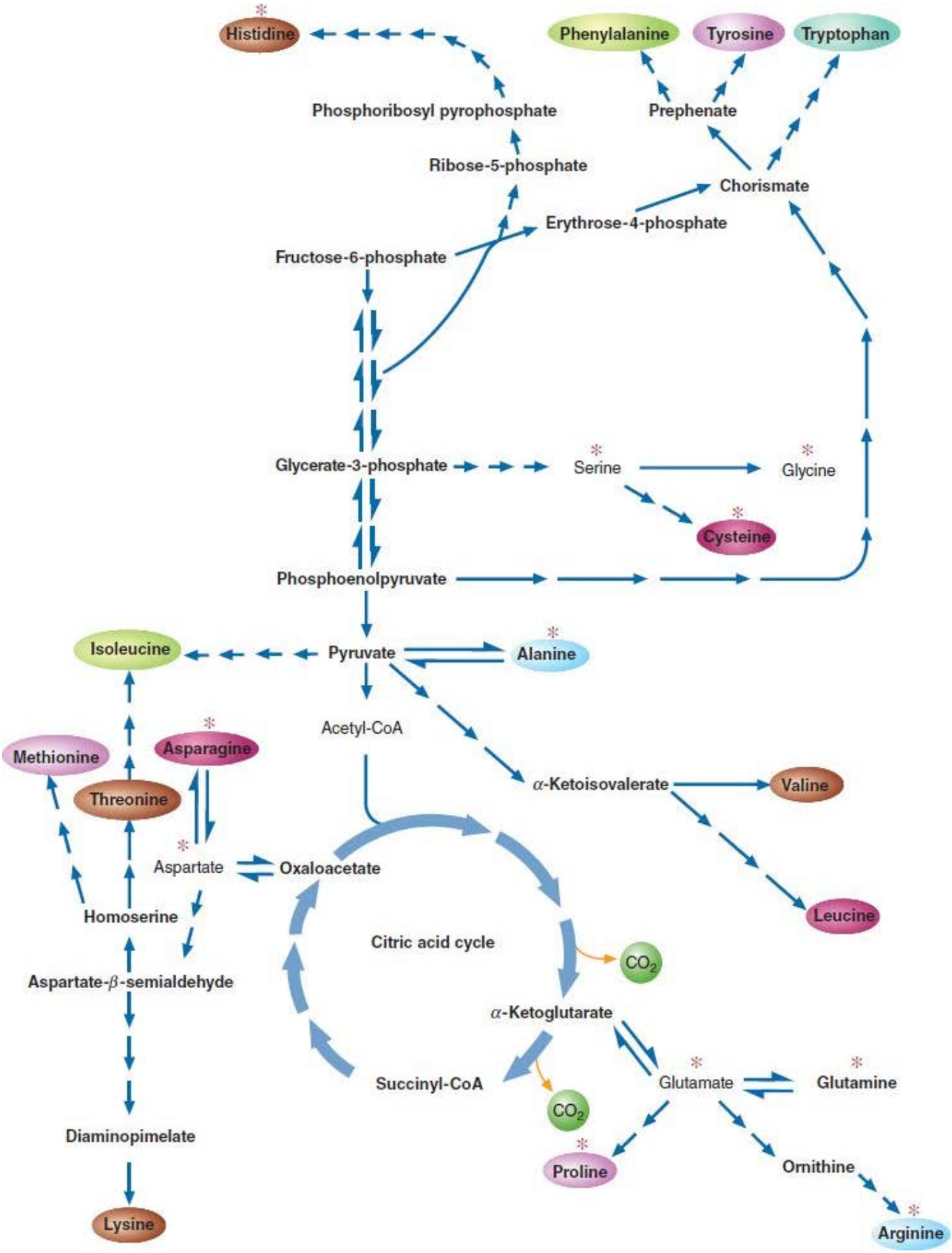


FIGURE 14.8

Biosynthesis of the Amino Acids

Intermediates in the central metabolic pathways provide the carbon skeleton precursor molecules required for

synthesis of each amino acid. The number of reactions in each pathway is indicated. The nonessential amino acids for mammals are indicated by asterisks. (In mammals, tyrosine can be synthesized from phenylalanine.)

BCAA are an important source of amino groups in glutamine synthesis. As mentioned, blood that leaves the liver is selectively enriched in BCAA. Many more BCAA are taken up by peripheral tissues than are needed for protein synthesis. BCAA amino groups may be used primarily for synthesis of glutamate. The conversion of glutamate to glutamine, catalyzed by glutamine synthetase, takes place in liver, brain, kidney, muscle, and intestine. In addition to its role in protein synthesis, glutamine is the amino group donor in numerous biosynthetic reactions (e.g., purine, pyrimidine, and amino sugar syntheses) and, as previously mentioned, as a safe storage and transport form of NH_4^+. Glutamine is therefore a major metabolite in living organisms. Other functions of glutamine vary, depending on the cell type being considered. For example, in the kidney and small intestine, glutamine is a major source of energy. In the enterocytes of the small intestine, approximately 55% of glutamine carbon is oxidized to CO_2.

Proline is a cyclized derivative of glutamate. As shown in **Figure 14.9**, a γ-glutamyl phosphate intermediate is reduced to glutamate-γ-semialdehyde. The enzyme catalyzing the phosphorylation of glutamate (γ-glutamyl kinase) is regulated by negative feedback inhibition by proline. Glutamate-γ-semialdehyde cyclizes spontaneously to form Δ^1-pyrroline-5-carboxylate. Then Δ^1-pyrroline-5-carboxylate reductase catalyzes the reduction of Δ^1-pyrroline-5-carboxylate to form proline. The interconversion of Δ^1-pyrroline-5-carboxylate and proline may act as a shuttle mechanism to transfer reducing equivalents derived from the pentose phosphate pathway into mitochondria. This process may partially explain the high turnover of proline in many cell types. Proline can also be synthesized from ornithine, a urea cycle intermediate. The enzyme catalyzing ornithine's conversion to glutamate-γ-semialdehyde, ornithine aminotransferase, is found in relatively high concentration in fibroblasts where the demand for proline incorporation into collagen is high.

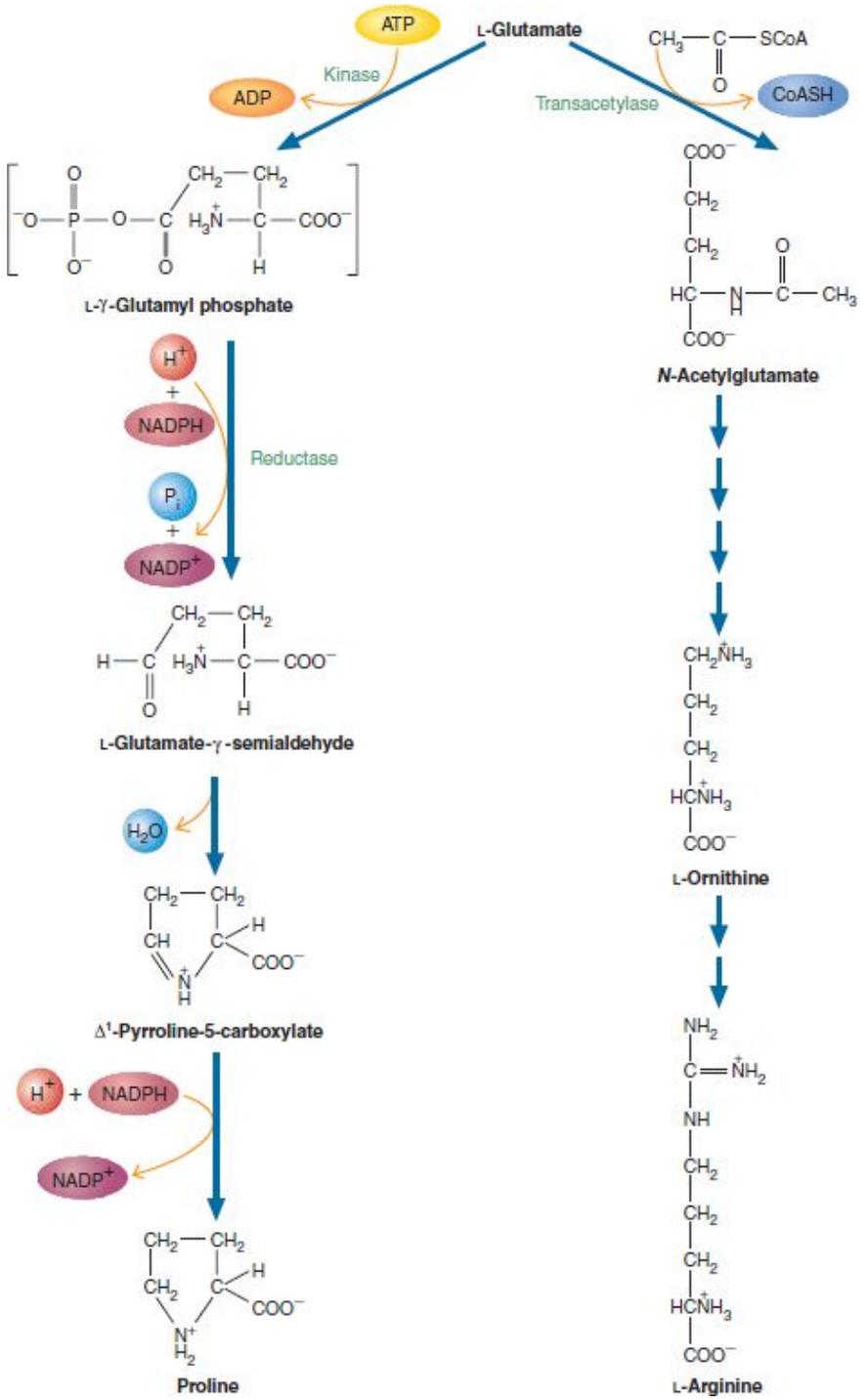


FIGURE 14.9

Biosynthesis of Proline and Arginine from Glutamate

Proline is synthesized from glutamate in four steps. The third step is a spontaneous cyclization reaction. In arginine synthesis, the acetylation of glutamate prevents the cyclization reaction. In mammals, the reactions that convert ornithine to arginine are part of the urea cycle.

Arginine synthesis begins with the acetylation of the α-amino group of glutamate. *N*-acetylglutamate is then converted to ornithine in a series of reactions that include a phosphorylation, a reduction, a transamination, and a deacetylation (removal of an acetyl group). The subsequent reactions in which ornithine is converted to arginine are part of the urea cycle (p. 577). In infants, in whom the urea cycle is insufficiently functional, arginine is an EAA.

THE SERINE FAMILY The members of the serine family—serine, glycine, and cysteine—derive

their carbon skeletons from the glycolytic intermediate glycerate-3-phosphate. The members of this group play important roles in numerous anabolic pathways. Serine is a precursor of ethanolamine and sphingosine. Glycine is used in the purine, porphyrin, and glutathione synthetic pathways. Together, serine and glycine contribute to a series of biosynthetic pathways that are referred to collectively as one-carbon metabolism (p. 543). Cysteine plays a significant role in sulfur metabolism (p. 586).

Serine is synthesized in a direct pathway from glycerate-3-phosphate that involves dehydrogenation, transamination, and hydrolysis by a phosphatase (**Figure 14.10**). Cellular serine concentration controls the pathway through feedback inhibition of phosphoglycerate dehydrogenase and phosphoserine phosphatase. The last-named enzyme catalyzes the only irreversible step in the pathway.

The conversion of serine to glycine consists of a single complex reaction catalyzed by serine hydroxymethyltransferase, a pyridoxal phosphate-requiring enzyme. During the reaction, which is an aldol cleavage, serine binds to pyridoxal phosphate. The reaction yields glycine and a chemically reactive formaldehyde group that is transferred to the coenzyme tetrahydrofolate (THF) (p. 545) to form N^5,N^{10}-methylene THF. Serine is the major source of glycine. Smaller amounts of glycine can be derived from choline when choline is present in excess. The synthesis of glycine from choline consists of two dehydrogenations and a series of demethylations.

Cysteine synthesis is a primary component of sulfur metabolism (p. 586). The carbon skeleton of cysteine is derived from serine (**Figure 14.11**). In animals the sulfhydryl group is transferred from methionine by way of a demethylated derivative homocysteine. Both enzymes involved in the conversion of serine to cysteine (cystathionine β-synthase, or CBS, and γ-cystathionase, or CSE) require pyridoxal phosphate.

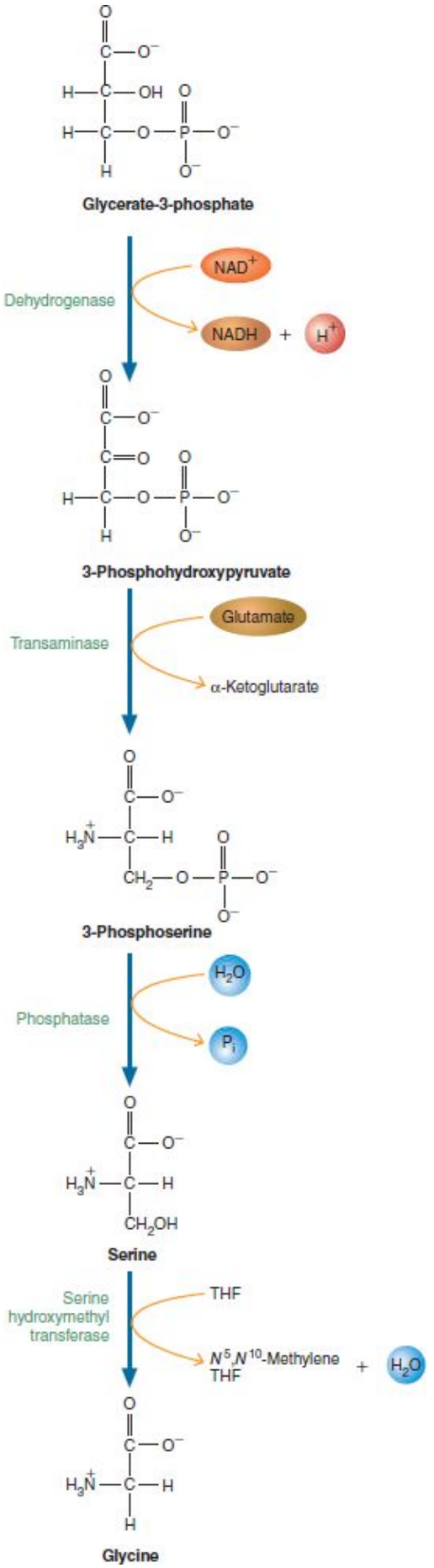


FIGURE 14.10

Biosynthesis of Serine and Glycine

Serine biosynthesis begins with the oxidation of glycerate-3-phosphate. The carbonyl-containing product 3-phosphohydroxypyruvate then undergoes a transamination reaction with glutamate to yield 3-phosphoserine. The hydrolysis of 3-phosphoserine yields serine. Serine hydroxymethyl transferase catalyzes the conversion of serine to glycine and N^5, N^{10}-methylene THF. Serine inhibits glycerate-3-phosphate dehydrogenase, the first reaction in the pathway.

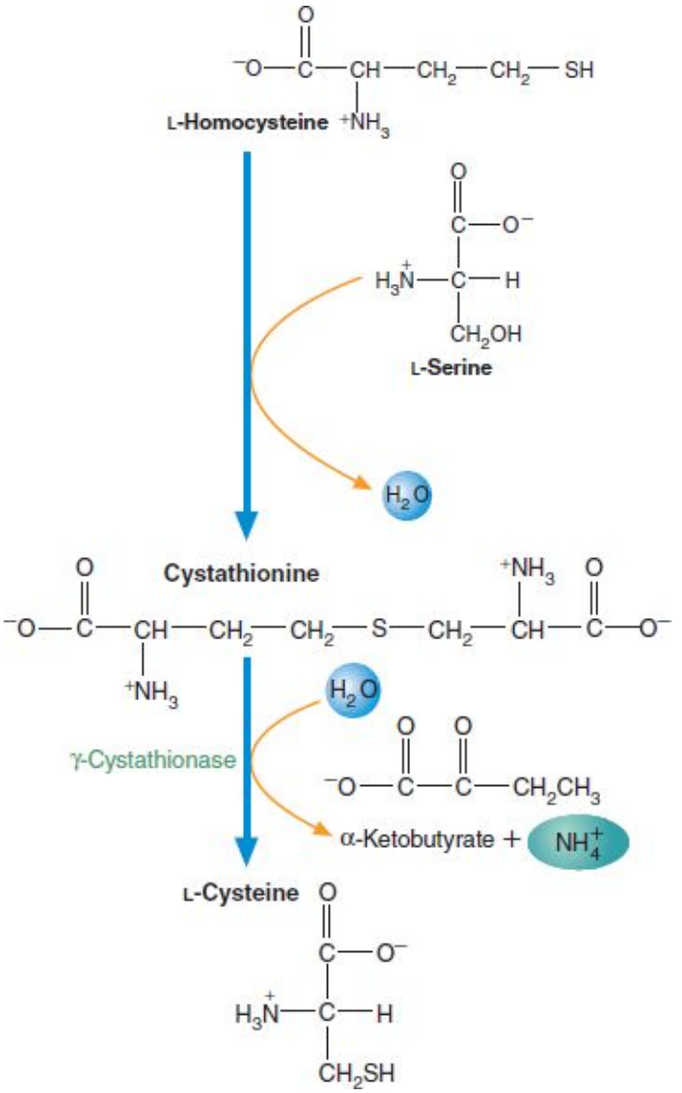


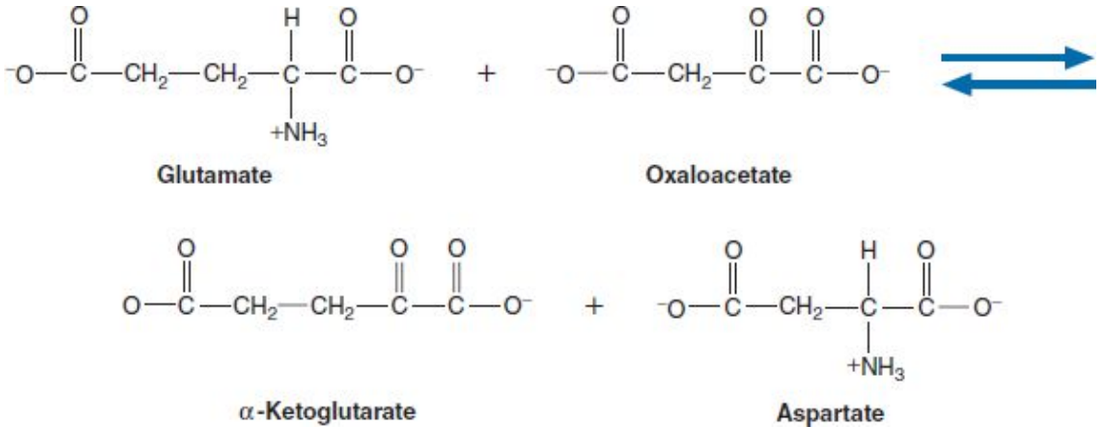
FIGURE 14.11

Biosynthesis of Cysteine

In animals, serine condenses with homocysteine (derived from methionine) to form cystathionine in a reaction catalyzed by cystathionine β-synthase (CBS). γ-Cystathionase (CSE) catalyzes the cleavage of cystathionine to yield cysteine, α-ketobutyrate, and NH_4^+.

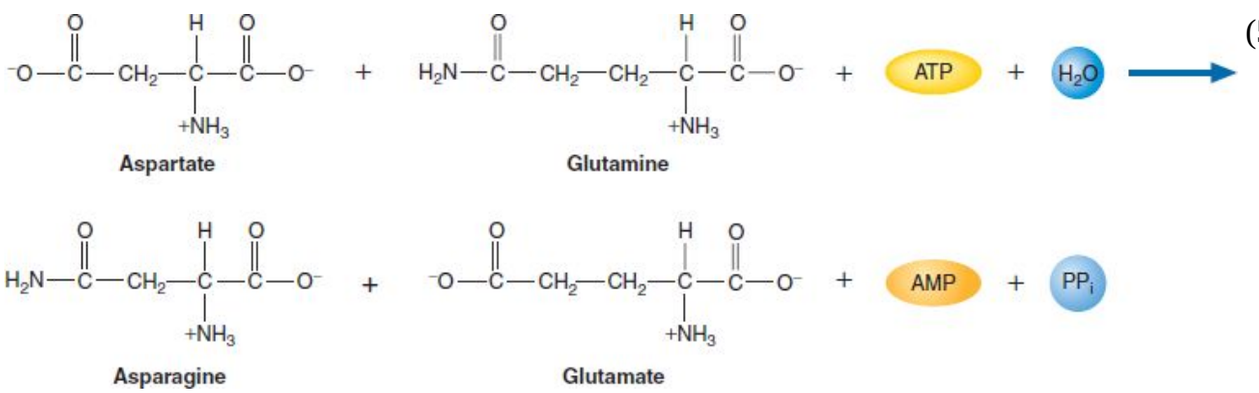
THE ASPARTATE FAMILY Aspartate, the first member of the aspartate family of amino acids, is derived from OAA in a transamination reaction:

(4)



Aspartate transaminase (AST; also known as glutamic oxaloacetic transaminase, or GOT) is the most active of the aminotransferases, found in most cells in both mitochondria and the cytoplasm. The reaction that AST catalyzes is reversible and significantly influences the flow of carbon and nitrogen within the cell. For example, excess glutamate is converted via AST to aspartate. Aspartate is then used as a source of both nitrogen (for urea formation) and the citric acid cycle intermediate OAA. Aspartate is also an important precursor in nucleotide synthesis.

The aspartate family also contains asparagine and the essential amino acids lysine, methionine, and threonine. Asparagine, the amide of aspartate, is not formed directly from aspartate and NH_4^+. Instead, the amide group of glutamine is transferred by amide group transfer during an ATP-requiring reaction catalyzed by asparagine synthase:



(5)

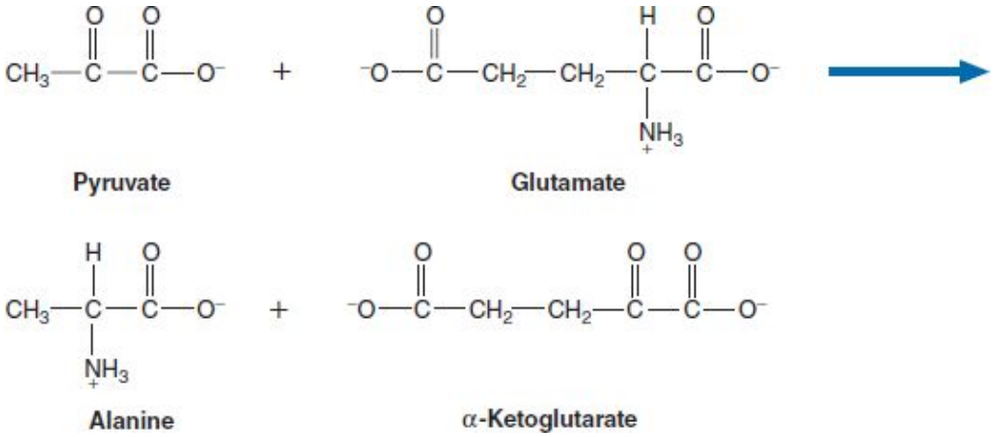
The biosynthesis of the EAA in the aspartate family (lysine, methionine, and threonine) is described briefly in the online Biochemistry in Perspective essay The Essential Amino Acids.

Threonine contributes to the reaction pathway in which isoleucine is synthesized. The synthesis of isoleucine, often considered a member of the pyruvate family, is also discussed in the online reading noted above.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on the essential amino acids.

THE PYRUVATE FAMILY The pyruvate family consists of alanine, valine, leucine, and isoleucine. Alanine is synthesized from pyruvate in a single step:

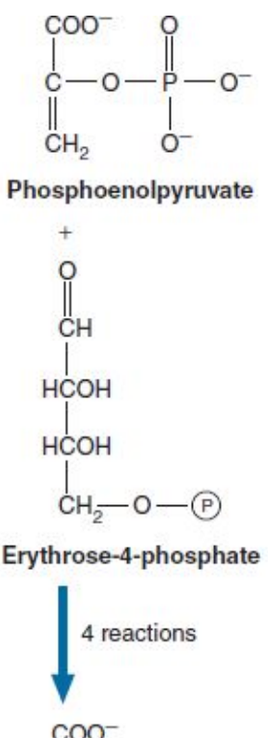
(6)



3D animation of alanine

Although the enzyme that catalyzes this reaction, alanine aminotransferase, has cytoplasmic and mitochondrial forms, the majority of its activity has been found in the cytoplasm. Recall that the glucose-alanine cycle (Chapter 8) contributes to the maintenance of blood glucose. BCAAs (leucine, isoleucine, and valine) are the ultimate source of many of the amino groups transferred from glutamate in the alanine cycle. Their biosynthesis is described in the online reading The Essential Amino Acids.

THE AROMATIC FAMILY The aromatic family of amino acids includes phenylalanine, tyrosine, and tryptophan. Of these, only tyrosine is considered nonessential in mammals. Either phenylalanine or tyrosine is required for the synthesis of dopamine, epinephrine, and norepinephrine, an important class of biologically potent molecules referred to as the *catecholamines*. (See the online essay entitled The Catecholamines.) Tryptophan is a precursor in the synthesis of NAD^+, $NADP^+$, and the neurotransmitter serotonin.



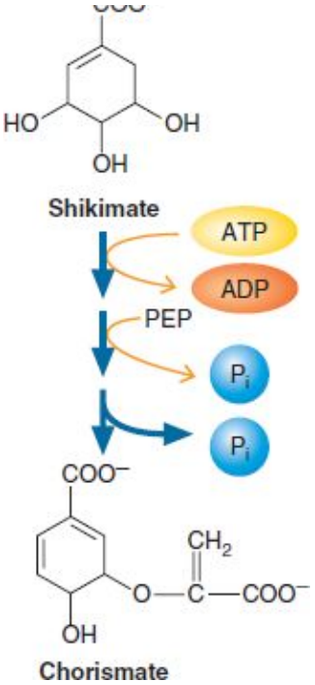


FIGURE 14.12

Chorismate Biosynthesis

Chorismate is an intermediate in the shikimate pathway. The formation of chorismate involves the ring closure of an intermediate (not shown) and the subsequent creation of two double bonds. The side chain of chorismate is derived from phosphoenolpyruvate (PEP).

The benzene ring of the aromatic amino acids is formed by the *shikimate pathway* (**Figure 14.12**). The carbons in the benzene ring are derived from erythrose-4-phosphate and PEP. These two molecules condense to form a molecule that is subsequently converted to chorismate. Chorismate is the branch point in the syntheses of various aromatic compounds. For example, the aromatic rings in the mixed terpenoids (e.g., the tocopherols, the ubiquinones, and plastoquinone) are derived from chorismate.

Tyrosine is not an EAA in animals because it is synthesized from phenylalanine in a hydroxylation reaction. The enzyme involved, phenylalanine-4-monooxygenase, requires the coenzyme tetrahydrobiopterin (BH_4, pp. 552 and 584), a folic acid-like molecule derived from GTP. Because this reaction is also a first step in phenylalanine catabolism, it is discussed further in Chapter 15. The synthesis of the aromatic family of amino acids is described in the online reading The Essential Amino Acids.

HISTIDINE Histidine is considered nonessential in prokaryotes and plants. In humans and many animals, histidine must be provided by the diet. Histidine contributes substantially to protein structure and function because of its unique chemical properties. Recall, for example, that histidine residues bind heme prosthetic groups in hemoglobin. In addition, histidine often acts as a general acid during enzyme-catalyzed reactions.

Of all the amino acids, histidine's biosynthesis is the most unusual. Histidine is synthesized from phosphoribosylpyrophosphate (PRPP), ATP, and glutamine (**Figure 14.13**). Synthesis begins with the condensation of PRPP with ATP to form phosphoribosyl-AMP, which is then hydrolyzed to phosphoribosyl-AMP. Subsequently, a hydrolysis reaction opens the adenine ring. After an isomerization and the transfer of an amino group from glutamine, imidazole glycerol phosphate is synthesized. (The other product of the latter reaction, 5′-phosphoribosyl-4-carboxamide-5-aminoimidazole, is used in the synthesis of purine nucleotides. See Section 14.3.) Histidine is produced from imidazole glycerol phosphate in a series of reactions that include a dehydration, a

transamination, a phosphorolysis, and an oxidation.

KEY CONCEPTS

- There are six families of amino acids: glutamate, serine, aspartate, pyruvate, the aromatics, and histidine.
- The nonessential amino acids are derived from precursor molecules available in many organisms.
- The essential amino acids are synthesized from metabolites produced only in plants and some microorganisms.

14.3 BIOSYNTHETIC REACTIONS INVOLVING AMINO ACIDS

As described, amino acids are precursors of many physiologically important nitrogen-containing molecules, in addition to serving as building blocks for polypeptides. The following discussion focuses on the syntheses of several examples of these molecules (neurotransmitters, glutathione, and nucleotides). The synthesis of heme and chlorophyll is described in a reading, available online. Because many biosynthetic processes involve the transfer of carbon groups, this section begins with a brief description of one-carbon metabolism.

One-Carbon Metabolism

One-carbon metabolism is a set of reactions in which single-carbon atoms are transferred from one molecule to another. Carbon atoms have several oxidation states. Those of biological interest are found in methanol (+1), formaldehyde (+2), and formate (+3). **Table 14.2** lists the equivalent one-carbon groups that are actually involved in biosynthetic reactions. The most important carriers of one-carbon groups in biosynthetic pathways are folic acid and *S*-adenosylmethionine. The metabolism of each is described briefly. (The function of biotin, a carrier of CO_2 groups, is discussed in Section 8.2.)

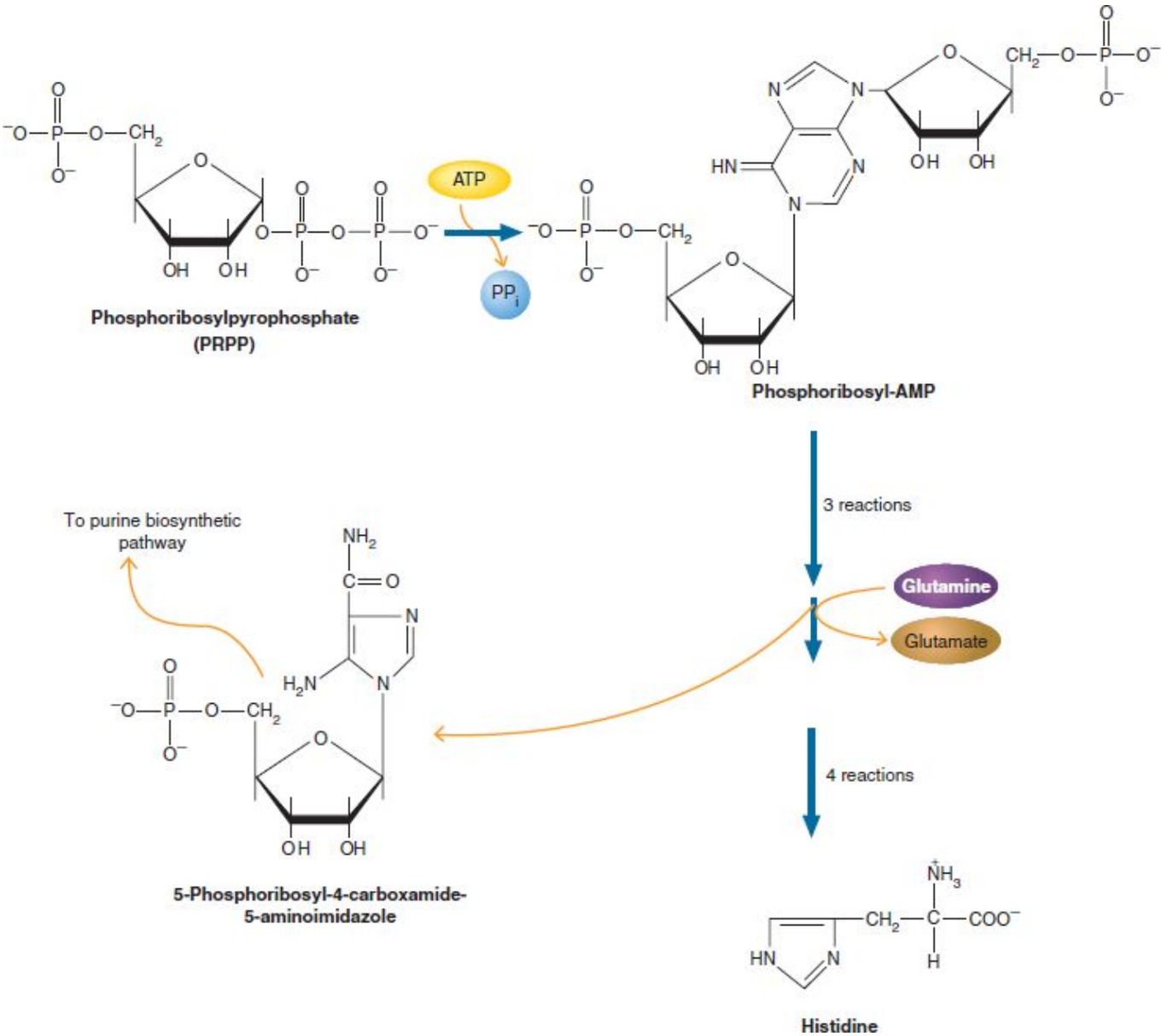


FIGURE 14.13

Histidine Biosynthesis

Histidine is derived from three biomolecules: phosphoribosylpyrophosphate (PRPP) (five carbons), the adenine ring from ATP (one nitrogen and one carbon), and glutamine (one nitrogen). The ATP used in the first reaction in the pathway is regenerated when 5-phosphoribosyl-4-carboxamide-5-aminoimidazole (released in a subsequent reaction) is diverted into the purine nucleotide biosynthetic pathway.

TABLE 14.2 One-Carbon Groups

Oxidation Level	Methanol (most reduced)	Formaldehyde	Formate (most oxidized)
One-carbon group	Methyl (—CH_3)	Methylene (—CH_2—)	Formyl (—CHO) Methenyl (—CH═)

FOLIC ACID Folic acid, also known as folate or folacin, is a B vitamin found in beans, peas, broccoli, beets, spinach, and sunflower seeds. Folic acid's structure consists of a pteridine nucleus and *para*-aminobenzoic acid, linked to one or more glutamic acid residues (**Figure 14.14**). Once absorbed by the body, folic acid is converted by dihydrofolate reductase to the biologically active

form, **tetrahydrofolate (THF)**. The carbon units carried by THF (i.e., methyl, methylene, methenyl, and formyl groups) are bound to N^5 of the pteridine ring and/or N^{10} of the para-aminobenzoate ring. **Figure 14.15** illustrates the interconversions of the one-carbon units carried by THF, as well as their origin and metabolic fate. A substantial number of one-carbon units enter the THF pool as N^5, N^{10}-methylene THF, produced during the conversion of serine to glycine and the cleavage of glycine (catalyzed by glycine synthase).

Methylcobalamin, a coenzyme form of vitamin B_{12}, is required for the N^5-methyl THF–dependent conversion of homocysteine to methionine (**Figure 14.15**). **Vitamin B_{12}** (cobalamin) is a complex, cobalt-containing molecule synthesized only by microorganisms (**Figure 14.16**). (During the purification of cobalamin, a cyanide group attaches to cobalt.) Another coenzyme form of vitamin B_{12}, 5′-deoxyadenosylcobalamin, is required for the isomerization of methylmalonyl-CoA to succinyl-CoA, which is catalyzed by methylmalonyl-CoA mutase (refer to **Figures 12.12** and **15.11**). Animals obtain cobalamin from intestinal flora and from consumption of foods derived from other animals (e.g., liver, eggs, shrimp, chicken, and pork).

Pernicious Anemia

A deficiency of vitamin B_{12} results in *pernicious anemia*. In addition to low red blood cell counts, the symptoms of this malady include weakness and various neurological disturbances. Pernicious anemia is most often caused by decreased secretion of intrinsic factor, a glycoprotein secreted by stomach cells, which is required for absorption of the vitamin in the intestine. Vitamin B_{12} absorption can also be inhibited by several gastrointestinal disorders, such as celiac disease and tropical sprue, both of which damage the lining of the intestine. A reduction in vitamin B_{12} absorption has also been observed in the presence of intestinal overgrowths of microorganisms induced by antibiotic treatments.

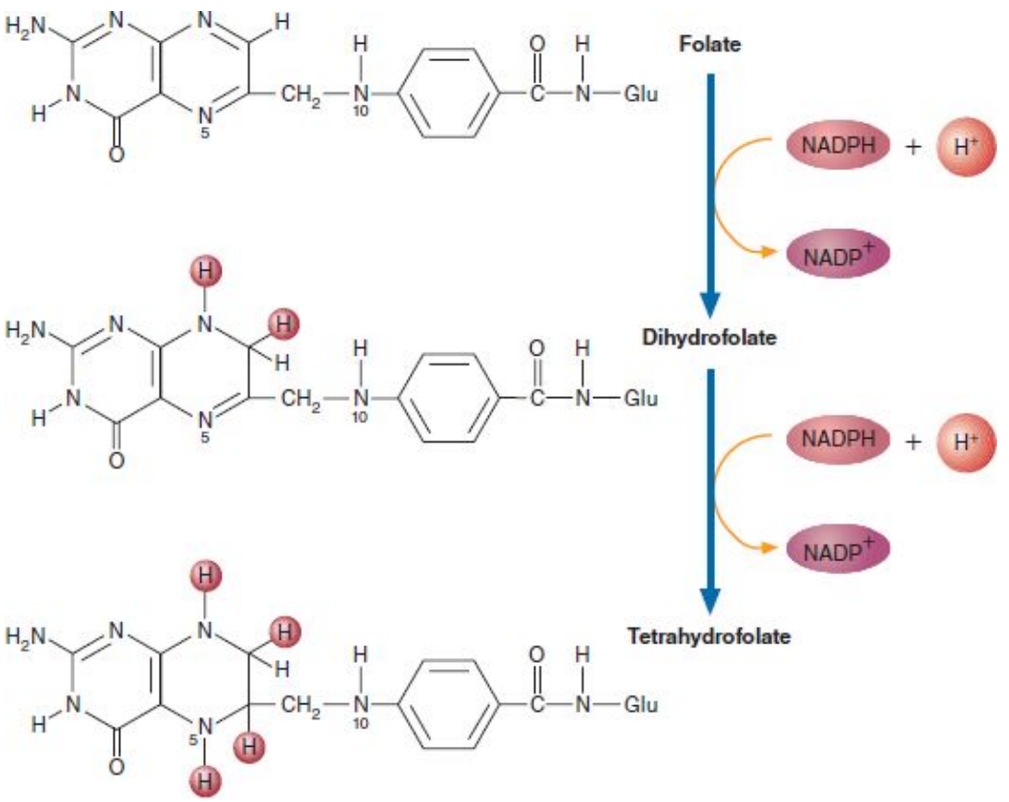


FIGURE 14.14

Biosynthesis of Tetrahydrofolate (THF)

The vitamin folic acid (folate) is converted to its biologically active form by two successive reductions of the pteridine ring. Both reactions are catalyzed by dihydrofolate reductase.

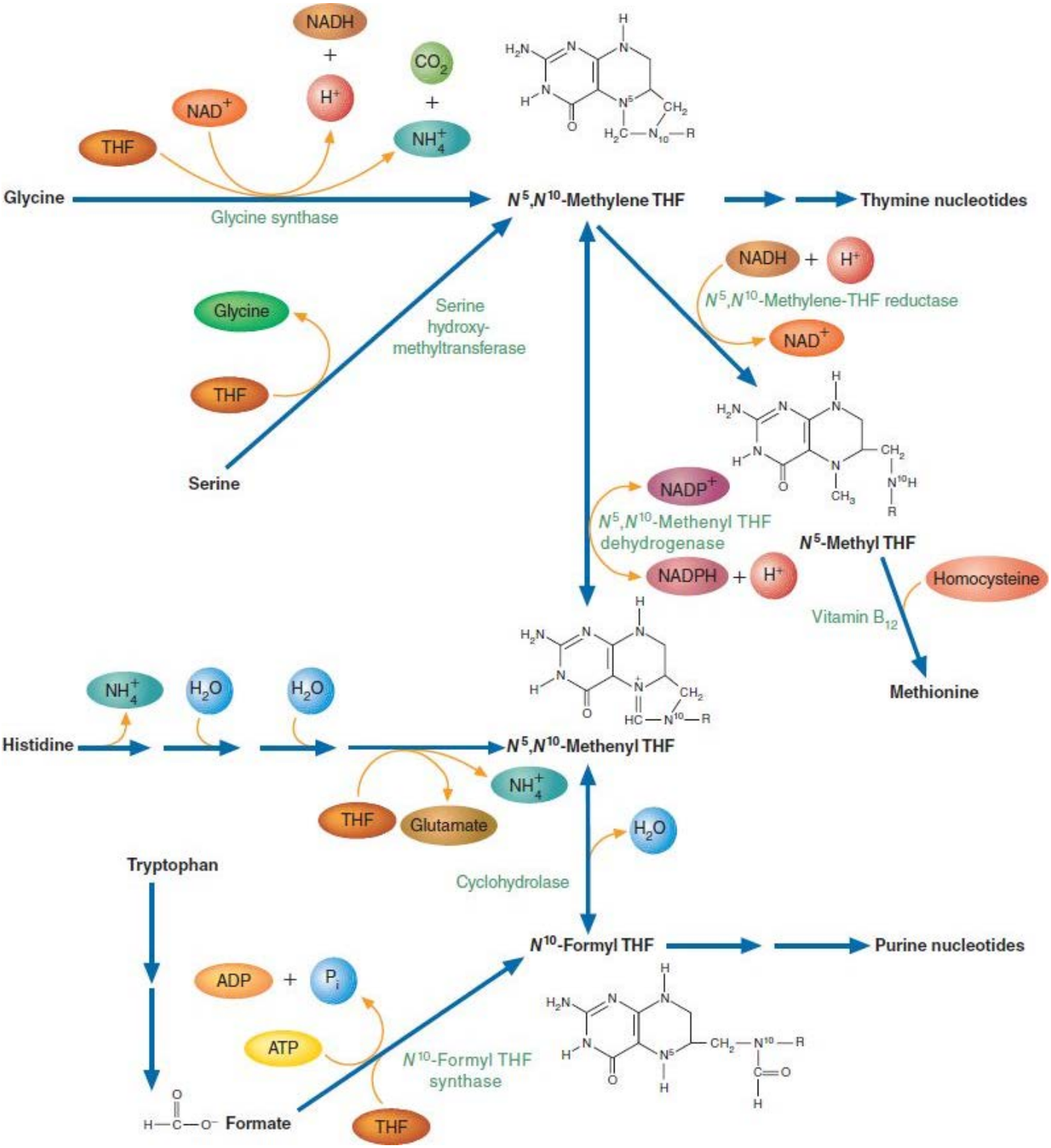


FIGURE 14.15

Structures and Enzymatic Interconversions of THF Coenzymes

The THF coenzymes play a critical role in one-carbon metabolism. The interconversions of the coenzymes are reversible except for the conversion of N^5,N^{10}-methylene THF to N^5-methyl THF.

S-ADENOSYLMETHIONINE ***S*-Adenosylmethionine (SAM)** is the major methyl group donor in one-carbon metabolism. Formed from methionine and ATP (**Figure 14.17**), SAM contains an "activated" methyl thioether group, which can be transferred to a variety of acceptor molecules (**Table 14.3**). *S*-Adenosylhomocysteine (SAH) is a product in these reactions. The loss of free energy that accompanies SAH formation makes the methyl transfer irreversible. SAM acts as a methyl donor in at least 115 transmethylation reactions, some of which occur in the synthesis of phospholipids,

several neurotransmitters, and glutathione. *DNA* methylation reactions play a significant role in gene expression. SAM also plays a role in the synthesis of the polyamines (p. 651) by donating aminopropyl groups. Polyamines are polycationic molecules that bind to negatively charged DNA, allowing the DNA to be compressed into chromosomes.

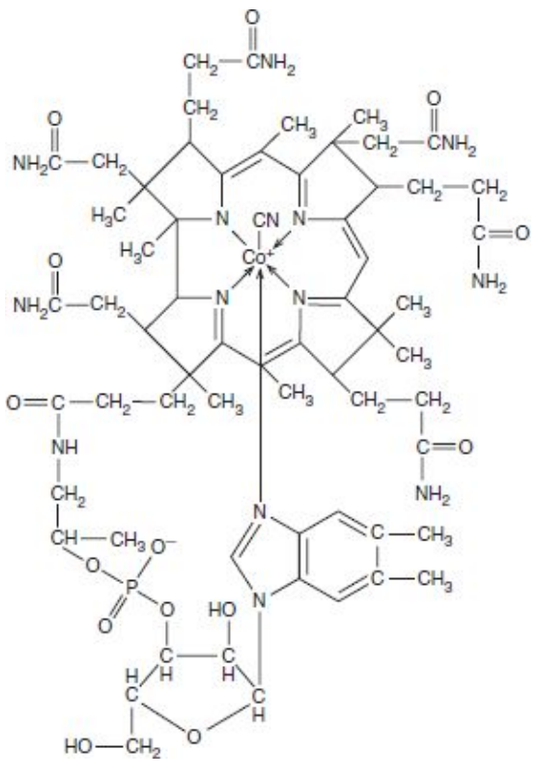


FIGURE 14.16

Structure of Cyanocobalamin, a Derivative of Vitamin B_{12}

During the purification of cobalamin, a cyanide group binds to cobalt.

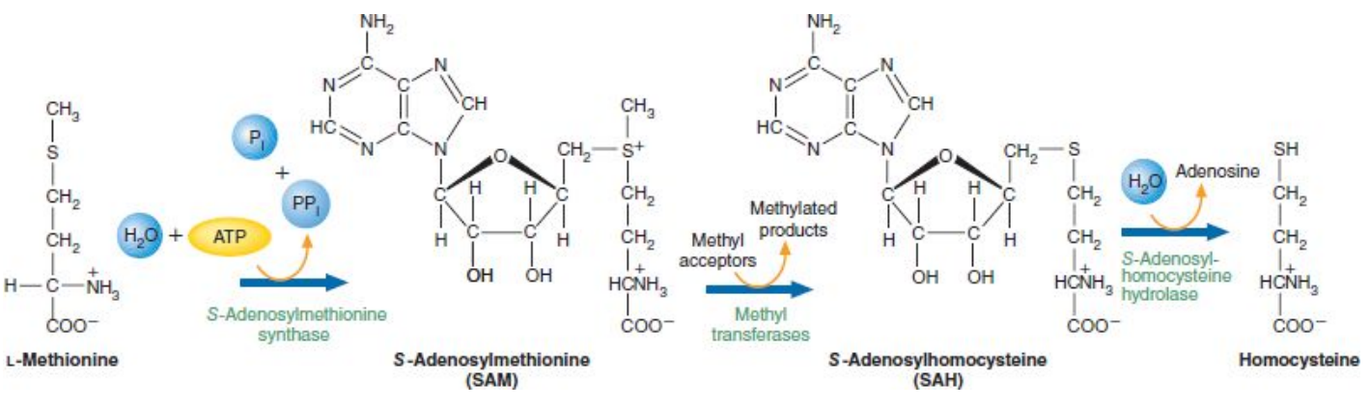


FIGURE 14.17

The Formation of *S*-Adenosylmethionine (SAM)

One of the principal functions of SAM is to serve as a methylating agent. The SAM product of these reactions is then hydrolyzed to form homocysteine.

TABLE 14.3 Examples of Transmethylation Acceptors and Products

Methyl Acceptors	Methylated Product
Phosphatidylethanolamine	Phosphatidylcholine (p. 410)
Norepinephrine*	Epinephrine* (p. 319)
Guanidinoacetate	Creatine (p. 549)
γ-Aminobutyric acid	Carnitine (p. 452)

* A catecholamine neurotransmitter molecule

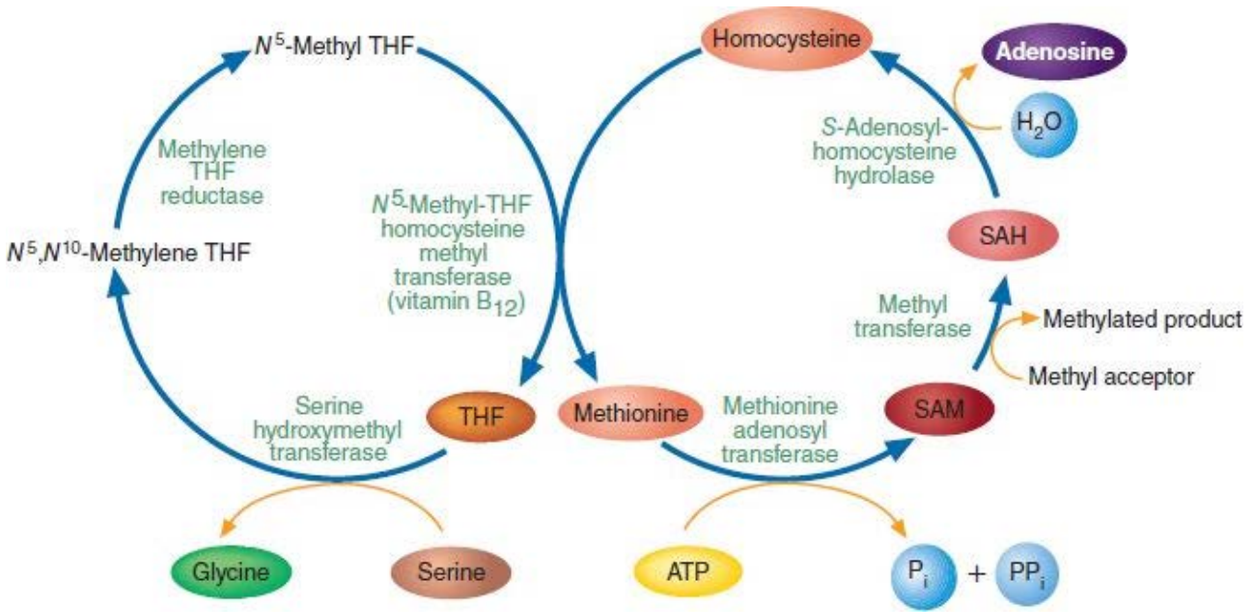


FIGURE 14.18

The Tetrahydrofolate (THF) and *S*-Adenosylmethionine (SAM) Pathways

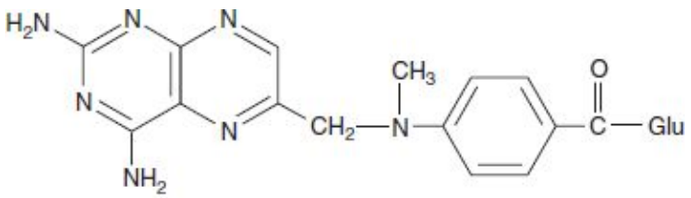
The THF and SAM pathways intersect at the reaction, catalyzed by N^5-methyl THF homocysteine methyltransferase, in which homocysteine is converted to methionine.

The importance of SAM in metabolism is reflected in the several mechanisms that provide for the synthesis of sufficient amounts of its precursor, methionine, when methionine is temporarily absent from the diet. For example, choline is used as a source of methyl groups to convert homocysteine into methionine. Homocysteine can also be methylated in a reaction utilizing N^5-methyl THF. This latter reaction is a bridge between the THF and SAM pathways (**Figure 14.18**).

Amethopterin and Disease

QUESTION 14.3

Amethopterin, also referred to as methotrexate, is a structural analogue of folate. Methotrexate has been used to treat several types of cancer and autoimmune diseases. It has been especially successful in childhood leukemia. Autoimmune diseases treated with methotrexate include rheumatoid arthritis, Crohn's disease, and psoriasis (excessive skin cell production and inflammation).



Amethopterin (methotrexate)

3D animation of amethopterin

Using your knowledge of cell biology and biochemistry, suggest a biochemical mechanism that explains why amethopterin is effective against cancer. [*Hints*: Compare the structures of folate and methotrexate. Review **Figure 14.14**.]

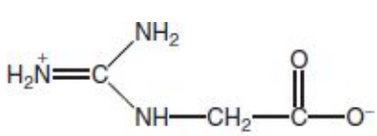
Circadian Rhythms

QUESTION 14.4

Melatonin is a hormone derived from serotonin. It is produced in the brain's light-sensitive pineal gland. The pineal's secretion of melatonin is depressed by nerve impulses that originate in the retina of the eye and other light-sensitive tissue in the body in response to light. Pineal function is involved in *circadian rhythms*, patterns of biochemical activity associated with light and dark, such as sleep/wake cycles. Melatonin is also a powerful antioxidant, especially in the central nervous system. After serotonin (5-hydroxytryptamine) is produced in the pineal gland, it is converted to 5-hydroxy-*N*-acetyltryptamine by *N*-acetyltransferase. 5-Hydroxy-*N*-acetyltryptamine is then methylated by *O*-methyltransferase. SAM is the methylating agent. With this information, draw the synthetic pathway of melatonin. Refer to Figure 5.4 for serotonin's structure.

QUESTION 14.5

Creatine is a nitrogen-containing organic acid found primarily in muscle and brain. Both of these tissues experience large and fluctuating energy demands. Phosphocreatine, the product of the reaction of creatine and ATP, serves as a short-term storage form of high-energy phosphate. (Refer to **Table 4.1**.) When energy demands are high and available ATP molecules are hydrolyzed, phosphocreatine donates its phosphoryl group to ADP to yield ATP (p. 379). Most creatine molecules are synthesized in the body in a two-reaction pathway. In the first step, arginine and glycine are converted in the kidney to guanidinoacetate and ornithine (p. 539) by L-arginine:glycine amidinotransferase (AGAT).



Guanidinoacetate

3D animation of guanidinoacetate

In liver, guanidinoacetate reacts with SAM to form creatine and SAH in a reaction catalyzed by *S*-adenosyl-L-methionine:*N*-guanidinoacetate methyltransferase (GAMT). With the information provided, write out the creatine biosynthetic pathway.

KEY CONCEPT

Tetrahydrofolate, the biologically active form of folic acid, and *S*-adenosylmethionine are important carriers of single-carbon atoms in a variety of biosynthetic reactions.

Glutathione

With concentrations that average about 5 mM in mammalian cells, glutathione (γ-glutamylcysteinylglycine; see **Table 5.3**) is the most abundant intracellular reducing agent. Glutathione (GSH) performs several critical functions. First, GSH is a major endogenous antioxidant molecule that quenches superoxide ($O_2^{\cdot-}$), hydroxyl radicals (·OH), and peroxynitrite anions ($ONOO^-$). GSH is a powerful antioxidant because the thiol group of its cysteine residue donates a reducing equivalent (H^+ and e^-) to unstable groups or radicals. GSH is also a cofactor for several antioxidant enzymes. Examples of enzymes that utilize GSH as a reducing agent include glutathione peroxidase (p. 389) and dehydroascorbate reductase, the enzyme that maintains vitamin C (p. 393) in its reduced form. GSH also protects cells from oxidative stress by maintaining the sulfhydryl groups of enzymes and other molecules in a reduced state. Second, GSH has roles in diverse biochemical processes. Examples include DNA synthesis and repair, protein synthesis, and leukotriene synthesis (LTC_4, **Figure 11.4c**). Third, GSH protects cells from xenobiotics. GSH conjugates of these molecules are formed either spontaneously or in reactions catalyzed by the GSH-S-transferases.

KEY CONCEPTS

- Glutathione (GSH), the most common intercellular thiol, is involved in many cellular activities.
- In addition to reducing sulfhydryl groups, GSH protects cells against toxins and promotes the transport of amino acids such as cysteine.

GSH is synthesized in a pathway composed of two reactions. In the first reaction, γ-glutamylcysteine synthase catalyzes the condensation of glutamate with cysteine (**Figure 14.19**). γ-Glutamylcysteine, the product of this reaction, then combines with glycine to form GSH in a reaction catalyzed by glutathione synthase.

An additional enzyme, γ-glutamyl transpeptidase, also called γ-glutamyl transferase, is synthesized in certain tissues, where it cleaves extracellular GSH and other molecules to increase the availability of amino acids, especially cysteine, a rate-limiting substrate for intracellular GSH synthesis. In this process, called the γ-glutamyl cycle (**Figure 14.19**), the γ-glutamyl transpeptidase catalyzes the reaction between an extracellular amino acid and GSH to yield an intracellular γ-glutamylamino acid and glycine and cysteine.

The γ-glutamylamino acid is then hydrolyzed by γ-glutamylcyclotransferase to yield 5-oxoproline and the amino acid. 5-Oxoproline is then converted to glutamate by 5-oxoprolinase (**Figure 14.20**).

GSH synthesis is regulated, in part, by cellular cysteine levels. Because homocysteine is a precursor in cysteine synthesis (**Figure 14.11**), GSH synthesis is linked to both the SAM (**Figure 14.17**) and transsulfuration (p. 586) pathways. (The transulfuration pathway, described in Chapter 15, is a biochemical pathway that controls cell concentrations of sulfur-containing biomolecules.)

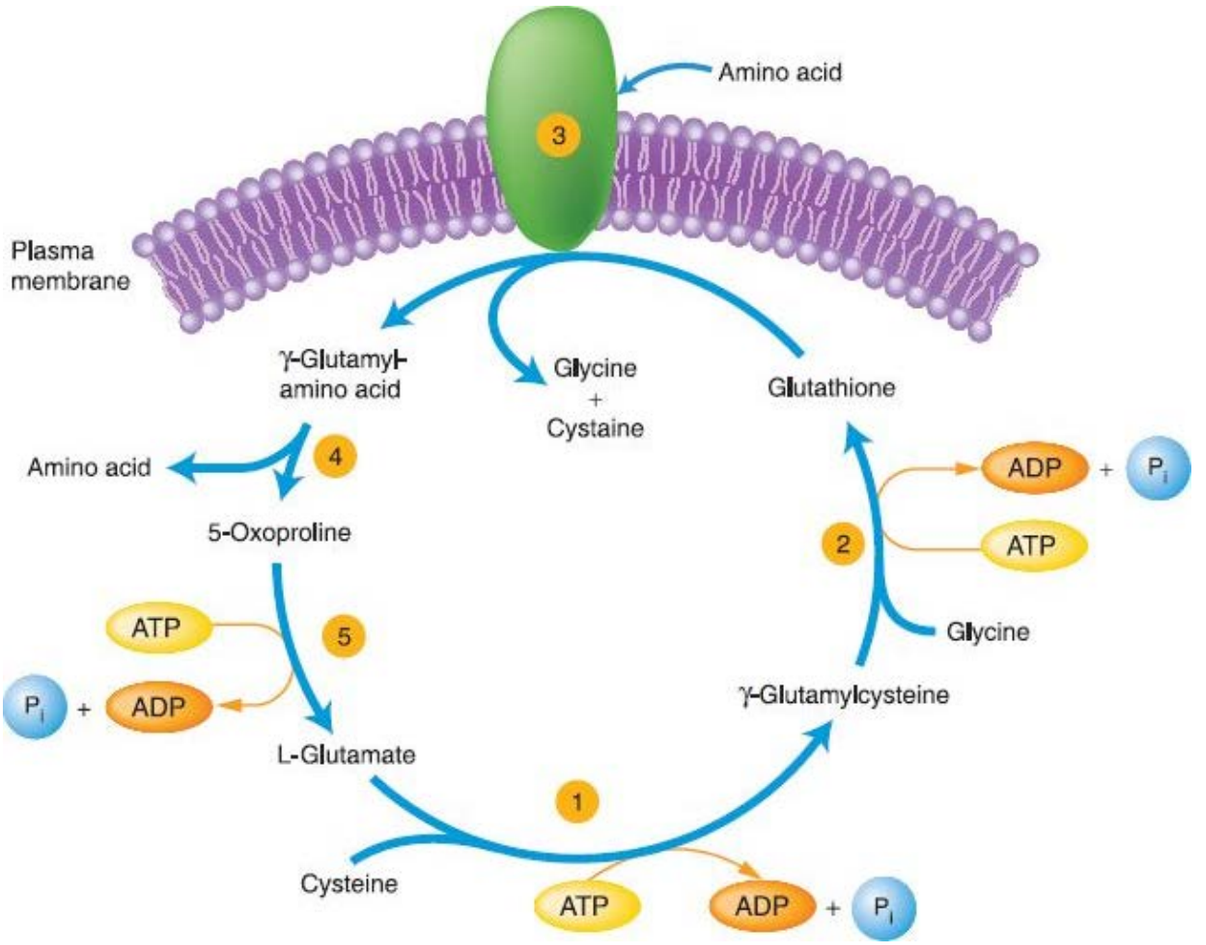


FIGURE 14.19

Glutathione Biosynthesis and the γ-Glutamyl Cycle

GSH is synthesized from glutamate, cysteine, and glycine in two ATP-requiring reactions. *γ*-Glutamylcysteine synthetase (1) converts glutamate and cysteine into *γ*-glutamylcysteine. Glutathione synthetase (2) then catalyzes the formation of a peptide bond that links *γ*-glutamylcysteine to glycine, yielding glutathione. In some tissues (e.g., intestinal and kidney cells), amino acids are transported across the plasma membrane because an additional enzyme, *γ*-glutamyl transpeptidase (3), is produced that catalyzes a reaction between GSH and an amino acid to yield a *γ*-glutamylamino acid and cysteinylglycine. Cysteinylglycine is subsequently hydrolyzed to cysteine and glycine. The *γ*-glutamylamino acid is hydrolyzed by *γ*-glutamyl cyclotransferase (4) to yield 5-oxoproline and the amino acid. Glutamate is regenerated from 5-oxoproline by 5-oxoprolinase (5), an ATP-requiring enzyme.

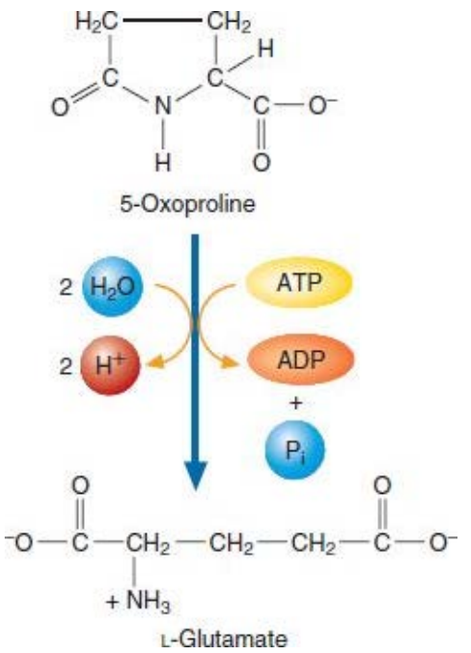


FIGURE 14.20

Conversion of 5-Oxyproline to L-Glutamate

The ATP-driven hydrolysis of 5-oxoproline to yield L-glutamate is catalyzed by 5-oxoprolinase.

Neurotransmitters

More than 30 different substances, including several amino acids, function as neurotransmitters. **Neurotransmitters**, signal molecules released from neurons, are either excitatory or inhibitory. *Excitatory neurotransmitters* (e.g., glutamate and acetylcholine) open sodium channels and promote the depolarization of the membrane in another cell (either another neuron or an effector cell, such as a muscle cell). If the second (postsynaptic) cell is a neuron, the wave of depolarization (referred to as an action potential) triggers the release of neurotransmitter molecules as it reaches the end of the axon. (Most neurotransmitter molecules are stored in numerous membrane-enclosed *synaptic vesicles*.) When the action potential reaches the nerve ending, the neurotransmitter molecules are released by exocytosis into the synapse. If the postsynaptic cell is a muscle cell, sufficient release of the excitatory neurotransmitter acetylcholine results in muscle contraction. *Inhibitory neurotransmitters* (e.g., glycine) open chloride channels and make the membrane potential in the postsynaptic cell even more negative; that is, they inhibit the formation of an action potential.

A significant percentage of neurotransmitter molecules are either amino acids or amino acid derivatives (**Table 14.4**). The latter class, referred to as the **biogenic amines**, includes γ-aminobutyric acid (GABA), the catecholamines (derivatives of tyrosine), serotonin, and histamine. The synthesis and inactivation of the inhibitory neurotransmitters γ-aminobutyric acid (GABA, derived from glutamate) and serotonin (derived from tryptophan) are described in the online reading The Amine Neurotransmitters. Catecholamine biosynthesis is described in an online reading. The link between the catecholamine dopamine and Parkinson's disease is described in the online reading Parkinson's Disease and Dopamine.

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essays on the amine neurotransmitters, the catecholamines, and dopamine and Parkinson's disease.

TABLE 14.4 Amino Acid and Amine Neurotransmitters

Amino Acids	Amines
Glycine	Norepinephrine*
Glutamate	Epinephrine*
γ-Aminobutyric acid (GABA)	Dopamine*
	Serotonin
	Histamine

* A catecholamine.

Biochemistry IN PERSPECTIVE

Gasotransmitters

How do gas molecules, previously thought to be toxic at any concentration, act as signal molecules in the mammalian body? **Gasotransmitters** are endogenous gaseous molecules that have been recently recognized as a class of signal molecules. These molecules, which are synthesized by highly regulated enzymes, have several characteristics in common: (1) they are lipid-soluble gases that can diffuse through cellular membranes; (2) their biological effects at physiological concentrations (usually μmolar) are mediated by second messenger molecules and/or ion channels; and (3) signals are not terminated by inactivating enzymes but largely by discontinued synthesis and diffusion away from cellular targets. Gasotransmitters include nitric oxide (NO•), carbon monoxide (CO), hydrogen sulfide (H_2S), and certain reactive oxygen species (ROS). Each of these substances is extremely toxic and/or potentially lethal at high concentrations.

Nitric Oxide

NO• is a highly reactive gas. Because of its free radical structure, until recently NO• was regarded only as a contributing factor in the destruction of the ozone layer in the Earth's atmosphere and as a precursor of acid rain. NO• is an important signal molecule, produced throughout the mammalian body. Physiological functions such as the regulation of blood pressure, the inhibition of blood clotting, and the destruction of foreign, damaged, or cancerous cells by macrophages, are triggered when NO• binds to the heme group of guanylate cyclase. The product of guanylate cyclase is the second messenger molecule cGMP (cyclic GMP; see p. 610). NO• is also a neurotransmitter and is produced in many areas of the brain where its formation has been linked to the excitatory neurotransmitter glutamate. When glutamate is released from a neuron and binds to a certain class of glutamate receptor, a transient flow of Ca^{2+} through the postsynaptic membrane stimulates NO• synthesis.

Once synthesized, NO• diffuses back to the presynaptic cell, where it signals further release of glutamate. Thus, NO• acts as a *retrograde neurotransmitter*; that is, it promotes a cycle in which glutamate is released from the presynaptic neuron and then binds to and promotes action potentials in the postsynaptic neuron. This potentiating mechanism is now believed to play a role in learning and

memory formation, as well as other functions in mammalian brain. The disruption of the normally precise regulation of NO• synthesis has been linked to numerous pathological conditions, including stroke, migraine headache, male impotence, septic shock, and inflammatory conditions such as multiple sclerosis and insulin-dependent diabetes (p. 615).

NO• is synthesized by NO synthase (NOS), a heme-containing metalloenzyme that catalyzes a two-step oxidation of L-arginine to L-citrulline (**Figure 14A**). In this complex reaction, electrons are transferred from NADPH to O_2 by an electron transport chain that involves several redox components. The functional enzyme is a homodimer (**Figure 14B**). Each monomer has two major domains. The reductase domain possesses binding sites for NADPH, FAD, and FMN. The oxygenase domain, binds BH_4 (tetrahydrobiopterin, a folic acid–like molecule; see Figure 15.9 on p. 584) and the substrates arginine and O_2. Between the two major domains is the binding site for calmodulin (CAM), a small calcium-binding protein that regulates a variety of enzymes. During NO• synthesis, CAM accelerates the rate of electron transfer from the reductase domain to the heme group.

The biosynthesis of NO• begins with the hydroxylation of L-arginine. NADPH donates two electrons to FAD, which in turn reduces FMN. BH_4 is essential for activation of O_2 by the electrons donated by NADPH. The product of this reaction, L-hydroxyarginine, remains bound to NOS. The steps in the subsequent reaction have not yet been resolved. It is believed that L-hydroxyarginine reacts with a heme-peroxy complex (R—O—OH) to give citrulline and NO•.

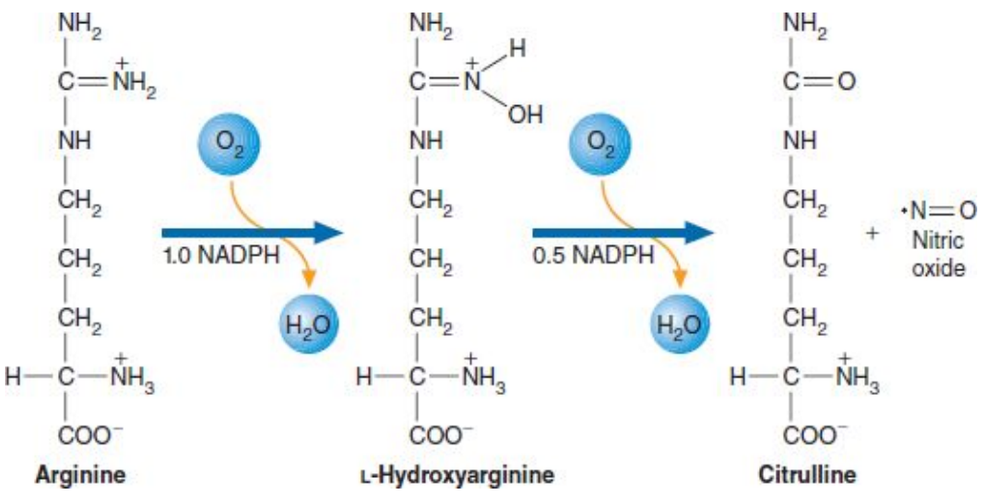


FIGURE 14A

The NOS-Catalyzed Reaction

The biosynthesis of NO• is a two-step process in which two successive monooxygenation reactions convert arginine into citrulline. During the reaction, 2 mol of O_2 and 1.5 mol of NADPH are consumed per mole of citrulline formed.

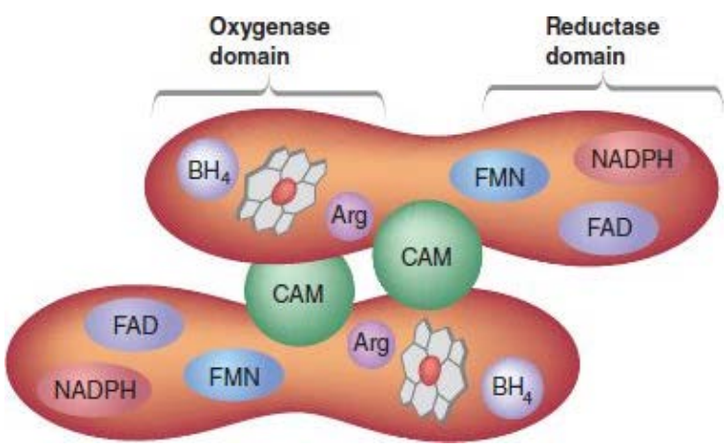


FIGURE 14B

Diagrammatic Structure of NOS

The catalytically active NOS is a homodimer. Each monomer, which has an oxygenase domain and a reductase domain, binds NADPH, FAD, FMN, BH_4, and CAM in addition to the substrates arginine and O_2. The reductase domain is homologous to cytochrome P_{450} reductase (pp. 485–86).

Carbon Monoxide

The most important physiological roles played by CO, a colorless and odorless gas, include neuromodulation (i.e., regulation of neurotransmitter release and other neural activities related to learning, memory, and thermal regulation), cardiac protection during hypoxia against ischemia–reperfusion damage (see the Biochemistry in Perspective essay in Chapter 10 entitled Myocardial Infarct: Ischemia and Reperfusion, p. 395), and vascular relaxation (i.e., promotion of smooth muscle relaxation and, therefore, blood vessel dilation). The reaction in which CO synthesis is catalyzed by heme oxygenase (HO), an ER enzyme, is described in the online Chapter 15 Biochemistry in Perspective essay on heme biotransformation. In addition to CO, the products of this reaction are biliverdin and free Fe(II) ions.

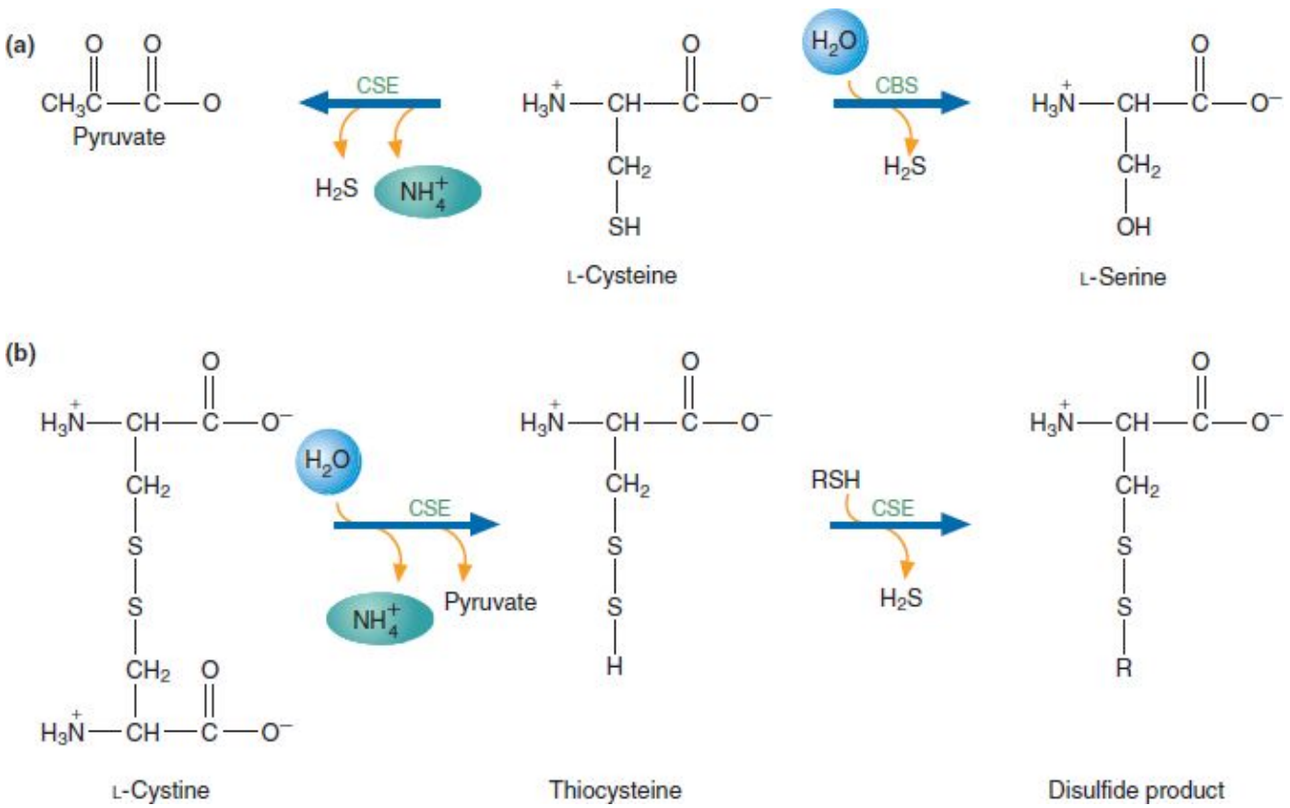


FIGURE 14C

H_2S Biosynthesis

(a) Most H_2S-producing reactions are catalyzed by either CBS or CSE with cysteine as the substrate. CBS hydrolyzes cysteine to form serine and H_2S, and CSE catalyzes the conversion of cysteine to pyruvate, NH_4^+, and H_2S. (b) CSE also catalyzes a two-reaction sequence in which cystine is converted to thiocysteine, pyruvate, and NH_4^+. Thiocysteine then reacts with a thiol (e.g., GSH or another cysteine) to yield H_2S and a cysteine disulfide product.

The catabolism of heme is an ongoing activity in all tissues but occurs most notably in spleen

(hemoglobin breakdown) and liver (cytochrome P_{450} turnover). Heme itself and its oxidized form, hemin, are strong oxidants and have the potential to confer significant tissue damage. Tobacco smoking, a major source of CO in blood, has been linked to the wet form of macular degeneration, a leading cause of blindness. Biliverdin and its reduced form, bilirubin, have much lower oxidant potential than the parent heme. All cells must have a pool of available free heme to serve as a resource for synthesis of heme-containing proteins, but that level is kept in a safe range by the action of HO. The free iron is sequestered by ferritin to prevent redox damage. Many actions of CO are mediated by its binding to heme-containing proteins or activation of K^+ channels. CO, at physiologically safe levels, has also been shown to reduce inflammation and suppress apoptosis (cell death).

Hydrogen Sulfide

H_2S is a toxic gas with an unpleasant, foul odor. Within the body, where it is produced in small quantities, H_2S is known to affect numerous physiological processes. The most prominent of these are its functions as a neuromodulator in brain and as a vascular relaxant. H_2S is synthesized from cysteine by several enzymes (**Figure 14C**). The most important of these are the pyridoxal phosphate–requiring enzymes cystathione β-synthetase (CBS) and γ-cystathionase (CSE), both of which catalyze other reactions in the transulfuration pathway (see later: Section 15.2). With cystine used as an alternate substrate, CSE catalyzes a two- reaction pathway that also yields H_2S. CSE is the dominant H_2S–producing enzyme in the liver and the cardiovascular system. CBS appears to be dominant in the brain. H_2S dilates arterioles by triggering the opening of ATP-sensitive K^+ channels, a process that results in the hyperpolarization of smooth muscle cell membrane.

H_2S exerts its neuromodulatory effects by enhancing the activation of receptors for NMDA (*N*-methyl-D-aspartate). *NMDA receptors* are a type of glutamate receptor; when activated, they open an ion channel that allows the inward flow of Ca^{2+} and other ions. NMDA receptors are believed to play an important role in *synaptic plasticity*, the capacity of a synapse (the connection between two neurons) to change in chemical strength. Changes in synaptic strength are believed to be the basis for memory formation.

SUMMARY At very low concentrations, NO, CO, and H_2S are signal molecules that diffuse easily through cell membranes and whose synthesis is rigorously controlled.

Nucleotides

Nucleotides are nitrogen-containing molecules required for cell growth and development. Not only are nucleotides the building blocks of the nucleic acids, they also play several essential roles in energy transformation and regulate many metabolic pathways. As described, each nucleotide is composed of three parts: a nitrogenous base, a pentose sugar, and one or more phosphate groups. The nitrogenous bases are derivatives of either purine or pyrimidine (**Figure 14.21**), which are planar heterocyclic aromatic molecules.

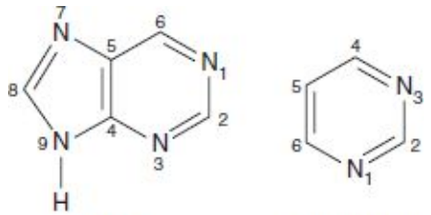


FIGURE 14.21

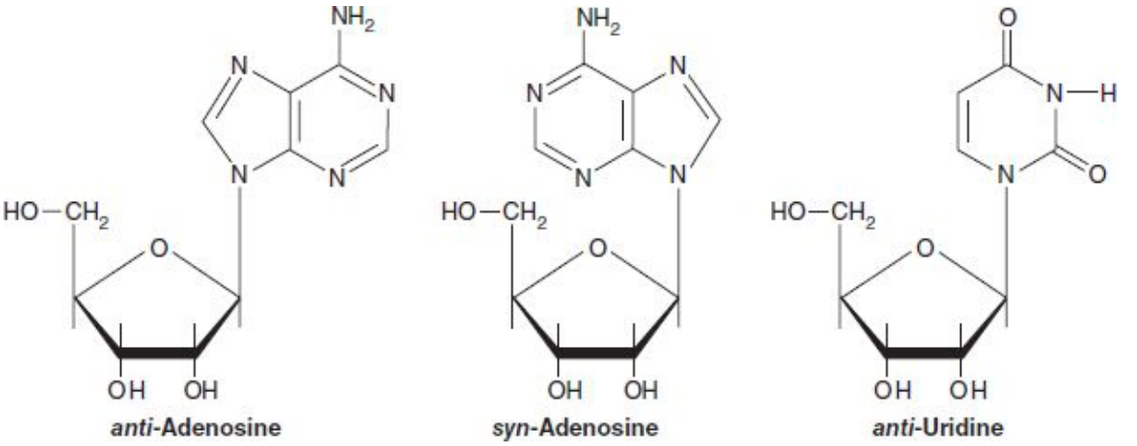
Structures of Purine (a) and Pyrimidine (b)

3D animation of purine

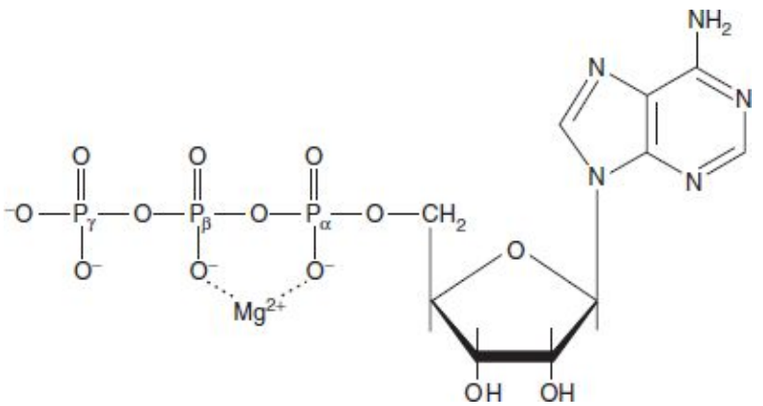
3D animation of pyrimidine

Common, naturally occurring **purines** include adenine, guanine, xanthine, and hypoxanthine; thymine, cytosine, and uracil are common **pyrimidines** (**Figure 14.22**). Because of their aromatic structures, the purines and pyrimidines absorb UV light. At pH 7, this absorption is especially strong at 260 nm. Purine and pyrimidine bases have tautomeric forms; that is, they undergo spontaneous shifts in the relative position of a hydrogen atom and a double bond in a three-atom sequence involving heteroatoms. This property is especially important because the precise location of hydrogen atoms on the oxygen and nitrogen atoms affects the interaction of bases in nucleic acid molecules. Adenine and cytosine have both amino and imino forms; guanine, thymine, and uracil have both keto (lactam) and enol (lactim) forms. At physiological pH, the amino and keto forms are the most stable. The amino and imino forms of adenine and the keto and enol forms of thymine are illustrated in **Figure 14.23**.

When a purine or pyrimidine base is linked through a β-N-glycosidic linkage to C-1 of a pentose sugar, the molecule is called a **nucleoside** (**Figure 14.24**). Nucleosides contain one of two types of sugar: ribose or deoxyribose. Ribose-containing nucleosides with adenine, guanine, cytosine, and uracil are referred to as adenosine, guanosine, cytidine, and uridine, respectively. When the sugar component is deoxyribose, the prefix *deoxy* is used. For example, the deoxy nucleoside with adenine is called deoxyadenosine. Deoxythymidine is called thymidine because the base thymine usually occurs only in deoxyribonucleosides. Possible confusion in the identification of atoms in the base and sugar components of nucleosides is avoided by using a superscript prime to denote the atoms in the sugar (**Figure 14.24**). Rotation around the N-glycosidic bond of nucleosides creates two conformations: *syn* and *anti*. Purine nucleosides occur as either *syn* or *anti* forms. In pyrimidine nucleosides, the *anti* conformation predominates because of steric hindrance between the pentose sugar and the carbonyl oxygen at C-2.



Nucleotides are nucleosides in which one or more phosphate groups are bound to the sugar (**Figure 14.25**). Most naturally occurring nucleotides are 5′-phosphate esters. If one phosphate group is attached at the 5′-carbon of the sugar, the molecule is named as a nucleoside monophosphate (e.g., adenosine-5′-monophosphate: AMP). Nucleoside di- and triphosphates contain two and three phosphate groups, respectively. Phosphate groups make nucleotides strongly acidic. (Protons dissociate from the phosphate groups at physiological pH.) Because of their acidic nature, nucleotides may also be named as acids. For example, AMP is often referred to as adenylic acid or adenylate. Nucleoside di- and triphosphates form complexes with Mg^{2+}. In nucleoside triphosphates such as ATP, Mg^{2+} can form α,β (shown) and β,γ complexes:



Purine and pyrimidine nucleotides can be synthesized in de novo and salvage pathways. Both types of pathway are described.

Visit the companion website at www.oup.com/us/mckee to read the related Biochemistry in Perspective essay The Nucleotides: IMP Biosynthesis.

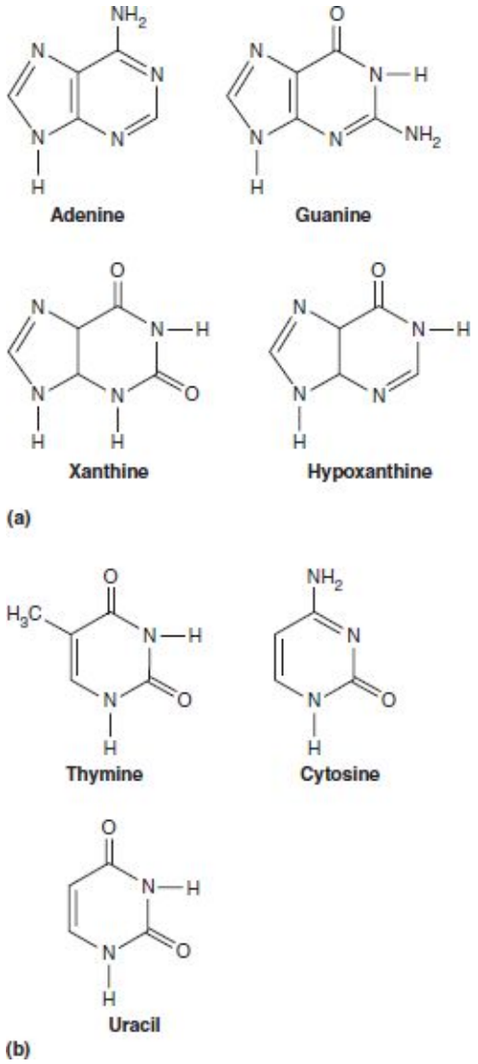


FIGURE 14.22

Structures of The Most Common Naturally Occurring Purine (a) and Pyrimidine (b) Bases

PURINE NUCLEOTIDE BIOSYNTHESIS The de novo synthesis of purine nucleotides begins with the formation of 5-phospho-α-D-ribosyl-1-pyrophosphate (PRPP) (**Figure 14.26**) catalyzed by ribose-5-phosphate pyrophosphokinase (PRPP synthetase). (The substrate for this reaction, α-D-ribose-5-phosphate, is a product of the pentose phosphate pathway.) The pathway by which PRPP is converted to inosine monophosphate (inosinate), the first purine nucleotide, is described in the online reading The Nucleotides: IMP Biosynthesis. The origin of the ring atoms of purines is illustrated in **Figure 14.27**.

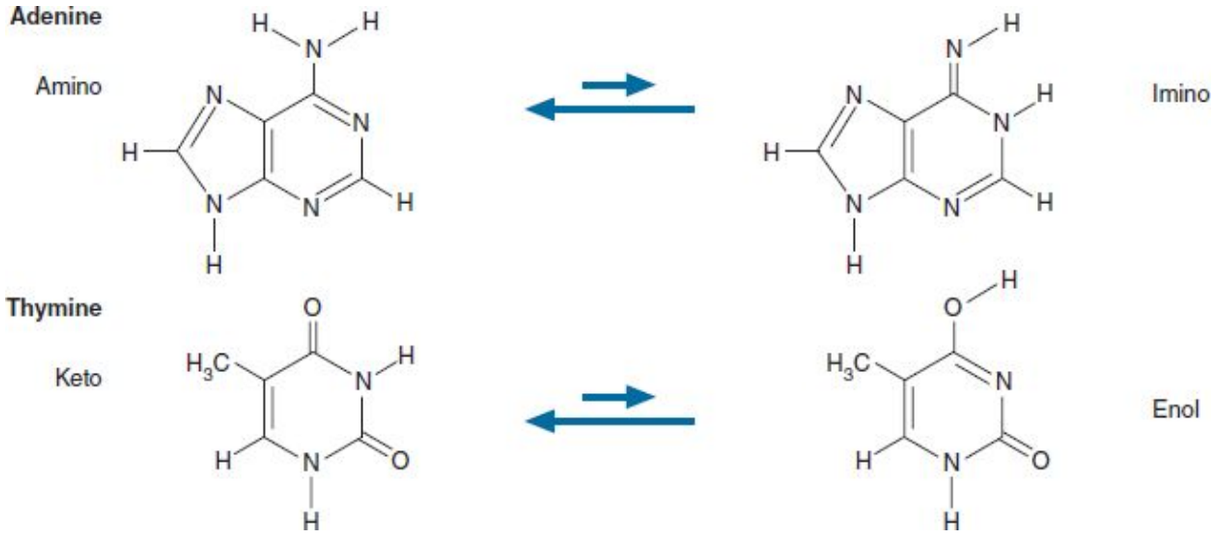


FIGURE 14.23

Tautomers of Adenine and Thymine

At physiological pH, the amino and keto tautomers of nitrogenous bases are the predominant forms. The hydrogen-binding properties of nucleotide base tautomers have important consequences in DNA replication (p. 642).

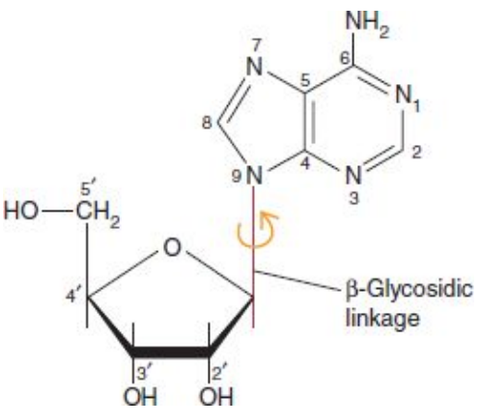


FIGURE 14.24

Nucleoside Structure

Nucleosides are molecules in which nitrogenous bases such as adenine (shown) are linked via a *β*-glycosidic bond to C-1 of a pentose sugar, in this case ribose.

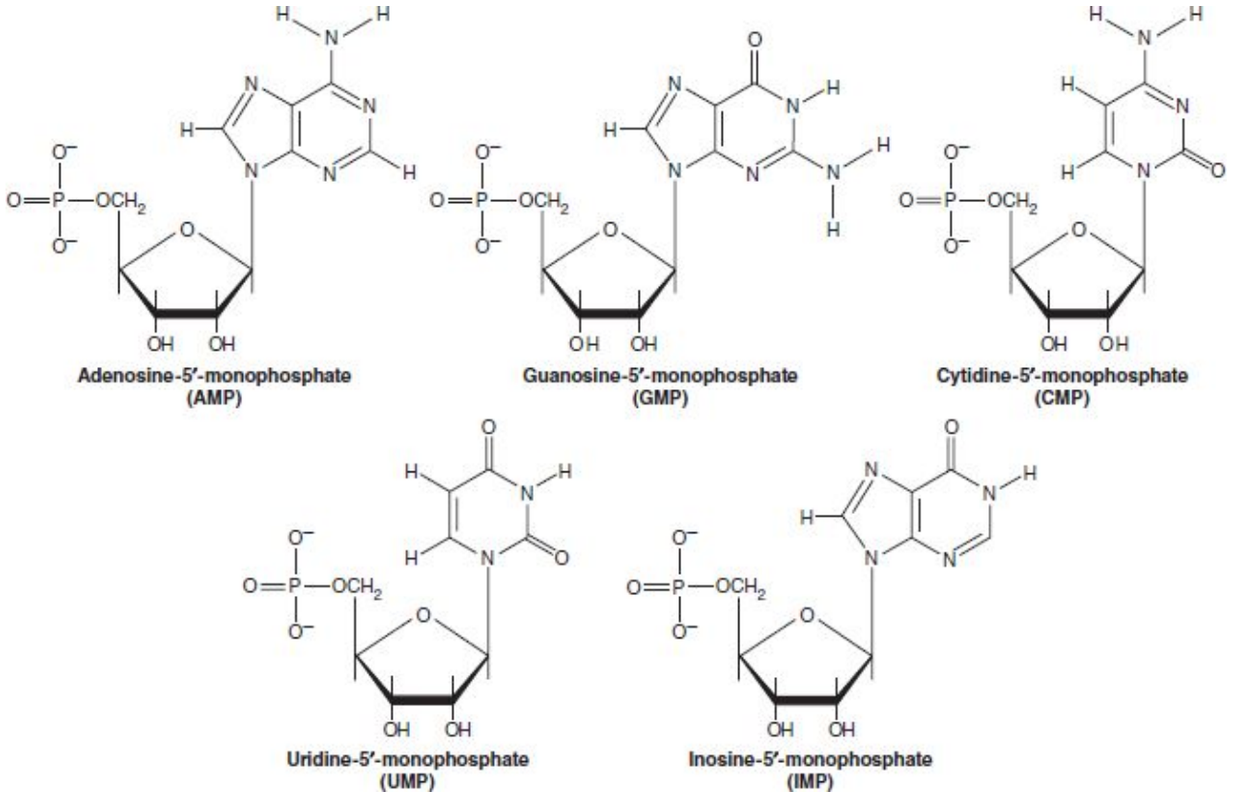


FIGURE 14.25

Common Ribonucleotides

Ribonucleotides contain ribose. When nucleotides contain deoxyribose instead of ribose, their names include the prefix *deoxy*. Inosine-5′-monophosphate (IMP) is an intermediate in purine nucleotide synthesis. The base component of IMP is hypoxanthine.

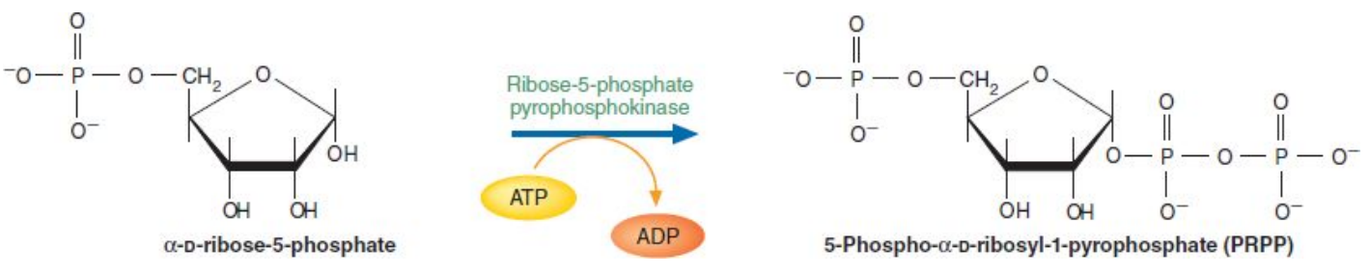


FIGURE 14.26

PRPP Synthesis

The de novo purine nucleotide pathway begins with synthesis of PRPP (5-phospho-α-D-ribosyl-1-pyrophosphate), which is catalyzed by ribose-5-phosphate pyrophosphatase.

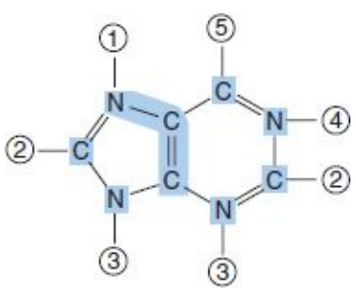


FIGURE 14.27

Origin of Purine Ring Atoms

The sources of the nitrogen and carbon atoms in purine rings, determined by isotopic labeling experiments, are as

follows: 1 = glycine; 2 = formate; 3 = glutamine's amide nitrogen; 4 = aspartate; 5 = CO_2.

The conversion of inosine-5′-monophosphate (IMP) to either AMP (or adenylate) or guanosine monophosphate (GMP or guanylate) requires two reactions (**Figure 14.28**). AMP differs from IMP in only one respect: an amino group replaces a keto oxygen at position 6 of the purine base. The amino nitrogen provided by aspartate becomes linked to IMP in a GTP-requiring reaction catalyzed by adenylosuccinate synthetase. In this step, the product adenylosuccinate eliminates fumarate to form AMP. (The enzyme that catalyzes this reaction also catalyzes a similar step in IMP synthesis.) The conversion of IMP to GMP begins with a dehydrogenation utilizing NAD^+, which is catalyzed by IMP dehydrogenase. The product, referred to as xanthosine monophosphate (XMP), is then converted to GMP by the donation of an amino nitrogen by glutamine in an ATP-requiring reaction catalyzed by GMP synthase.

Nucleoside triphosphates are the most common nucleotides used in metabolism. They are formed in the following manner. Recall that ATP is synthesized from ADP and P_i during certain reactions in glycolysis and aerobic metabolism. ADP is synthesized from AMP and ATP in a reaction catalyzed by adenylate kinase:

$$\text{AMP} + \text{ATP} \rightleftharpoons 2\ \text{ADP} \tag{7}$$

Other nucleoside triphosphates are synthesized in ATP-requiring reactions catalyzed by a series of nucleoside monophosphate kinases:

$$\text{NMP} + \text{ATP} \rightleftharpoons \text{NDP} + \text{ADP} \tag{8}$$

Nucleoside diphosphate kinase catalyzes the formation of nucleoside triphosphates,

$$\text{N}_1\text{DP} + \text{N}_2\text{TP} \rightleftharpoons \text{N}_1\text{TP} + \text{N}_2\text{DP} \tag{9}$$

where N_1 and N_2 are purine or pyrimidine bases.

In the purine salvage pathway, purine bases obtained from the normal turnover of cellular nucleic acids or (to a lesser extent) from the diet are reconverted into nucleotides. Because the de novo synthesis of nucleotides is metabolically expensive (i.e., relatively large amounts of phosphoryl bond energy are used), many cells have mechanisms to retrieve purine bases. Hypoxanthine-guaninephosphoribosyltransferase (HGPRT) catalyzes nucleotide synthesis using PRPP and either hypoxanthine or guanine. The hydrolysis of pyrophosphate makes these reactions irreversible.

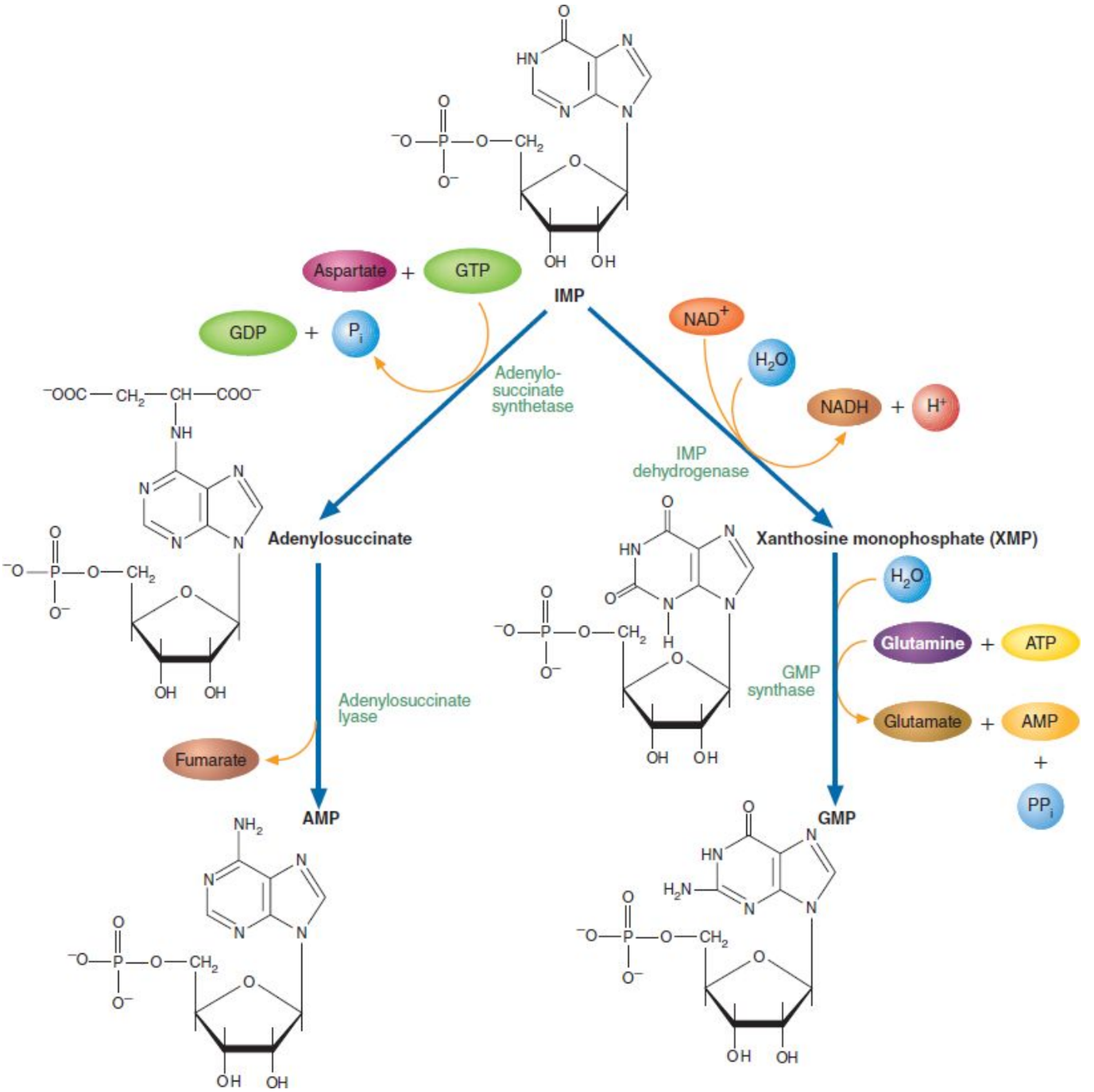


FIGURE 14.28

Biosynthesis of AMP and GMP from IMP

In the first step of AMP synthesis, the amino group of aspartate replaces the C-6 keto oxygen of the hypoxanthine base moiety of IMP. In the second step, the product of the first reaction, adenylosuccinate, is hydrolyzed to form AMP and fumarate. GMP synthesis begins with the oxidation of IMP to form XMP. GMP is produced as the amide nitrogen of glutamine replaces the C-2 keto oxygen of XMP. Note that AMP formation requires GTP and that GMP formation requires ATP.

(10)

(11)

Deficiency of HGPRT causes *Lesch–Nyhan syndrome*, a devastating X-linked disease that occurs primarily in males. It is characterized by excessive production of uric acid, the degradation product of purine nucleotides (Section 15.3), and certain neurological symptoms (self-mutilation, involuntary movements, and mental retardation). Although a powerful antioxidant, uric acid can act as a pro-oxidant when present in large amounts. Oxidative stress is now believed to contribute to the symptoms of Lesch–Nyhan syndrome. Affected children appear normal at birth but begin to deteriorate at about 3 to 4 months of age. Death, usually caused by renal failure, occurs in childhood. A partially defective HGPRT enzyme causes one form of *gout* (a condition in which high blood uric acid concentrations result in the accumulation of sodium urate crystals in joints, especially those in feet).

Lesch-Nyhan Syndrome

Visit the companion website at www.oup.com/us/mckee to read the Biochemistry in Perspective essay on gout in Chapter 15.

Adenine phosphoribosyltransferase (ARPT) catalyzes the transfer of adenine to PRPP, thus forming AMP:

(12)

The relative importance of the de novo and salvage pathways is unclear. However, the severe symptoms of hereditary HGPRT deficiency indicate that the purine salvage pathway is vitally important. In addition, the purine salvage pathway has a critical function in terminally differentiated neurons in the adult brain. Investigations of purine nucleotide synthesis inhibitors for treating cancer indicate that both pathways must be inhibited for significant tumor growth suppression.

The regulation of purine nucleotide biosynthesis is summarized in **Figure 14.29**. The pathway is controlled to a considerable degree by PRPP availability. Several products of the pathway inhibit both ribose-5-phosphate pyrophosphokinase and glutamine-PRPP amidotransferase (the rate-limiting enzyme in IMP synthesis). The combined inhibitory effect of the end products is synergistic (i.e., the net inhibition is greater than the inhibition of each nucleotide acting alone). At the IMP branch point, both AMP and GMP regulate their own syntheses by feedback inhibition of adenylosuccinate synthetase and IMP dehydrogenase, respectively. The hydrolysis of GTP drives the synthesis of adenylosuccinate, whereas ATP drives XMP synthesis. This reciprocal arrangement facilitates the maintenance of appropriate cellular concentrations of adenine and guanine nucleotides.

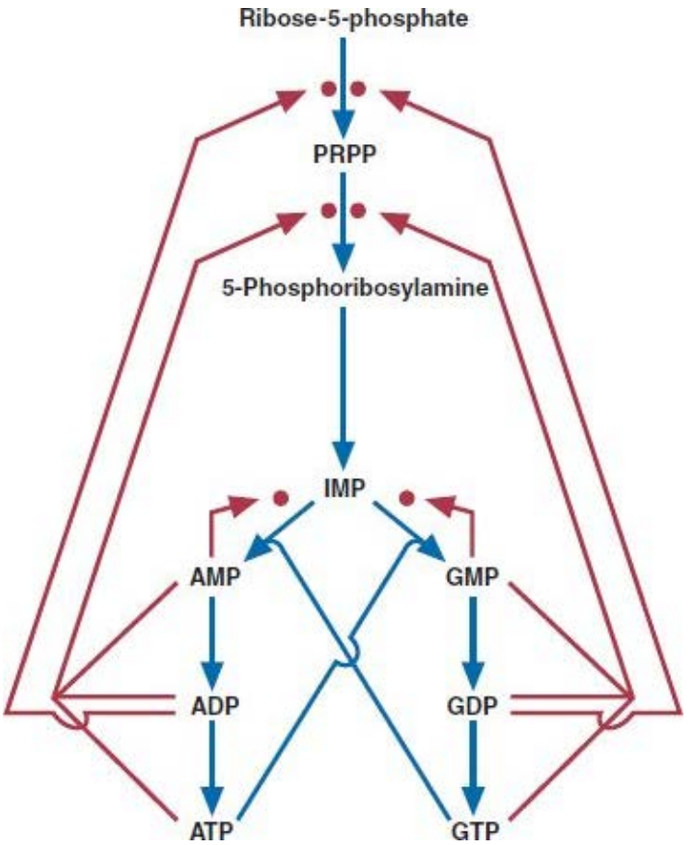


FIGURE 14.29

Purine Nucleotide Biosynthesis Regulation

Feedback inhibition is indicated by red arrows. The stimulation of AMP synthesis by GTP and GMP synthesis by ATP ensures a balanced synthesis of both purine nucleotide families.

PYRIMIDINE NUCLEOTIDE BIOSYNTHESIS Pyrimidine nucleotide synthesis occurs in the cytoplasm where the pyrimidine ring is assembled first and then linked to ribose phosphate. The carbon and nitrogen atoms in the pyrimidine ring are derived from bicarbonate, aspartate, and glutamine. Synthesis begins with the formation of carbamoyl phosphate in an ATP-requiring reaction catalyzed by carbamoyl phosphate synthetase II (CPSII) (**Figure 14.30**). (Carbamoyl phosphate synthetase I is a mitochondrial enzyme involved in the urea cycle and is described in Chapter 15.) One molecule of ATP provides a phosphate group, whereas the hydrolysis of another ATP drives the reaction. Aspartate transcarbamoylase (ATCase) catalyzes the reaction of carbamoyl phosphate with aspartate to form carbamoyl aspartate. The closure of the pyrimidine ring is then catalyzed by dihydroorotase. The product, dihydroorotate, is oxidized to form orotate. Dihydroorotate dehydrogenase, the enzyme that catalyzes this reaction, is an FMN flavoprotein associated with the inner mitochondrial membrane. (The NADH produced in this reaction donates its electrons to UQ in the ETC.) Once synthesized on the cytoplasmic face of the inner mitochondrial membrane, orotate is converted by orotate pyrophosphoribosyl transferase to orotidine-5′-phosphate (OMP), the first nucleotide in the pathway, by reacting with PRPP. Uridine-5′-phosphate (UMP) is produced when OMP is decarboxylated in a reaction catalyzed by OMP decarboxylase. Both orotate pyrophosphoribosyl transferase and OMP decarboxylase activities occur on a protein referred to as UMP synthase. UMP is a precursor for the other pyrimidine nucleotides. Two sequential phosphorylation reactions form UTP, which then accepts an amide nitrogen from glutamine to form CTP. The origin of the ring atoms in pyrimidines is illustrated in **Figure 14.31**.

Orotic Aciduria

In a rare genetic disease called *orotic aciduria*, there is excessive urinary excretion of orotic acid because UMP synthase is defective. Symptoms include megaloblastic anemia (not cured by vitamin B_{12} or folic acid) and delays in mental and physical growth. Treatment with a combination of pyrimidine nucleotides, which inhibit the production of orotate and provide the building blocks for nucleic acid synthesis, partially reverses the disease process.

DEOXYRIBONUCLEOTIDE BIOSYNTHESIS All the nucleotides discussed so far are ribonucleotides, molecules that are principally used as the building blocks of RNA, as nucleotide derivatives of molecules such as sugars, or as energy sources. The nucleotides required for DNA synthesis, the 2′-deoxyribonucleotides, are produced by reducing ribonucleoside diphosphates in a reaction catalyzed by ribonucleotide reductase (**Figure 14.32**). NADPH ultimately donates the electrons used in the synthesis of 2′-deoxyribonucleotides. Thioredoxin mediates the transfer of hydrogen atoms from NADPH to ribonucleotide reductase. The regeneration of reduced thioredoxin is catalyzed by thioredoxin reductase.

Ribonuclease reductase I, found in mammals, is a tetramer of two different subunits. Subunit 1 has a number of reactive thiols required for catalysis plus the allosteric sites involved in regulation. Subunit 2 possesses a critical binuclear Fe(III) center that generates and stabilizes a tyrosyl radical essential for enzyme function. The interface of the four subunits forms the active site. The tyrosyl radical initiates the radical-mediated reduction of substrate NDPs by abstracting an H atom from one of the thiols in subunit 1, generating a transient thiyl radical. The thiyl radical abstracts an H atom from C3′ of the substrate, generating a radical. A nearby thiol protonates the C2′—OH group, and it departs as H_2O, leaving behind a carbocation. Hydride ion shift from an active site thiol resolves the carbocation, and a disulfide bridge forms in the active site. The H atom abstracted by the thiyl radical is returned to C3′, and tyrosine transfers an H atom to resolve the thiyl radical. The product dNDP leaves the active site, and the enzyme is returned to its reduced free thiol state by electron transfer from NADPH mediated through thioredoxin (see p. 390) (**Figure 14.33**).

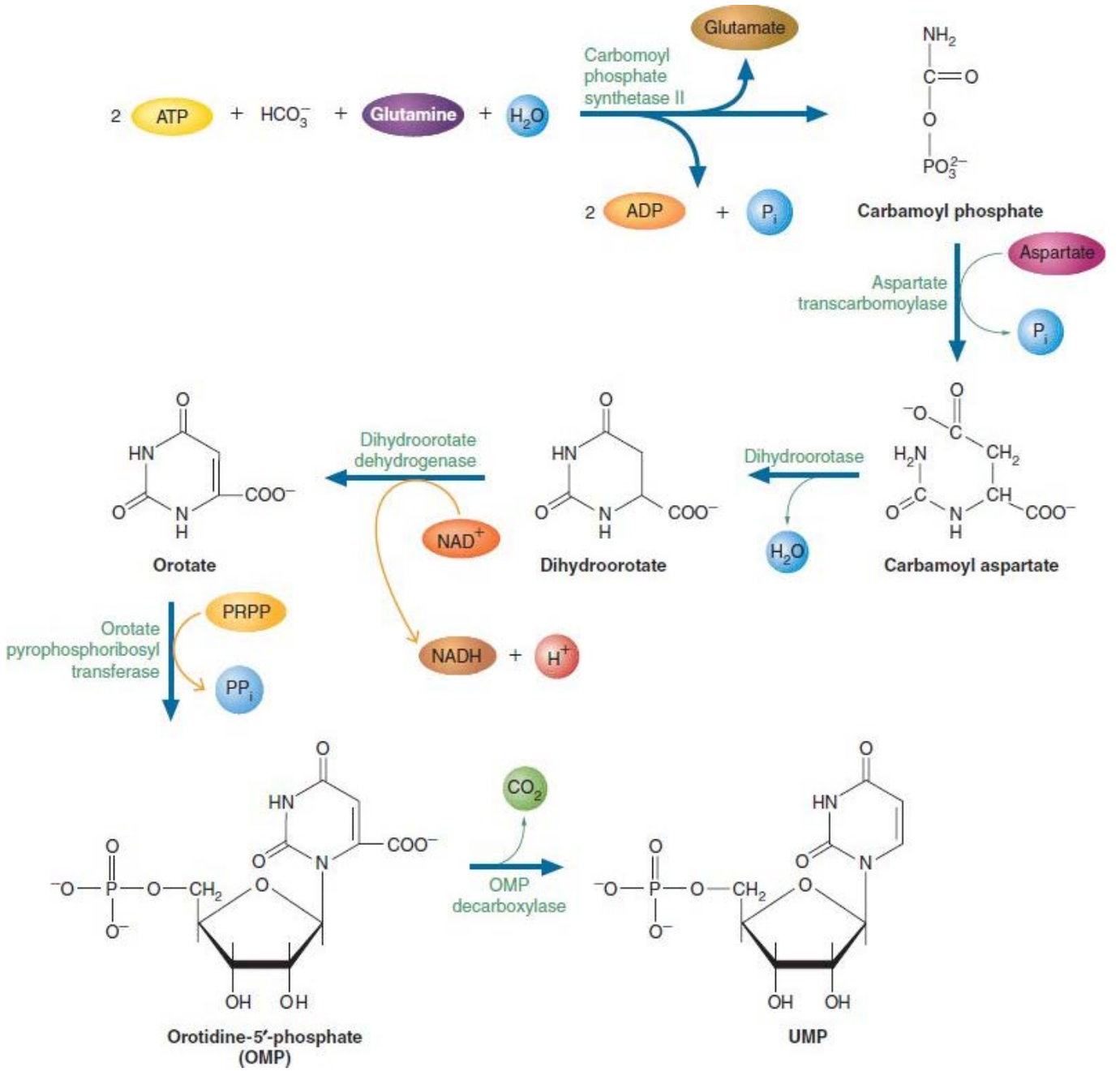


FIGURE 14.30

Pyrimidine Nucleotide Synthesis

The metabolic pathway in which UMP is synthesized is composed of six enzyme-catalyzed reactions. In mammals, the first three enzymatic activities in the pathway (**c**arbamoyl phosphate synthetase II, **a**spartate transcarbamoylase, and **d**ihydroorotate dehydrogenase) are located on a single polypeptide, referred to as CAD. Unlike the other enzymes in pyrimidine biosynthesis, which are cytoplasmic, CAD is located on the outer face of the inner mitochondrial membrane. The carbamoyl phosphate synthetase (CPS) II domain is activated by ATP and inhibited by UTP and CTP. In *E. coli*, the rate-limiting reaction in pyrimidine synthesis is catalyzed by the 12-subunit complex aspartate transcarbamoylase (**Figure 6.26**). Bacterial ATCase is stimulated by ATP and inhibited by UTP and CTP.

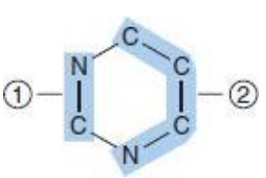


FIGURE 14.31

Origin of Pyrimidine Ring Atoms

The sources of the nitrogen and carbon atoms in pyrimidine rings are as follows: 1 = carbamoyl phosphate; 2 = aspartate.

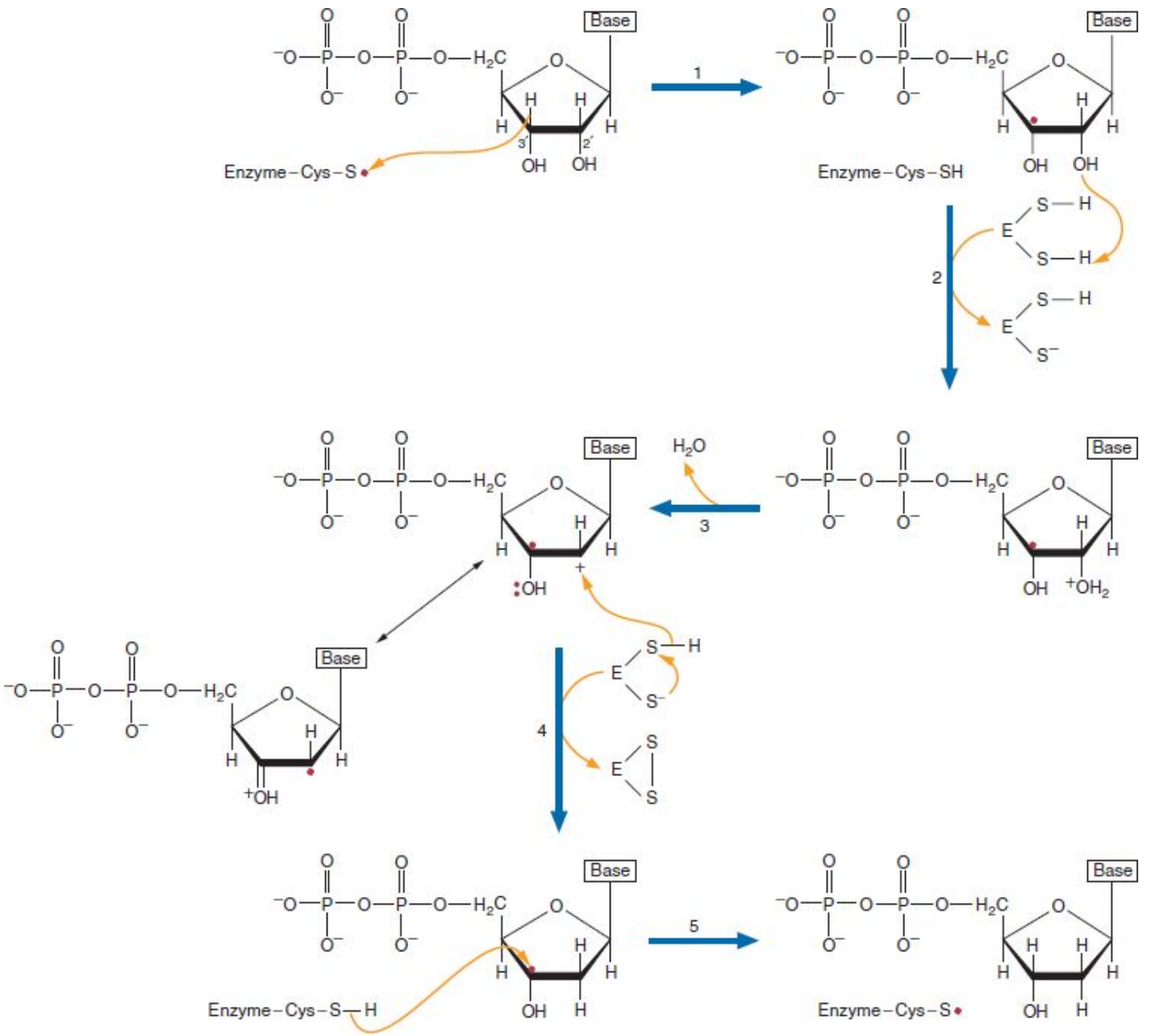


FIGURE 14.32

Mechanism of Ribonucleotide Reductase

The reaction begins with the tyrosyl radical-induced formation of a transient thiyl radical and the binding of NDP. (1) The thiyl radical abstracts a H atom from C3′. The C2′—OH (2) is protonated by a reactive thiol, and H_2O is eliminated (3) to generate a carbocation. A dithiol reduces the cation radical (4) and (5) an H atom is transferred from the initiating thiyl S to the C3′ and the product dNDP leaves the active site. The subsequent reduction of a disulfide mediated by thioredoxin/NADPH and regeneration of the tyrosyl radical returns the enzyme to its ready state to receive new substrate.

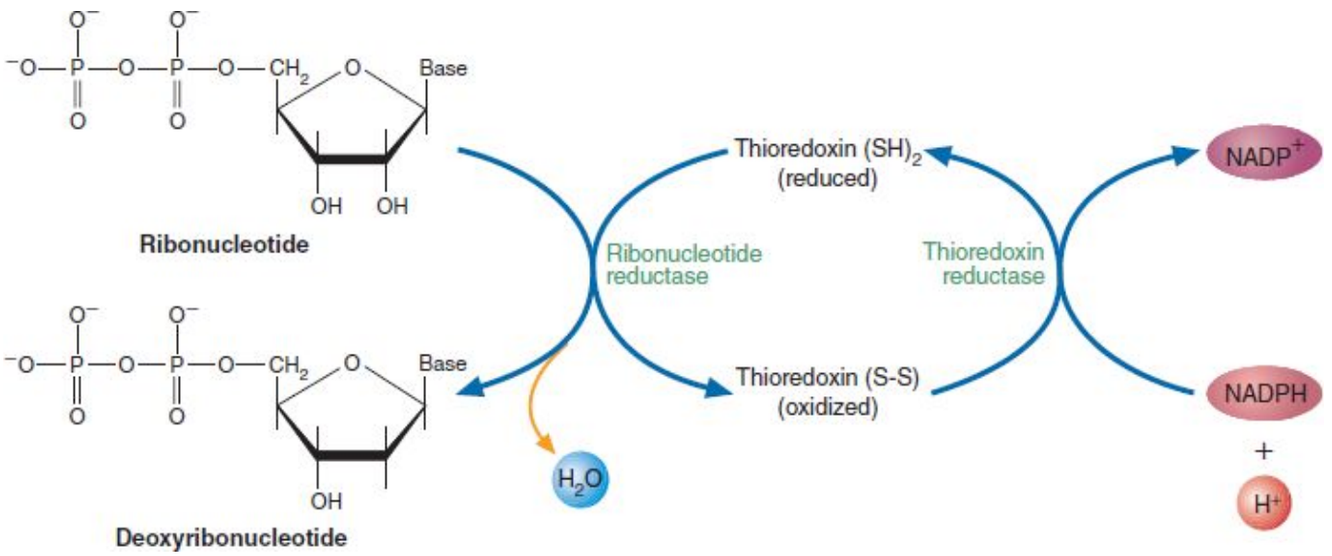


FIGURE 14.33

Deoxyribonucleotide Biosynthesis

Electrons for the reduction of ribonucleotides ultimately come from NADPH. Thioredoxin, a small protein with two thiol groups, mediates the transfer of electrons from NADPH to ribonucleotide reductase.

Regulation of ribonucleotide reductase is intricate. The binding of deoxyadenosine triphosphate to a regulatory site on the enzyme decreases catalytic activity. The binding of deoxyribonucleoside triphosphates to several other enzyme sites alters substrate specificity so that there are differential increases in the concentrations of each of the deoxyribonucleotides. This latter process balances the production of the 2′-deoxyribonucleotides required for cellular processes, especially that of DNA synthesis.

The deoxyuridylate (dUMP) produced by dephosphorylation of the dUDP product of ribonucleotide reductase is not a component of DNA, but its methylated derivative deoxythymidylate (dTMP) is. The methylation of dUMP is catalyzed by thymidylate synthase, which utilizes N^5,N^{10}-methylene THF. As the methylene group is transferred, it is reduced to a methyl group, while the folate coenzyme is oxidized to form dihydrofolate. THF is regenerated from dihydrofolate by dihydrofolate reductase and NADPH. (This reaction is the site of action of some anticancer drugs, such as methotrexate.) Deoxyuridylate can also be synthesized from dCMP by deoxycytidylate deaminase.

In mammals, carbamoyl phosphate synthetase II is the key regulatory enzyme in the biosynthesis of pyrimidine nucleotides. The enzyme is inhibited by UTP, the product of the pathway, and stimulated by purine nucleotides. In many bacteria, aspartate carbamoyl transferase is the key regulatory enzyme. It is inhibited by CTP and stimulated by ATP. The pyrimidine salvage pathway, which uses preformed pyrimidine bases from dietary sources or from nucleotide turnover, is of minor importance in mammals.

KEY CONCEPTS

- Nucleotides are the building blocks of the nucleic acids. They also regulate metabolism and transfer energy.
- The purine and pyrimidine nucleotides are synthesized in both de novo and salvage pathways.

Heme

Heme (**Figure 5.35**), one of the most complex nitrogen-containing molecules synthesized by mammalian cells, has an iron-containing porphyrin ring. Heme is an essential structural component

of hemoglobin (**Figure 5.40**), myoglobin (**Figure 5.37**), peroxidase, and the cytochromes (**Figure 10.8**). Although it occurs in almost all aerobic cells, the heme biosynthetic pathway is especially prominent in liver and reticulocytes (the nucleus-containing precursor cells of red blood cells in bone marrow). Both heme and chlorophyll are synthesized from the relatively simple components glycine and succinyl-CoA.

The pathways that produces heme and chlorophyll are outlined in a Biochemistry in Perspective essay available online. Note that heme biosynthesis is inhibited by heavy metal poisoning, a topic discussed in the online Biochemistry in Perspective essay Lead Poisoning.

Chapter Summary

1. Nitrogen, found in proteins, nucleic acids, and myriad other biomolecules, is an essential element in living systems. Biologically useful nitrogen, a scarce resource, is produced in a process referred to as nitrogen fixation. The nitrogenase enzyme complex that converts N_2 to NH_3 resides in some free soil bacteria, cyanobacteria, and symbiotic root nodule bacteria.
2. Organisms vary widely in their ability to synthesize amino acids. Some organisms (e.g., plants and some microorganisms) can produce all required amino acid molecules from fixed nitrogen. Animals can produce only some amino acids. Nonessential amino acids are produced from readily available precursor molecules, whereas essential amino acids must be acquired in the diet.
3. Two types of reaction play prominent roles in amino acid metabolism. In transamination reactions, new amino acids are produced when α-amino groups are transferred from donor α-amino acids to acceptor α-keto acids. Because transamination reactions are reversible, they play an important role in both amino acid synthesis and degradation. Ammonium ions or the amide nitrogen of glutamine can also be directly incorporated into amino acids and eventually other metabolites.
4. On the basis of the biochemical pathways in which they are synthesized, the amino acids can be divided into six families: glutamate, serine, aspartate, pyruvate, aromatics, and histidine.
5. Amino acids are precursors of many physiologically important biomolecules. Many of the processes that synthesize these molecules involve the transfer of carbon groups. Because many of these transfers involve one-carbon groups (e.g., methyl, methylene, methenyl, and formyl), the overall process is referred to as one-carbon metabolism. *S*-adenosylmethionine (SAM) and tetrahydrofolate (THF) are the most important carriers of one-carbon groups.
6. Molecules derived from amino acids include several neurotransmitters (e.g., GABA, the catecholamines, serotonin, and histamine) and hormones (e.g., melatonin). Glutathione is a tripeptide that plays an essential role in cells. The nucleotides, molecules that serve as the building blocks of the nucleic acids (as well as energy sources and metabolic regulators), possess heterocyclic nitrogenous bases as part of their structures. These bases, called the purines and the pyrimidines, are derived from various amino acid molecules. Heme is an example of a complex heterocyclic ring system that is derived from glycine and succinyl-CoA. The biosynthetic pathway that produces heme is similar to the one that produces chlorophyll in plants.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee**, where you can complete a multiple-choice quiz on synthesis of nitrogen-containing biomolecules to help you prepare for exams.

Chapter 14 Review Quiz

Suggested Readings

Burgos-Barragan G, et al. 2017. Mammals divert endogenous genotoxic formaldehyde into one-carbon metabolism. Nature 548:549–54.

Burrage LC, et al. 2014. Branched chain amino acid metabolism: from rare Mendelian diseases to more common disorders. Human Mol Gen 23(R1):R1–8.

Chen I. 2013. Lead's buried legacy. Sci Amer 309(3):28, 30.

Desai A, Sequeira JM, Quadros EV. 2016. The metabolic basis for developmental disorders due to defective folate transport. Biochimie 126:31–42.

Kuypers MM. 2015. A division of labor combined. Nature 528:487–8.

Lane AN, Fan TW-M. 2015. Regulation of mammalian nucleotide metabolism and biosynthesis. Nuc Acids Res 43(4):2466–85.

Newman AC, Maddocks ODK. 2017. One-carbon metabolism in cancer. British J Cancer 116:1499–504.

Stokstad E. 2016. The nitrogen fix. Science 353:1225–7.

Key Words

S-adenosylmethionine, *548*
amino acid pool, *533*
biogenic amine, *551*
branched-chain amino acid, *532*
catecholamine, *542*
essential amino acid, *532*
gasotransmitter, *552*
neurotransmitter, *551*
nitrogen fixation, *528*
nonessential amino acid, *532*
nucleoside, *554*
one-carbon metabolism, *543*
purine, *554*
pyrimidine, *554*
racemization, *534*
tetrahydrobiopterin, *543*
tetrahydrofolate, *545*
transamination, *533*
vitamin B_{12}, *545*

Review Questions

SECTION 14.1

Comprehension Questions

1. Define the following terms:
    a. nitrogen fixation

b. MoFe protein
c. Fe protein
d. nitrogen assimilation
e. dead zone

Fill in the Blanks

2. Usable nitrogen in the soil for use by living organisms such as plants is usually scarce because of a limited number of _____.
3. The incorporation of inorganic nitrogen into organic molecules is called _____ .
4. _____ is the product of the Haber–Bosch process.
5. Organisms such as *Nitrosomas* oxidize NH_4^+ to form _____.

Short-Answer Questions

6. Why are nitrogen-containing molecules limited in the biosphere? Give two reasons.
7. Nitrogenase complexes are irreversibly inactivated by oxygen. Explain how nitrogen-fixing bacteria solve this problem.
8. Use reaction equations to show how NH_4^+ is incorporated into glutamate in plants.
9. Describe what happens to organic nitrogen-containing molecules when organisms such as plants die.
10. What is the net reaction for nitrogen fixation?
11. What is the function of ATP in the conversion of glutamate to glutamine?
12. In the nitrogen reductase system, for every molecule of ammonia released, one molecule of hydrogen gas is also produced. What us the source of the hydrogen gas?

Critical-Thinking Questions

13. Both oxygen and nitrogen are present in the atmosphere as gases. Oxygen (O_2) is reactive, and nitrogen (N_2) is relatively inert. What structural features of the two molecules account for this difference?
14. Describe the mechanism of the nitrogenase complex in which N_2 is converted to NH_3.
15. Long before living organisms developed the nitrogenase system, there was ammonia in Earth's atmosphere. Suggest a naturally occurring method of fixing nitrogen.

SECTION 14.2

Comprehension Questions

16. Define the following terms:

    a. nonessential amino acids
    b. essential amino acids
    c. BCAA
    d. amino acid pool
    e. transamination

17. Define the following terms:

    a. pyruvate/alanine
    b. racemization

c. PLP
d. ping-pong mechanism
e. de novo

18. Define the following terms:

a. glutamate family
b. THF
c. AST
d. GOT
e. shikimate pathway

19. Define the following terms:

a. BH_4
b. PRPP
c. CBS
d. CSE
e. quinonoid intermediate

Fill in the Blanks

20. The amino acid _____ is a cyclized derivative of glutamate.
21. _____ is an amino acid derivative of phenylalanine.
22. Amino acids that must be provided in the diet are called _____ amino acids.
23. _____, _____, and _____ are gasotransmitters.
24. The branched chain amino acids are _____,_____, and _____.

Short-Answer Questions

25. In PLP-catalyzed reactions, the pyridinium ring acts as an electron sink. Describe this process.
26. What are the two major classes of neurotransmitter in reference to their effects on postsynaptic cells? How do their modes of action differ? Give an example of each type of neurotransmitter.
27. Illustrate the pathways to synthesize the following amino acids: glutamine, serine, arginine, glycine, and cysteine.
28. To which synthetic family does each of the following amino acids belong?

a. alanine
b. phenylalanine
c. methionine
d. tryptophan
e. histidine
f. serine

29. Pyridoxal phosphate acts as an intermediate carrier of amino groups during transamination reactions. Write a series of reactions to show the role of pyridoxal phosphate in the transamination reaction involving alanine and α-ketoglutarate.

Critical-Thinking Questions

30. By definition, essential amino acids are not synthesized by an organism. Arginine is classified as an essential amino acid in infants, even though it is part of the urea cycle. Explain.
31. In PLP-catalyzed reactions, the bond broken in the substrate molecule must be perpendicular to the plane of the pyridinium ring. Considering the bonds present in this ring, describe why this

arrangement stabilizes the carbanion.

32. During your experimental investigation of the conversion of glutamine to proline, you have labeled the γ-carbonyl group of glutamine with ^{14}C. What carbon in the proline product will be labeled?
33. Transamination reactions have been described as ping-pong reactions. Use the reaction of alanine with α-ketoglutarate to indicate how this ping-pong reaction works.
34. Tyrosine is a nonessential amino acid in humans. Under what circumstance would it become an essential amino acid?

SECTION 14.3

Comprehension Questions

35. Define the following terms:
    a. one-carbon metabolism
    b. para-aminobenzoic acid
    c. cobalamin
    d. pernicious anemia
    e. intrinsic factor
36. Define the following terms:
    a. SAM
    b. SAH
    c. polyamine
    d. methotrexate
    e. glutathione
37. Define the following terms:
    a. neurotransmitter
    b. biogenic amines
    c. γ-glutamyl cycle
    d. gasotransmitter
    e. NOS
38. Define the following terms:
    a. purine
    b. pyrimidine
    c. nucleoside
    d. lactam
    e. lactim
39. Define the following terms:
    a. *anti* nucleoside
    b. HGPRT
    c. Lesch–Nyhan syndrome
    d. purine salvage
    e. gout

Fill in the Blanks

40. The reactions in which single-carbon atoms are transferred from one molecule to another are referred to collectively as _____ .
41. Neurotransmitters that are either amino acids or amino acid derivatives are referred to as _____ .
42. Adenine and guanine are examples of the _____ class of nitrogenous bases.
43. Uracil is an example of the _____ class of nitrogenous bases.
44. When a purine or a pyrimidine is linked through a β-N-glycosidic link to C-1 of a pentose, the molecule is called a _____ .
45. The most abundant intracellular thiol is _____ .
46. The catecholamines (dopamine, norepinephrine, and epinephrine) are all derived from the amino acid _____ .

Short-Answer Questions

47. Describe the cause and symptoms of Lesch–Nyhan syndrome.
48. Orotic aciduria is an autosomal recessive disease caused by a deficiency of the enzyme UMP synthase (a bifunctional protein composed of orotate pyrophosphoribosyl transferase and OMP decarboxylase activities). What symptom accounts for the disease's name? Suggest a possible treatment.
49. Describe the functions of the gasotransmitter CO.
50. How is nitric oxide synthesized? What cofactors are required?
51. In pyrimidine nucleosides, the *anti* conformation predominates. Explain. Do the purine nucleosides have similar interactions?
52. Vegans (individuals who do not consume meat or meat products such as milk and eggs) may develop pernicious anemia if they are treated with antibiotics. Explain.
53. Describe the γ-glutamyl cycle. How many ATPs are required to synthesize glutathione from oxoproline?
54. What is the biologically active form of folic acid? How is it formed?
55. What carbon in uracil is derived from CO_2? Draw the uracil structure to illustrate your answer.
56. Individuals with inadequate tyrosine metabolism are often light sensitive and easily develop severe cases of sunburn. Explain. [*Hint*: the skin pigment melanin is derived from L-DOPA (L-3,4-dihydroxyphenylalanine, a precursor in catecholamine biosynthesis). Also, refer to the online Biochemistry in Perspective essay in this chapter entitled The Catecholamines.]

Critical-Thinking Questions

57. Radiation exerts part of its damaging effect by causing the formation of hydroxyl radicals. Write a reaction equation to explain how glutathione acts to protect against this form of radiation damage.
58. In purine nucleotide synthesis, the carbon and nitrogen atoms are derived from bicarbonate, aspartate, and glutamine. Devise a simple experiment to prove the source of the nitrogen atoms. Do not forget to take into account the nitrogen exchange in amino acids.
59. In water, cytosine will gradually convert to uracil. Using reactions that you have learned in this and previous chapters, show how this conversion takes place.
60. Why do marathon runners prefer beverages with sugar instead of amino acids during a long run?
61. Considering the genetics of Lesch–Nyhan syndrome, explain how a female could inherit the disease considering that afflicted male patients do not have offspring.
62. Describe the synthesis of epinephrine (a hormone and neurotransmitter) from tyrosine. [*Hint*: Refer to the online reading The Catecholamines.] Describe the possible effects of inadequate epinephrine biosynthesis.

63. Consumption of nitrate-rich vegetables such as spinach have been shown to reduce blood pressure in prehypertensive individuals (i.e., those having no history of clinical hypertension), a phenomenon attributed to increased nitric oxide synthesis. Describe the reactions that convert nitrate (NO_3^-) to nitric oxide (NO•) in the body.

MCAT Study Questions

64. Branched-chain amino acids are important in metabolic processes for all of the following reasons except

    a. They serve as nonessential amino acids precursors.
    b. Their major role is in liver amino acid metabolism.
    c. They are a major transport form of amino nitrogen in blood.
    d. They are a critical source of hydrophobic side chains in DNA-binding proteins.

65. In transamination reactions, which of the following is not a donor amino acid–acceptor α-keto acid pair?

    a. aspartate and pyruvate
    b. alanine and pyruvate
    c. glutamate and α-ketoglutarate
    d. aspartate and oxaloacetate

66. All of the following are neurotransmitters except

    a. glutamate
    b. glycine
    c. H_2S
    d. GSH

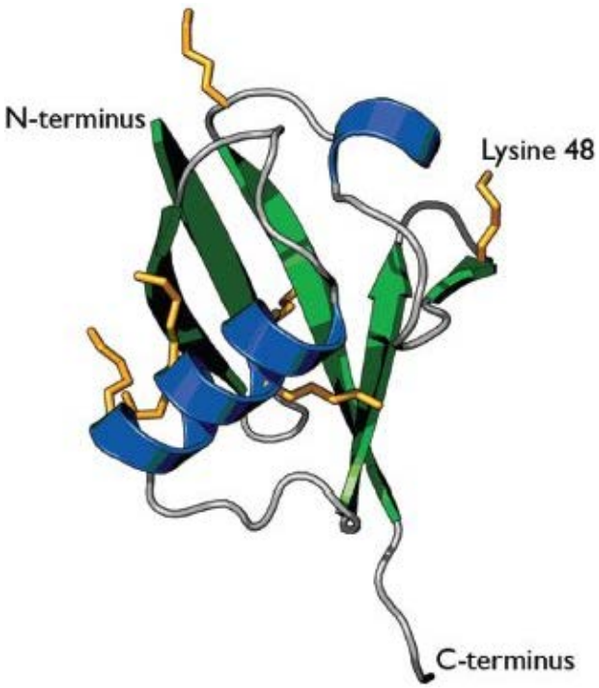
67. Which of the following statements about folic acid is not true?

    a. Folic acid is activated by folate synthase.
    b. Folic acid structure contains a pteridine ring and para-amino benzoic acid.
    c. NADPH is required in the reactions activating folic acid.
    d. Tetrahydrofolate is the biologically active derivative of folic acid.

68. Which of the following is required in transamination reactions?

    a. folic acid
    b. vitamin B_{12}
    c. niacin
    d. pyridoxal phosphate

CHAPTER 15 Nitrogen Metabolism II: Degradation



Ubiquitin Ubiquitin is a cellular protein in all eukaryotes. Its best-understood function is its involvement in protein degradation. Once the first ubiquitin molecule is covalently bonded to a target protein, additional ubiquitin molecules are linked to form a polyubiquitin chain via the formation of an amide bond between the C-terminal glycine of each incoming ubiquitin and the side chain amino group of lysine 48 in the previous ubiquitin molecule. Proteins with at least four linked ubiquitins are directed to the proteasome, a large protein-degrading machine. Ubiquitin has six other lysine residues that also serve as ubiquitination substrates, leading to polyubiquitin chains that together play regulatory roles in almost all cell processes.

OUTLINE

DEGRADATIVE PATHWAYS AND HUMAN DISORDERS

15.1 PROTEIN TURNOVER
Ubiquitin Proteosomal System
Autophagy-Lysosomal System

15.2 AMINO ACID CATABOLISM
Deamination
Urea Synthesis

Control of the Urea Cycle
Catabolism of Amino Acid Carbon Skeletons

15.3 NUCLEOTIDE DEGRADATION
Purine Catabolism
Pyrimidine Catabolism

Biochemistry in Perspective
Disorders of Amino Acid Catabolism

AVAILABLE ONLINE
Biochemistry in Perspective
Hyperammonemia
Biochemistry in Perspective
Gout
Biochemistry in Perspective
Catecholamine Inactivation
Biochemistry in Perspective
Heme Biotransformation

Degradative Pathways and Human Disorders

What do the following four people have in common? A 62-year-old man has just been diagnosed with Parkinson's disease, a movement disorder characterized by tremor, rigidity, and muscle weakness. An 84-year-old woman is in the last stage of Alzheimer's disease, the most common form of dementia. A 7-year-old child is suffering symptoms (unsteady walking, slurred speech, tremor, and enlarged spleen and liver) that have led to a diagnosis of Niemann–Pick disease type C. Tuberculosis, an infectious disease caused by *Mycobacterium tuberculosis*, has just been confirmed in a diabetic patient with symptoms of chronic cough, night sweats, chills, appetite loss, and fatigue.

Puzzled? In all of these cases, there is either a genetic defect or an altered functioning of the *endosomal lysosomal system*. In this system, the fusion of several types of cellular vesicles with lysosomes is the culmination of several degradative pathways. Endocytosis and phagocytosis (p. 52) are pathways in which cells degrade extracellular material. Autophagy is a mechanism whereby intracellular components (e.g., proteins and organelles) are degraded. Intense and ongoing research into the detailed biochemical mechanisms of these lysosome-related diseases provides new and unanticipated opportunities for developing effective treatments.

Both Parkinson's and Alzheimer's diseases (pp. 436–37) are neurodegenerative disorders in which cells fail to remove accumulating aggregates of specific misfolded proteins. Because protein aggregates are poor substrates for proteasomes (protein-degrading machines, p. 572), autophagy is activated. Over time, however, genetic or acquired functional alterations in the autophagic process (e.g., mutations in proteins that result in slower protein aggregate clearance or the sequestration of autophagic proteins in protein aggregates) can impede autophagy, resulting in impaired cell function. In healthy neurons, autophagy is so efficient that few autophagic vacuoles are normally observed. The large numbers of these structures that are seen in neurons from affected brain regions in both diseases are an indication of the dire consequences of a failed autophagic process.

Niemann–Pick disease type C (NPC) is a lysosomal storage disease (p. 54). In the better-known forms of this disease—type A (infantile onset) and type B (adolescent onset)—the enzyme sphingomyelinase is missing or deficient. NPC, caused by the accumulation of cholesterol and glycolipids in lysosomes, is linked to a mutated transmembrane protein that is believed to have a transport function. Symptoms that lead to a diagnosis of NPC, usually in school-age children, include liver and/or spleen enlargement and progressive neurological problems (e.g., unsteady walking, difficulty in swallowing, epilepsy, and abnormal movements

and posture). Eventually, NPC patients become bedridden and severely demented.

Tuberculosis (TB), usually acquired by inhalation of the bacterium *M. tuberculosis* into the lungs, is a chronic and potentially fatal infection. Over many thousands of years, *M. tuberculosis* has evolved several strategies that evade the human immune system. One of these strategies is to actively subvert the endosomal lysosomal system of macrophages. Macrophages are one of several types of phagocytic immune system cells that defend the body from pathogenic microbes. Phagocytosis results in the formation of a phagolysosome (**Figure 10.24**) as a result of the fusion of a phagosome, an intracellular membrane-bound vesicle that contains ingested foreign cells, with a lysosome. This process usually leads to the destruction of pathogens. The release of a fusion-preventing protein by *M. tuberculosis* is one factor in a complex mechanism that allows this organism's survival and growth within the macrophage. Tubercles, aggregates of infected macrophages, are the characteristic lesions of TB.

Overview

THE METABOLISM OF NITROGEN-CONTAINING MOLECULES SUCH AS PROTEINS AND NUCLEIC ACIDS DIFFERS SIGNIFICANTLY FROM THAT OF CARBOHYDRATES AND LIPIDS. Whereas the latter molecules can be stored and mobilized as needed for biosynthetic reactions or for energy generation, there is no nitrogen-storing molecule. (One exception to this rule is storage protein in seeds.) Organisms must constantly replenish their supply of usable nitrogen to replace organic nitrogen that is lost in catabolism. For example, animals must have a steady supply of amino acids in their diets to replace the nitrogen excreted as urea, uric acid, and other nitrogenous waste products.

Despite their apparent stability, most living cells are constantly undergoing renovation. One of the most obvious aspects of cellular renovation is the turnover of protein and nucleic acids, a process that results in the continuous flow of nitrogen atoms through living organisms. Living organisms recycle organic nitrogen into a variety of metabolites before the element is reconverted to its inorganic form.

Animals dispose of excess nitrogen-containing molecules (i.e., amino acids and nucleotides) by producing and then excreting nitrogenous waste molecules. Although many variations are observed among species, the following generalizations can be made. The nitrogen in amino acids is removed by deamination reactions and converted to ammonia. The toxic nature of this molecule requires that it be detoxified and/or excreted as fast as it is generated. Aquatic invertebrates such as crustaceans, called *ammonotelic* organisms, release ammonia directly into surrounding water. *Uricotelic* organisms such as insects, lizards, and birds, in which water conservation is a significant issue, excrete uric acid, which is less toxic than ammonia and can be released in a nearly solid form. Terrestrial animals such as amphibians and mammals are referred to as *ureotelic* because they convert ammonia to urea, which can be excreted with small amounts of water. Mammals also excrete uric acid, the nitrogenous waste product of purine nucleotide catabolism.

The chapter begins with a discussion of protein turnover, a critically important process that allows cells both to respond efficiently to ever-changing metabolic circumstances and to maintain proteostasis by disposing of damaged, and potentially toxic, proteins. The nitrogen atoms in proteins are traced from amino acids to their degradation products. The chapter ends with descriptions of the catabolic pathways of the nucleotides. A brief review of neurotransmitter degradation is provided in the online reading Catecholamine Inactivation. The degradation of the nitrogen-containing porphyrin molecule heme is described in the online reading Heme Biotransformation.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essays on catecholamine inactivation and heme biotransformation.

15.1 PROTEIN TURNOVER

The cellular concentration of each type of protein is a consequence of a balance between its synthesis and its degradation. Although it appears to be wasteful, the continuous degradation and synthesis of proteins, a process referred to as **protein turnover**, serves several purposes. First, metabolic flexibility is afforded by relatively quick changes in the concentrations of key regulatory enzymes, peptide hormones, and receptor molecules. Numerous physiological processes are just as dependent on timely degradative reactions as they are on synthetic ones. For example, the progression of eukaryotic cells through the phases of the cell cycle (Section 18.1) is regulated by the precisely timed synthesis and degradation of a class of proteins called the *cyclins*. Protein turnover also provides the amino acids required for the synthesis of proteins when nutrient levels are low and protects cells from the accumulation of abnormal proteins. Despite numerous safeguards, protein synthesis and the folding process are error prone. As many as one-third of all proteins are degraded within a few minutes of their synthesis because of errors in transcription, translation, or folding.

Proteins differ significantly in their turnover rates, which are measured in half-lives. (A *half-life* is the time required for 50% of a specified amount of a protein to be degraded.) Proteins that play structural roles typically have long half-lives. For example, many connective tissue proteins (e.g., the collagens) have half-lives that are measured in years. In contrast, the half-lives of regulatory enzymes are typically measured in minutes. **Table 15.1** lists several examples.

TABLE 15.1 Human Protein Half-Lives

Protein	Approximate Value of Half-Life (h)
Ornithine decarboxylase	0.5
Tyrosine aminotransferase	2
Tryptophan oxygenase	2
PEP carboxykinase	5
Arginase	96
Aldolase	118
Glyceraldehyde-3- phosphate dehydrogenase	130
Cytochrome c	150
Hemoglobin	2880

Although some proteins are degraded by proteolytic enzymes in cytoplasm (e.g., Ca^{2+}-activated calpains), most cellular proteins are degraded by two major systems: the ubiquitin proteasomal system and the autophagy lysosomal system (**Figure 15.1**).

Ubiquitin Proteasomal System

In the **ubiquitin proteasomal system**, or **UPS** (**Figure 15.1a**), protein degradation is initiated with a covalent modification referred to as ubiquitination, in which **ubiquitin**, a highly conserved 8.5-kDa protein, is attached to substrate proteins. UPS degrades most short-lived proteins (e.g., regulatory proteins such as transcription factors). UPS is also triggered by ERAD (ER-associated protein degradation: p. 49).

The mechanisms that target protein for destruction by ubiquitination are not fully understood. It is known that many proteins have sequence motifs, called *degrons*, which mark them for proteolytic destruction. Examples include the following:

1. **N-terminal residues**. Very short-lived proteins often have basic (Arg, Lys, His) or bulky hydrophobic (Leu, Phe, Tyr, Trp, Ile) N-terminal residues. More stable proteins characteristically have sulfur-containing, hydroxyl-containing, or nonbulky hydrophobic amino acids at the N-terminus.
2. **Peptide motifs**. Proteins with certain homologous sequences are rapidly degraded. For example, proteins that have extended sequences containing proline, glutamate, serine, and threonine have half-lives of less than 2 hours. (PEST sequences are named for the one-letter abbreviations for these amino acids. See **Table 5.1**.) Ensuring rapid ubiquitination is the *cyclin destruction box*, a set of homologous nine-residue sequences near the N-terminus of cyclins.

Once ubiquitinated, target proteins are transferred to massive proteolytic molecular machines called **proteasomes** that cleave them into peptide fragments with an average of seven to eight amino acid residues. The peptide fragments are further degraded by cytoplasmic proteases to amino acids that can be recycled into new protein molecules. **Ubiquitination**, the attachment of the 76-residue ubiquitin to worn-out or damaged proteins or short-lived regulatory proteins, occurs in several stages and involves three enzyme classes: E1, E2, and E3.

In the first step (**Figure 15.1a**), E1 (*ubiquitin-activating enzyme*) activates a ubiquitin molecule via *adenylylation* (the covalent linkage of an AMP to a hydroxyl side chain of a protein) and transfers it to an active site thiol of E1 to form a high-energy thioester. The C-terminal glycine carboxyl group of the ubiquitin molecules participates in this reaction. Ubiquitin is then transferred to an active site thiol of E2 (*ubiquitin-conjugating enzyme*) via a transthiolation reaction. There is only one E1 but at least 50 E2 enzymes in mammalian cells. The E2 enzymes vary in their specificity for association with E3 (*ubiquitin ligase*). The E3 enzymes (as many as 1000 in mammals) determine the substrate specificity of ubiquitination because they interact with E2 and the target protein. E3 enzymes transfer ubiquitin to a specific internal lysine side chain of the target protein via a thioester-to-amide transition. An E3 enzyme recognizes its target protein by binding to a degron, which may be revealed either by posttranslational modification (e.g., acetylation, hydroxylation, or proteolytic cleavage) or by a change in a protein's conformation that exposes hydrophobic residues. Subsequent ubiquitination lengthens the ubiquitin tag on the protein from 4 to 50 units. The specificity of ubiquitination—and therefore regulated proteolysis—derives in part from the substrate specificities of the large number of E2s and E3s.

KEY CONCEPTS

- Protein turnover, the continuous synthesis and degradation of proteins, provides living organisms with metabolic flexibility and protects cells from the accumulation of abnormal proteins.
- Most short-lived cellular proteins are degraded by the ubiquitin proteasomal system to yield short peptides. Long-lived proteins and organelles are degraded by the autophagy lysosomal system.
- The amino acid products of the peptides that are cleaved by cytoplasmic proteases enter the amino acid pool

and are available for incorporation into new protein molecules.

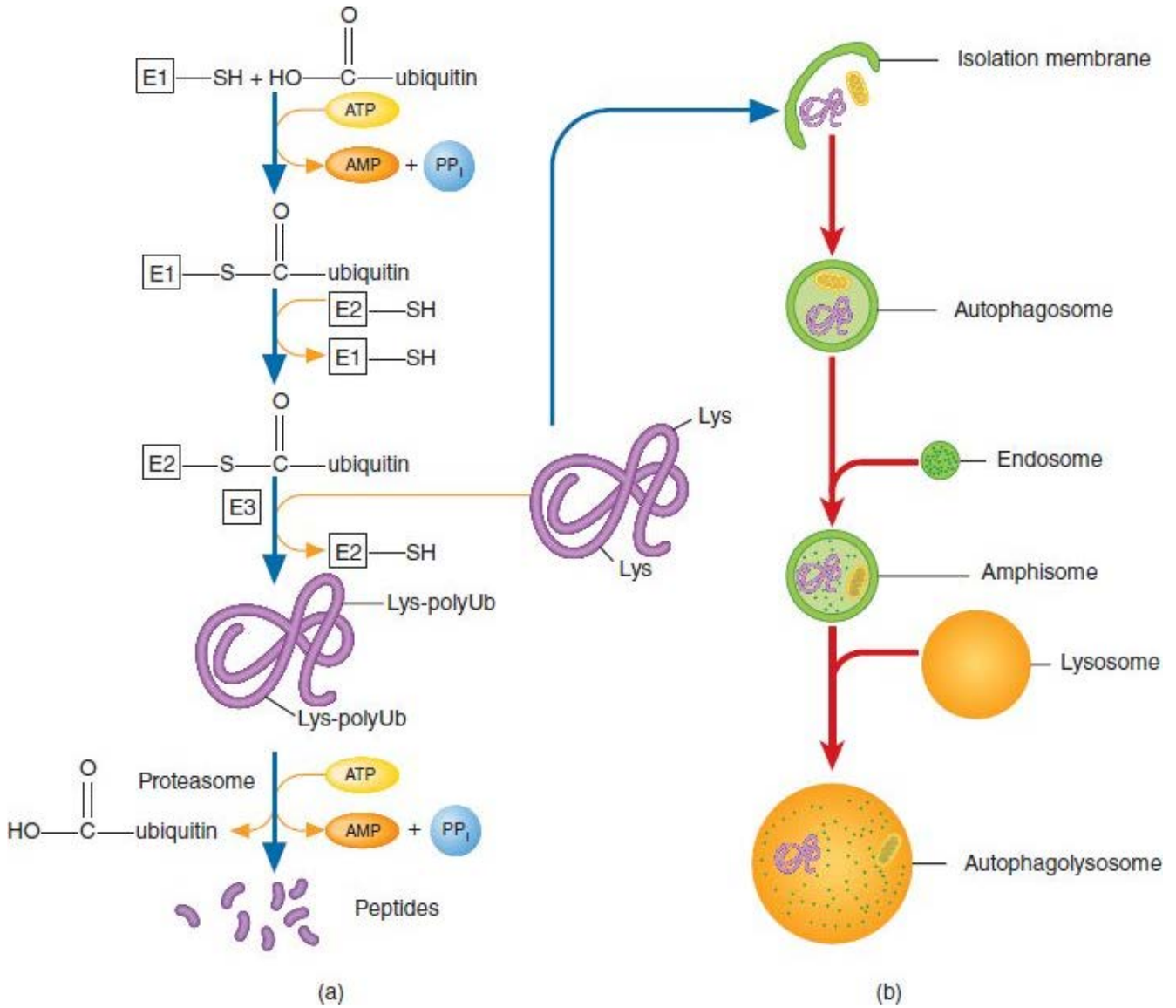


FIGURE 15.1

The Ubiquitin Proteasomal and the Autophagy Lysosomal Systems

(a) Protein ubiquitination begins with the ATP hydrolysis–driven formation of a thiol ester bond between E1 (ubiquitin-activating enzyme) and ubiquitin. Ubiquitin is then transferred to E2 (ubiquitin-conjugating enzyme). E3 (ubiquitin ligase) transfers ubiquitin from E2 to a lysine side chain on the target protein and subsequently to ubiquitin moieties already on the target protein. Polyubiquitination continues until there are 4 to 50 ubiquitins attached to the target lysine residues. The polyubiquitinated protein is then degraded by a proteasome. (b) In autophagy, a seemingly random process, an isolation membrane begins to form that surrounds and sequesters cytoplasmic components. The expanding isolation membrane eventually seals to form the autophagosome. When the autophagosome fuses with an endosome, forming an amphisome, the internal pH begins to drop. Fusion of the amphisome with a lysosome results in the formation of the autophagolysosome. Lysosomal enzymes then degrade the cargo. The products are recycled or degraded to generate energy. Once thought to be independent pathways, ubiquitin proteasomal system (UPS) and autophagy are interrelated. They share some regulatory proteins, and under certain conditions either pathway may degrade some protein substrates.

Ubiquitination also has nonproteolytic functions. The reversible covalent attachment of ubiquitin to target proteins is a pervasive mechanism used to control diverse cell processes such as cell cycle

control, transcription, DNA repair, and cell signaling.

The proteasome (**Figure 15.2**) is a large (2.5 MDa) multisubunit complex in the form of a hollow cylinder with dimensions of 15 by 11.5 nm. The interior chamber at its widest and narrowest is 5.3 nm and 1.3 nm, respectively. Referred to as the 26S proteasome, this structure consists of a 20S core particle and a 900-kDa 19S regulatory particle. (The Svedberg unit [S] is a measure of the rate at which a particle sediments in an ultracentrifuge. Since S values are related to both the mass and shape of particles, they are not additive.) The 20S particle consists of four heptameric protein rings ($\alpha_7\beta_7\beta_7\alpha_7$). The two inner β rings possess three different types of proteolytic activity that face the inner chamber. Chymotrypsin-like, trypsin-like, and caspase-like enzymatic activities cleave peptide bonds after hydrophobic, basic, and acidic amino acid residues, respectively. As a result, proteasomes are capable of degrading any naturally occurring polypeptides. The α-ring components have an N-terminal segment that limits access to the narrow proteolytic chamber to unfolded peptide chains. The 19S particles consist of a nine-subunit *lid* that binds directly to the α-ring of the core particle and a nine-subunit *base*. Lid subunits participate in substrate selection (by binding to ubiquitinated substrate protein) and ubiquitin removal. Six of the base subunits possess ATPase activity. ATP binding to the ATPase subunit is required for the association of 19S and 20S particles, the opening of the narrow 20S gate, the translocation of the unfolded polypeptide through it into the 20S chamber, and proteolysis. It appears that ATP hydrolysis is the driving force that unfolds a de-ubiquitinated polypeptide, which is required for threading it through the narrow gate into the catalytic chamber. One protein is degraded at a time, and the 6- to 10-residue peptide products are released from the proteasome for hydrolysis to free amino acids by cytoplasmic proteases.

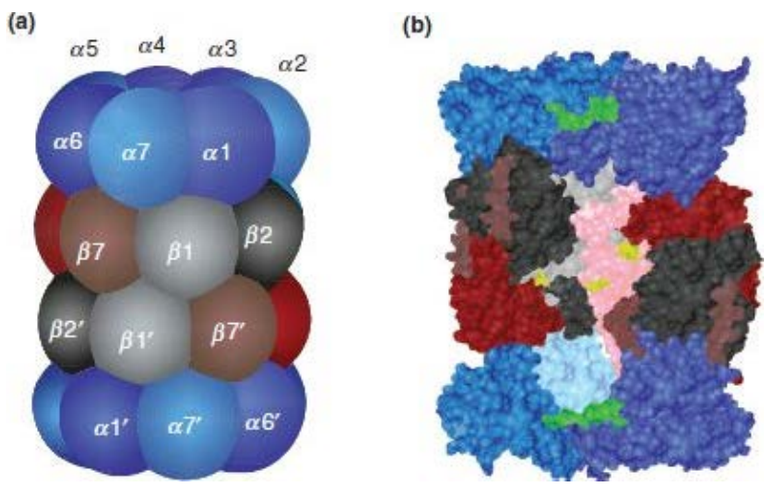


FIGURE 15.2

The Proteasome

(a) The 20S proteasome (core particle) is a 700-kDa barrel-shape structure that contains 28 proteins: two α rings with seven subunits each and two β rings with seven subunits each. (b) A cutaway view reveals the inner catalytic chamber.

Autophagy-Lysosomal System

Autophagy ("self-eating") is a cellular degradation pathway in which cell components, most notably long-lived proteins and organelles, are degraded by the hydrolytic enzymes within lysosomes. Autophagy has several roles in cells. In addition to its obvious role in degrading worn-out or damaged cell components, autophagy provides a recycling mechanism that maintains vital functions when nutrient levels are low (during fasting, overnight, for example). Autophagy is also involved in the regulation of development (cellular remodeling) and the destruction of invading microorganisms.

There are three forms of autophagy: chaperone-mediated autophagy, microautophagy, and

macroautophagy. **Chaperone-mediated autophagy** is a receptor-mediated process in which specific proteins that are bound to a chaperone complex are unfolded and then translocated into a lysosome, where they are degraded by lysosomal proteases. In **microautophagy**, small amounts of cytoplasm are directly engulfed by lysosomes. **Macroautophagy** (**Figure 15.1b**), often referred to as autophagy, uses a lysosomal pathway for bulk degradation of cytoplasmic components. It is the major catabolic mechanism used by eukaryotic cells to maintain optimal function and respond to changing environmental conditions.

Autophagy is induced by a large number of stressors including ER stress (e.g., stress caused by the unfolded protein response; see p. 49), hypoxia, oxidative stress, nutrient deprivation, high temperature, and viral infections. Autophagy begins with the formation of a double-membrane structure called the *isolation membrane* (probably originating from a ribosome-free zone of the RER). The isolation membrane surrounds and sequesters cytoplasmic material as it expands and eventually closes to form the autophagosome. The autophagosome then fuses with an endosome (p. 52) to form an *amphisome*. Endosomal membrane-bound ATPases begin pumping protons into the amphisome lumen. Eventually, the outer membrane of the amphisome fuses with a lysosome to form an autophagolysosome. Lysosomal enzymes then proceed to digest the cytoplasmic cargo and the inner amphisome membrane. The degradation products (e.g., amino acids and sugars) are exported to the cytoplasm where they may be used in biosynthesis and/ or degraded to generate energy. Any autophagolysosomes that contain substances that resist digestion, referred to as residual bodies, remain in the cytoplasm. For example, *lipofuscin granules*, which contain indigestible brown-pigmented debris, are found in aging nerve, heart, kidney, and adrenal cells.

Autophagy is a housekeeping process that operates at a basal level in almost all eukaryotic cells. It can be rapidly upregulated when cells are under stress or when energy and nutrient levels are low. For example, the protein synthesis initiation factor 2α (eIF2α) and AMPK stimulate autophagy when the levels of nutrients and energy, respectively, are low. When nutrient and energy levels are high, autophagy is inhibited by the serine-threonine kinase mammalian target of rapamycin (mTOR). (Rapamycin is a bacterial molecule used clinically to prevent the rejection of transplanted organs.) mTOR, which integrates intracellular signals (e.g., nutrient and energy levels and redox status) with extracellular signals (e.g., hormones and growth factors), is a central regulator of cell metabolism.

15.2 AMINO ACID CATABOLISM

The catabolism of the amino acids usually begins with removal of the amino group. Amino groups can then be disposed of in urea synthesis (pp. 576–80). The resulting carbon skeletons are then degraded to form one or more of seven possible metabolic products: acetyl-CoA, acetoacetyl-CoA, pyruvate, α-ketoglutarate, succinyl-CoA, fumarate, or OAA. Depending on the animal's current metabolic requirements, these molecules are used to synthesize fatty acids or glucose or to generate energy. Amino acids degraded to form acetyl-CoA or acetoacetyl-CoA are referred to as **ketogenic** because they can be converted to either fatty acids or ketone bodies. The carbon skeletons of the **glucogenic** amino acids, which are degraded to pyruvate or a citric acid cycle intermediate, can then be used in gluconeogenesis. Several amino acids are both glucogenic and ketogenic (**Table 15.2**). Most amino acids are glucogenic. Discussions of deamination pathways and urea synthesis are followed by descriptions of the pathways that degrade carbon skeletons.

Deamination

The removal of the α-amino group from α-amino acids involves two types of biochemical reaction: transamination and oxidative deamination, both of which were described in Section 14.2. Because these reactions are reversible, amino groups are easily shifted from abundant amino acids and used to synthesize those that are scarce. Amino groups become available for urea synthesis when amino

acids are in excess. Urea is synthesized in especially large amounts when the diet is high in protein or when there is massive breakdown of protein, for example, during starvation.

In muscle, excess amino groups are transferred to α-ketoglutarate to form glutamate:

$$\alpha\text{-Ketoglutarate} + \text{L-Amino acid} \rightleftharpoons \text{L-Glutamate} + \alpha\text{-Keto acid} \quad (1)$$

TABLE 15.2 Amino Acid Carbon Skeleton Fate

Glucogenic		Glucogenic or Ketogenic*	Ketogenic
Alanine	Glycine	Isoleucine	Leucine
Arginine	Histidine	Phenylalanine	Lysine
Asparagine	Methionine	Threonine	
Aspartate	Proline	Tryptophan	
Cysteine	Serine	Tyrosine	
Glutamate	Valine		
Glutamine			

* These amino acids may generate both glucogenic and ketogenic products, or they are degraded by different pathways in different cells or organisms.

The amino groups of glutamate molecules are transported in blood to the liver by the glucose-alanine cycle (**Figure 8.13**):

$$\text{Pyruvate} + \text{L-Glutamate} \rightleftharpoons \text{L-Alanine} + \alpha\text{-Ketoglutarate} \quad (2)$$

In the liver, glutamate is formed as the reaction catalyzed by alanine transaminase is reversed. The oxidative deamination of glutamate yields α-ketoglutarate and NH_4^+. (The reverse of this reaction is illustrated on p. 536.)

Most ammonia molecules produced by extrahepatic tissues is carried to the liver as the amide group of glutamine. The ATP-requiring reaction in which glutamate is converted to glutamine is catalyzed by glutamine synthetase:

$$\text{L-Glutamate} + NH_4^+ + ATP \rightarrow \text{L-Glutamine} + ADP + P_i \quad (3)$$

After its transport to the liver, glutamine is hydrolyzed by glutaminase to form glutamate and NH_4^+. Additional NH_4^+ is generated as glutamate dehydrogenase converts glutamate to α-ketoglutarate:

$$\text{L-Glutamine} + H_2O \rightarrow \text{L-Glutamate} + NH_4^+ \quad (4)$$

$$\text{L-Glutamate} + H_2O + NAD^+ \rightarrow \alpha\text{-Ketoglutarate} + NADH + H^+ + NH_4^+ \quad (5)$$

KEY CONCEPTS

- The degradation of most amino acids begins with removal of the α-amino group.

- Two types of biochemical reaction are involved in amino group removal: transamination and oxidative deamination.

Most of the ammonia generated in amino acid degradation is produced by the oxidative deamination of glutamate. Additional ammonia is produced in several other reactions catalyzed by the following enzymes.

The L-*amino acid oxidases* are flavin mononucleotide (FMN)-requiring liver and kidney enzymes that convert some of the amino acids to α-keto acids, NH_4^+, and H_2O_2. The *serine* and *threonine dehydratases* are hepatic pyridoxal-requiring enzymes that convert serine and threonine to pyruvate and α-ketobutyrate, respectively. Large quantities of ammonia are produced by intestinal bacterial *urease*, which hydrolyzes urea circulating in the bloodstream. Afterward, ammonia diffuses into the blood and is transported to the liver. *Adenosine deaminase* (p. 590) releases NH_4^+ from the adenine ring of AMP in a nucleotide catabolic pathway.

Urea Synthesis

Approximately 90% of surplus nitrogen in ureotelic organisms is converted to urea. As shown in **Figure 15.3**, urea is formed from ammonia, CO_2, and aspartate in a cyclic pathway referred to as the **urea cycle**. The urea cycle is often referred to as the **Krebs urea cycle** or the *Krebs–Henseleit cycle* because it was discovered by Hans Krebs and Kurt Henseleit. The overall equation for urea synthesis is

$$\begin{aligned} &CO_2 + NH_4^+ + \text{Aspartate} + 3\ ATP + 2\ H_2O \rightarrow \\ &\text{Urea} + \text{Fumarate} + 2\ ADP + 2\ P_i + AMP + PP_i + 5\ H^+ \end{aligned} \qquad (6)$$

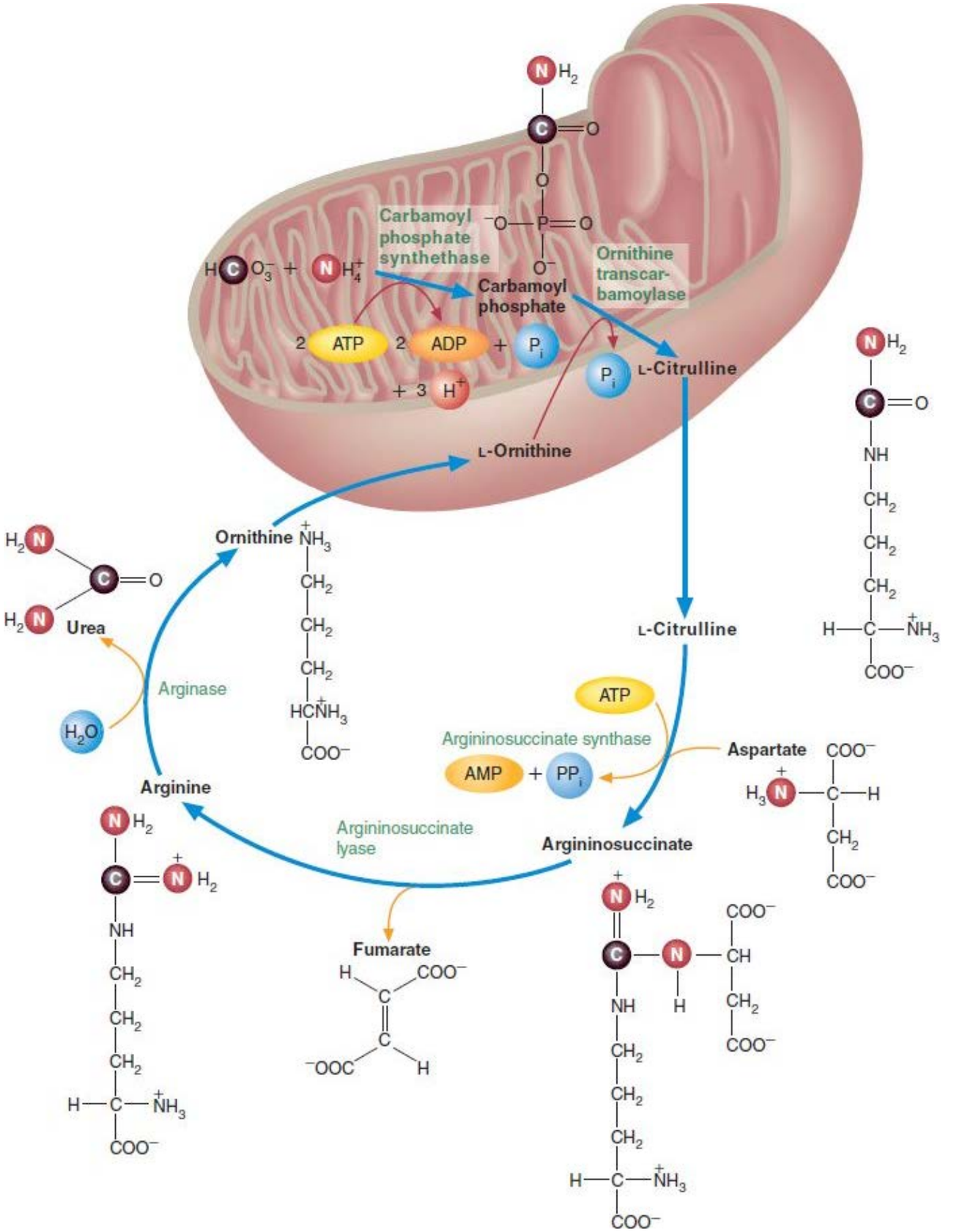
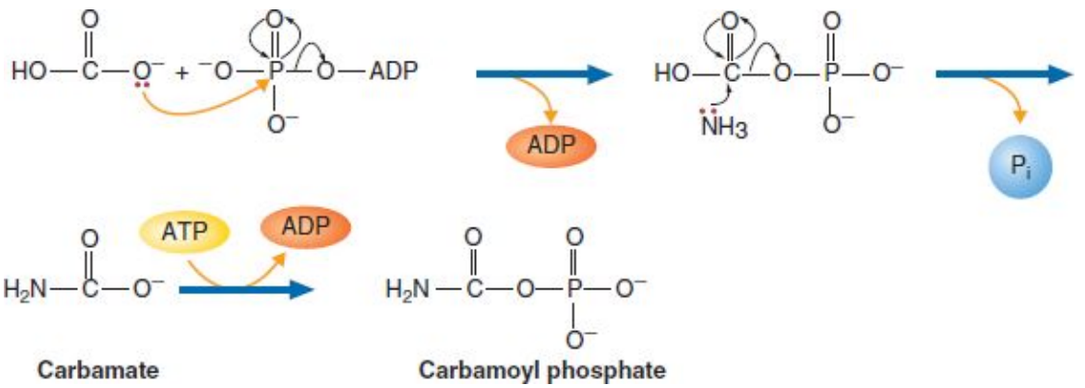


FIGURE 15.3

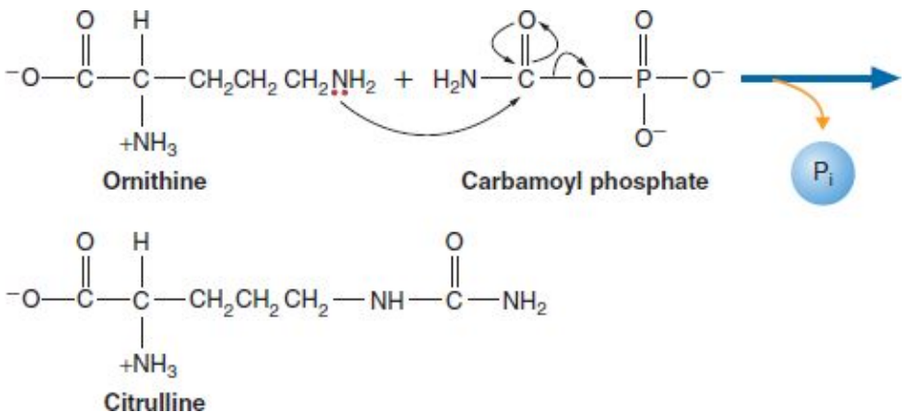
The Urea Cycle

The urea cycle converts NH_4^+ to urea, a less toxic molecule. The sources of the atoms in urea are shown in color. Citrulline is transported across the inner membrane by a carrier for neutral amino acids. Ornithine is transported by ornithine translocase in exchange for H^+ and citrulline. Fumarate is transported back into the mitochondrial matrix (for reconversion to malate) by carriers for α-ketoglutarate or tricarboxylic acids.

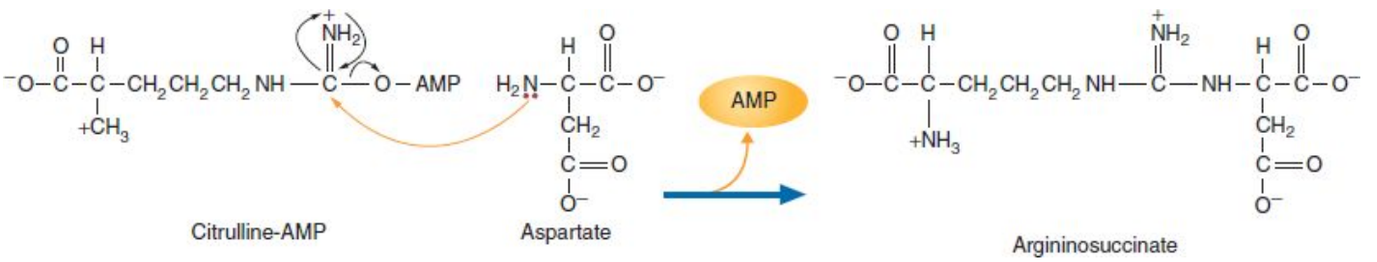
Urea synthesis, which occurs in hepatocytes, begins with the formation of carbamoyl phosphate in the matrix of mitochondria. The substrates for this reaction, catalyzed by carbamoyl phosphate synthetase (CPSI), are NH_4^+ and HCO_3^-.



Carbamoyl phosphate synthesis is essentially irreversible because two molecules of ATP are consumed. (One is used to activate HCO_3^-; the other is used to phosphorylate carbamate.) Carbamoyl phosphate subsequently reacts with ornithine to form citrulline. Citrulline is synthesized in a nucleophilic acyl substitution reaction in which the side chain amino group of ornithine is the nucleophile and phosphate is the leaving group.

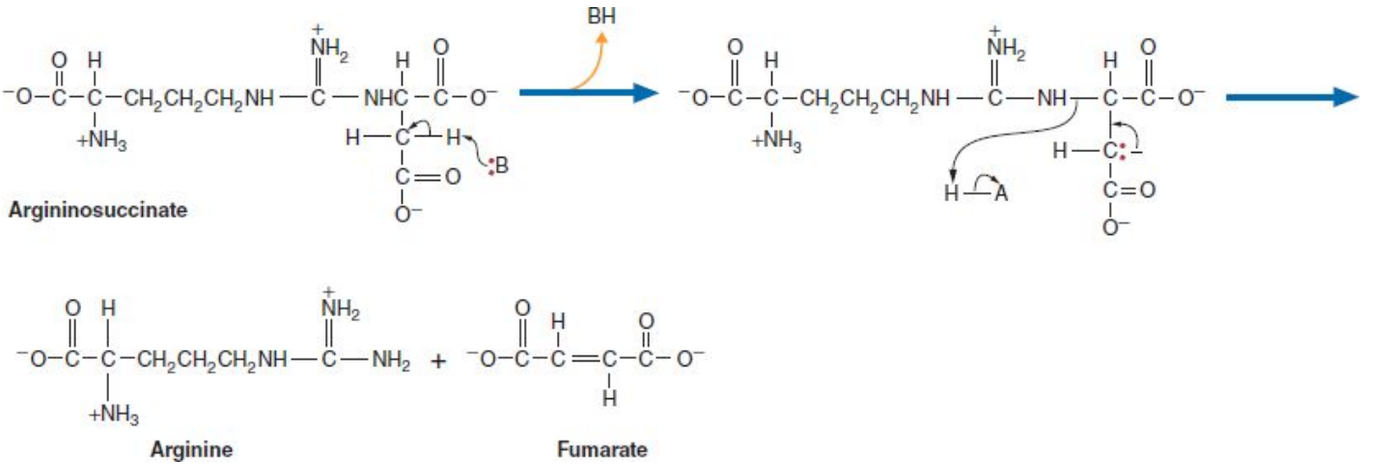


This reaction, catalyzed by ornithine transcarbamoylase, is driven to completion because phosphate is released from carbamoyl phosphate. (Recall from **Table 4.1** that carbamoyl phosphate has a high phosphoryl group transfer potential.) Once formed, citrulline is transported to the cytoplasm, where it reacts with aspartate to form argininosuccinate. (The α-amino group of aspartate, formed from OAA by transamination reactions in the liver, provides the second nitrogen in the urea molecule.) In this reaction, which is catalyzed by argininosuccinate synthase, citrulline is activated by reacting with ATP to form a citrulline–AMP intermediate and pyrophosphate. The amino nitrogen of aspartate, acting as a nucleophile, displaces the AMP to form argininosuccinate.



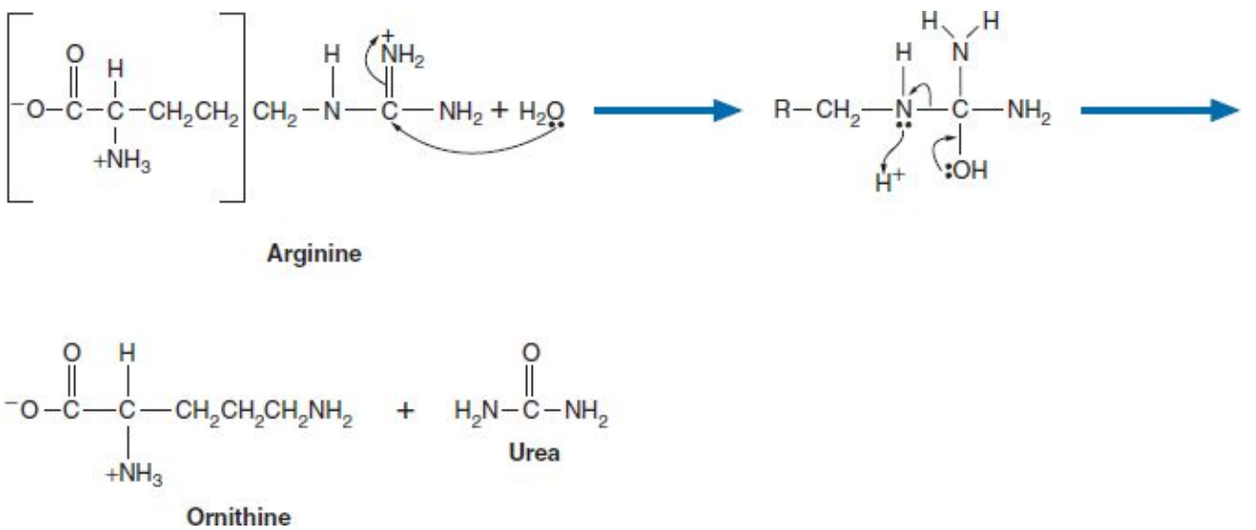
This acyl substitution reaction is driven forward by the cleavage of pyrophosphate catalyzed by

pyrophosphatase. Argininosuccinate lyase subsequently cleaves argininosuccinate to form arginine (the immediate precursor of urea) and fumarate.



A histidine residue within the enzyme's active site, acting as a base (B:), removes a proton from the substrate to form a carbanion. The carbanion then expels the nitrogen to form a carbon–carbon double bond. The nitrogen accepts a proton from a proton donor (HA) (perhaps the protonated histidine).

In the final reaction of the urea cycle, arginase catalyzes the hydrolysis of arginine to form ornithine and urea.



Once it forms, urea diffuses out of the hepatocytes and into the bloodstream. It is ultimately eliminated in urine by the kidney. Ornithine returns to the mitochondria for condensation with carbamoyl phosphate to begin the cycle again. Because arginase is found in significant amounts only in the ureotelic animal liver, urea is produced only in this organ.

After its transport back into the mitochondrial matrix, fumarate is hydrogenated to form malate, a component of the citric acid cycle. The OAA product of the citric acid cycle can be used in energy generation, or it can be converted to glucose or aspartate. The relationship between the urea cycle and the citric acid cycle, often referred to as the **Krebs bicycle**, is outlined in **Figure 15.4**.

Hyperammonemia, a potentially fatal condition in which blood levels of NH_4^+ become excessive when the liver's capacity to synthesize urea is compromised, is discussed in the online Biochemistry in Perspective essay on that topic.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on hyperammonemia.

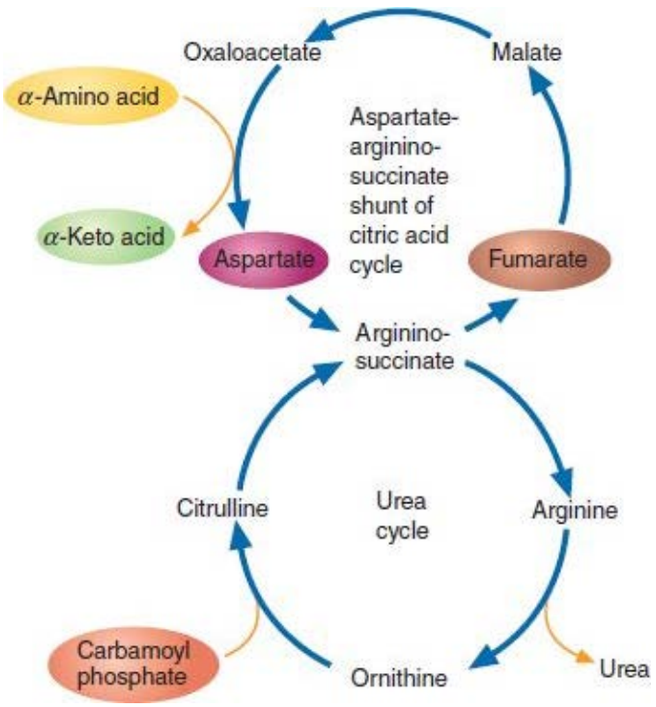


FIGURE 15.4

The Krebs Bicycle

The aspartate used in urea synthesis is generated from oxaloacetate, a citric acid cycle intermediate. This transamination reaction removes amino nitrogen from many amino acids.

WORKED PROBLEM 15.1

Review the urea cycle and then determine the number of ATP molecules used to synthesize one urea molecule.

SOLUTION

Two ATP are required for the synthesis of carbamoyl phosphate from NH_4^+ and CO_2. The synthesis of argininosuccinate involves the conversion of one ATP to an AMP product. Two ATP equivalents are required to convert AMP to ATP. The total number of ATP equivalents used to synthesize one urea molecule is therefore four. ■

Control of the Urea Cycle

Urea is a toxic molecule. Its synthesis is, therefore, stringently regulated. There are long- and short-term regulatory mechanisms. The urea cycle enzymes are controlled in the short term by the concentrations of their substrates. For example, urea synthesis is stimulated by a high-protein diet or by fasting. Carbamoyl phosphate synthetase I (CPSI) is also allosterically activated by *N*-

acetylglutamate. This last-named molecule is a sensitive indicator of the cell's concentration of glutamate, a source of NH_4^+. *N*-Acetylglutamate (NAG) is produced from glutamate and acetyl-CoA in a reaction catalyzed by *N*-acetylglutamate synthetase, which is allosterically activated by arginine. CPSI activation by NAG is a positive-feedback regulatory process because an increase in arginine concentration results in an increase in NAG synthesis. Substrate channeling (p. 232) also enhances the efficiency of the urea cycle. Of all the urea cycle metabolites, only urea, the product of the pathway, has been observed to mix freely with other cytoplasmic metabolites. Long-term regulation is affected by variations in dietary protein consumption. Within several days after a significant dietary change, there are twofold to threefold changes in enzyme levels. Several hormones are involved in the altered rates of enzyme synthesis. Glucagon and the glucocorticoids activate the transcription of urea cycle enzymes, whereas insulin represses their synthesis.

KEY CONCEPTS

- Urea is synthesized from ammonia, CO_2, and aspartate.
- The urea cycle is carefully regulated.

QUESTION 15.1

Although arginine is an intermediate in the urea cycle, it is an essential amino acid (EAA) in young animals. Suggest a reason for this phenomenon.

QUESTION 15.2

In some clinical circumstances, patients with hyperammonemia are treated with antibiotics. Suggest a rational basis for this therapy.

Catabolism of Amino Acid Carbon Skeletons

The α-amino acids can be grouped into classes according to their end products: acetyl-CoA, acetoacetyl-CoA, pyruvate, and several citric acid cycle intermediates. Each group is briefly discussed. The major degradation pathways for the 20 α-amino acids found in proteins are outlined in **Figure 15.5**.

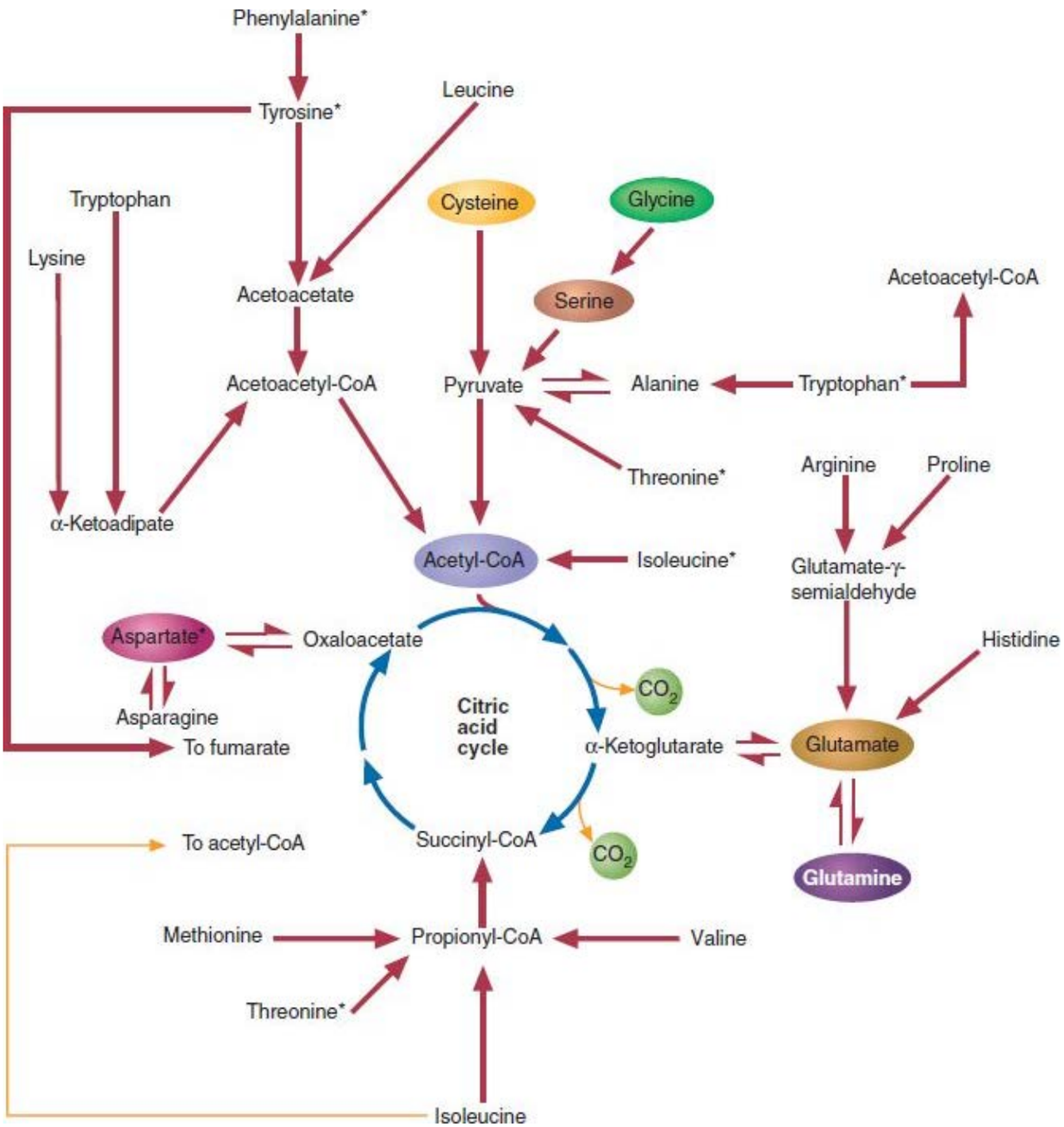


FIGURE 15.5

Degradation of the 20 α-Amino Acids Found in Proteins

The α-amino groups are removed early in the catabolic pathways. Carbon skeletons are converted to common metabolic intermediates. An asterisk marks the amino acids with more than one degradative pathway. Note that some amino acids can be classified as ketogenic or glucogenic, depending on which catabolic pathway is used (see **Table 15.1**).

AMINO ACIDS FORMING ACETYL-CoA In all, 10 α-amino acids yield acetyl-CoA. This group is further divided according to whether pyruvate is an intermediate in acetyl-CoA formation. The amino acids whose degradation involves pyruvate are alanine, serine, glycine, cysteine, and threonine. These amino acids can be ketogenic or glucogenic, depending on the relative activities of pyruvate dehydrogenase and pyruvate carboxylase. Depending on a cell's metabolic requirements, pyruvate may be converted to acetyl-CoA to be oxidized or used to synthesize fatty acids, or it can be converted to OAA, which can be diverted into gluconeogenesis. The other five amino acids converted to acetyl-CoA by pathways not involving pyruvate are lysine, tryptophan, tyrosine, phenylalanine, and leucine. The two reaction sequences are outlined in **Figures 15.6** and **15.7**.

The individual catabolic pathways for these molecules are as follows:

1. **Alanine**. Recall that the reversible transamination reaction involving alanine and pyruvate is an important component of the glucose–alanine cycle discussed previously (Section 8.2).
2. **Serine**. As described, serine is converted to pyruvate by the pyridoxal phosphate-requiring enzyme serine dehydratase.
3. **Glycine**. Glycine can be converted to serine by serine hydroxymethyltransferase. (The hydroxymethyl group is donated by N^5,N^{10}-methylene THF as described in Section 14.3.) Then serine is converted to pyruvate, as previously described. Most glycine molecules, however, are degraded by glycine cleavage enzyme to CO_2, NH_4^+, and N^5,N^{10}-methylene THF.

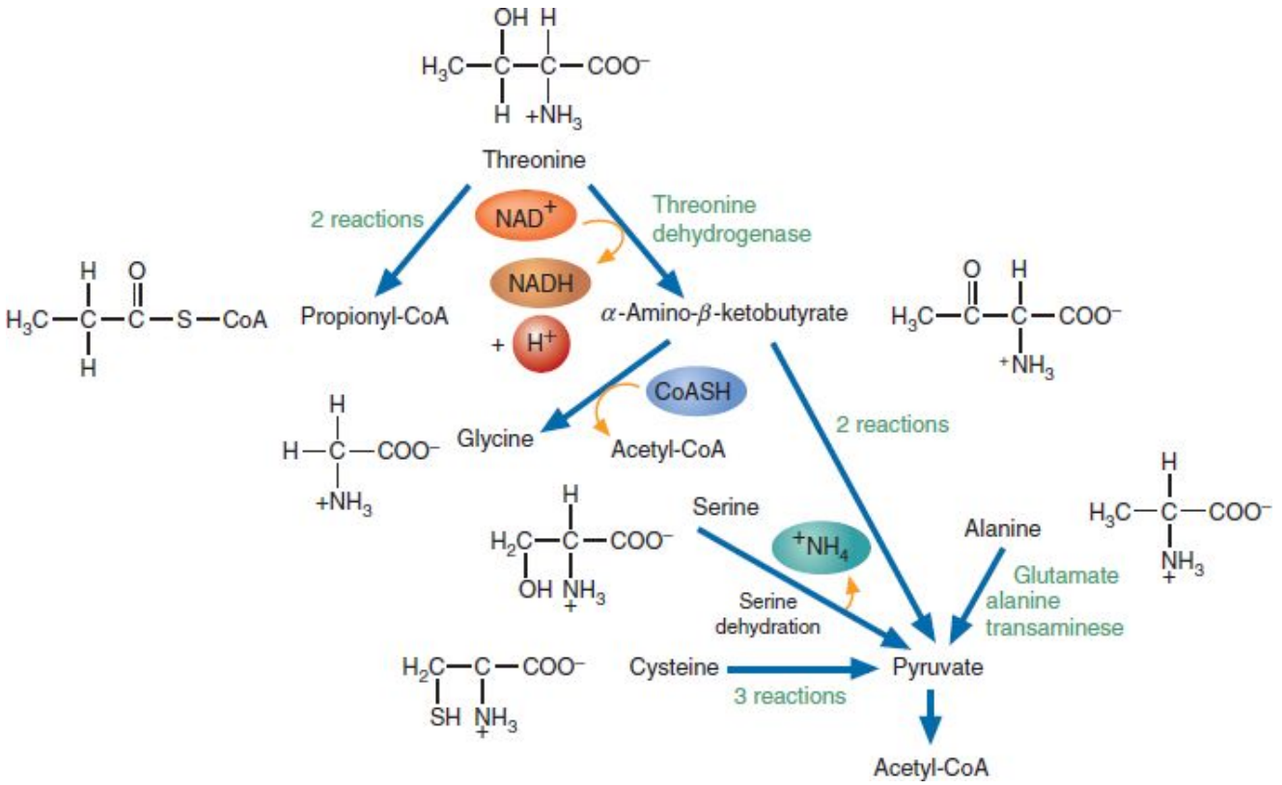


FIGURE 15.6

The Catabolic Pathways of Threonine, Glycine, Serine, Cysteine, and Alanine

Pyruvate is an intermediate in the conversion of these amino acids to acetyl-CoA. Threonine can be degraded by two pathways. In one pathway threonine is converted to pyruvate by a three-reaction pathway. α-amino–β-ketobutyrate, an intermediate in this pathway, can also react with CoASH to yield acetyl-CoA and glycine. Note that glycine is also degraded by glycine cleavage enzyme to form CO_2, NH_4^+, and N^5,N^{10}-methylene THF in an NAD^+-requiring reaction (not shown). In primates, most threonine molecules are degraded to propionyl-CoA, a molecule, which is then converted to the citric acid cycle intermediate succinyl-CoA (**Figure 12.12**). Serine is converted to pyruvate with the release of NH_4^+ by serine dehydratase. The major pathway for cysteine catabolism is a three-reaction pathway in which the oxidized intermediate cysteine sulfinate undergoes a transamination and then a desulfuration to yield pyruvate and bisulfite (HSO_3^-). A transamination reaction converts alanine to pyruvate.

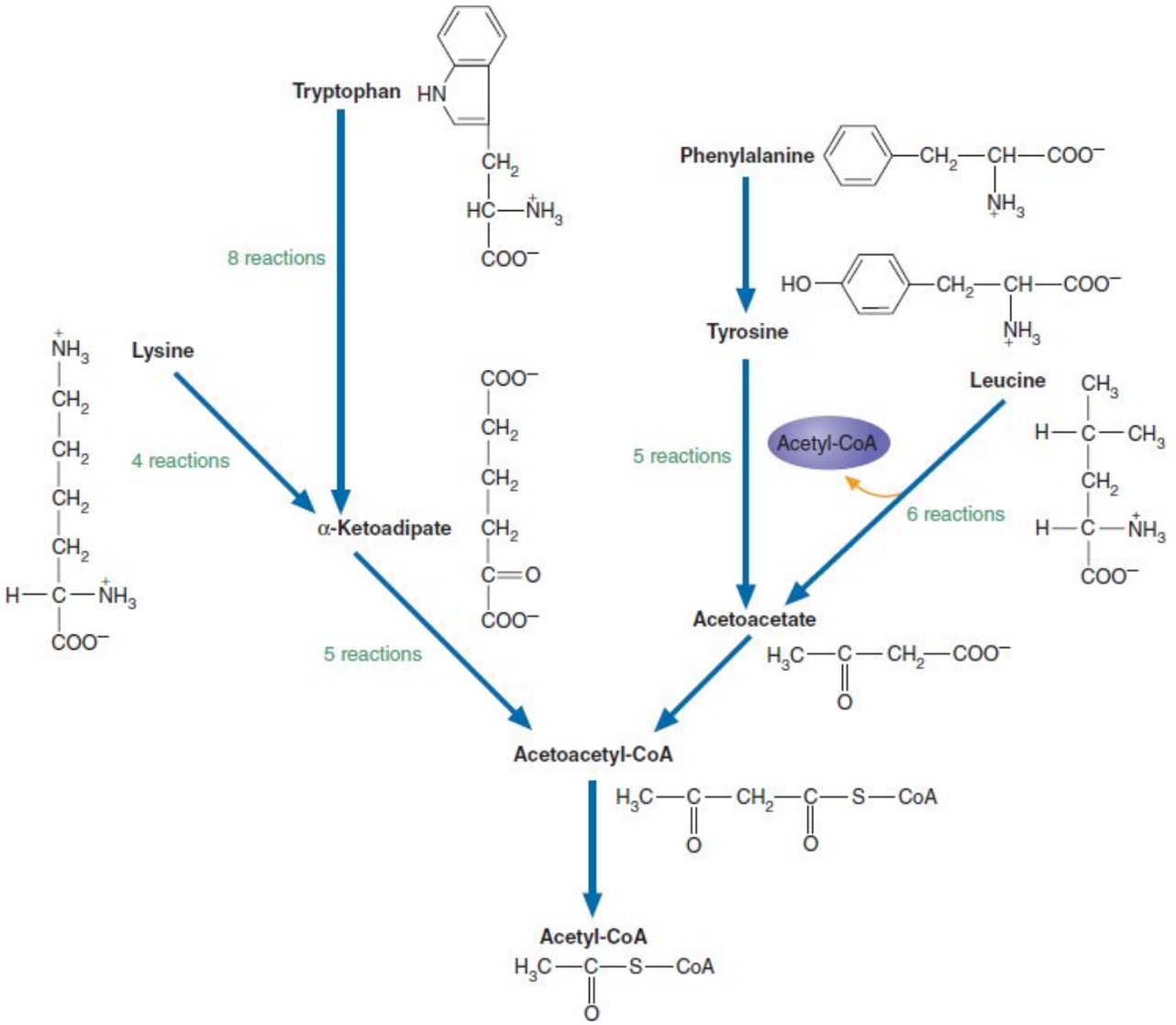


FIGURE 15.7

The Catabolic Pathways of Lysine, Tryptophan, Phenylalanine, Tyrosine, and Leucine

These pathways are long and complicated. The number of reactions in each segment is indicated.

4. **Cysteine**. In animals, cysteine is converted to pyruvate by several pathways. In the principal pathway, cysteine is oxidized to cysteine sulfinate. Pyruvate is then produced after a transamination and a desulfuration reaction.
5. **Threonine**. In the major degradative pathway, threonine is oxidized by threonine dehydrogenase to form *α*-amino–*β*-ketobutyrate. The latter molecule is metabolized further to form pyruvate, or it can be cleaved by *α*-amino–*β*-ketobutyrate lyase to form acetyl-CoA and glycine. Alternatively, threonine can be degraded to *α*-ketobutyrate by threonine dehydratase and subsequently to propionyl-CoA. Propionyl-CoA is then converted to succinyl-CoA (see p. 232).
6. **Lysine**. Lysine is converted to *α*-ketoadipate in a series of reactions that include two oxidations, removal of the side chain amino group, and a transamination. Acetoacetyl-CoA is produced in a reaction pathway that includes several oxidations, a decarboxylation, and a hydration. Acetoacetyl-CoA can be converted to acetyl-CoA in a reaction that is the reverse of a step in ketone body formation (p. 457).
7. **Tryptophan**. Tryptophan is converted to *α*-ketoadipate in eight reactions, which also yield formate and alanine. The alanine produced in this pathway is converted to acetyl-CoA via

pyruvate. Acetyl-CoA is synthesized from α-ketoadipate as described for lysine.

8. **Tyrosine**. Tyrosine catabolism begins with a transamination and a dehydroxylation. Homogentisate is synthesized in the latter reaction, catalyzed by the ascorbate-requiring enzyme parahydroxyphenylpyruvate dioxygenase. Homogentisate is converted to maleylacetoacetate by homogentisate oxidase. Acetoacetate and fumarate are then generated in isomerization and hydration reactions.

9. **Phenylalanine**. Phenylalanine is hydroxylated to form tyrosine by phenylalanine-4-monooxygenase (**Figure 15.8**) in a reaction requiring O_2 and tetrahydrobiopterin (BH_4) (**Figure 15.9**). Tyrosine is degraded to form acetoacetate and fumarate.

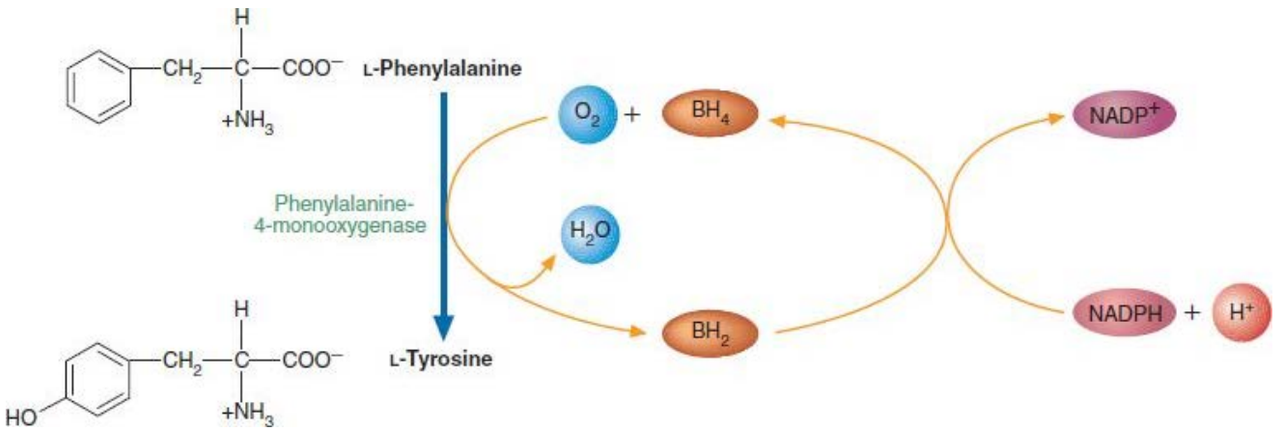


FIGURE 15.8

The Conversion of Phenylalanine to Tyrosine

The reaction catalyzed by phenylalanine-4-monooxygenase is irreversible. The electrons required for the hydroxylation of phenylalanine are carried to O_2 from NADPH by tetrahydrobiopterin (BH_4).

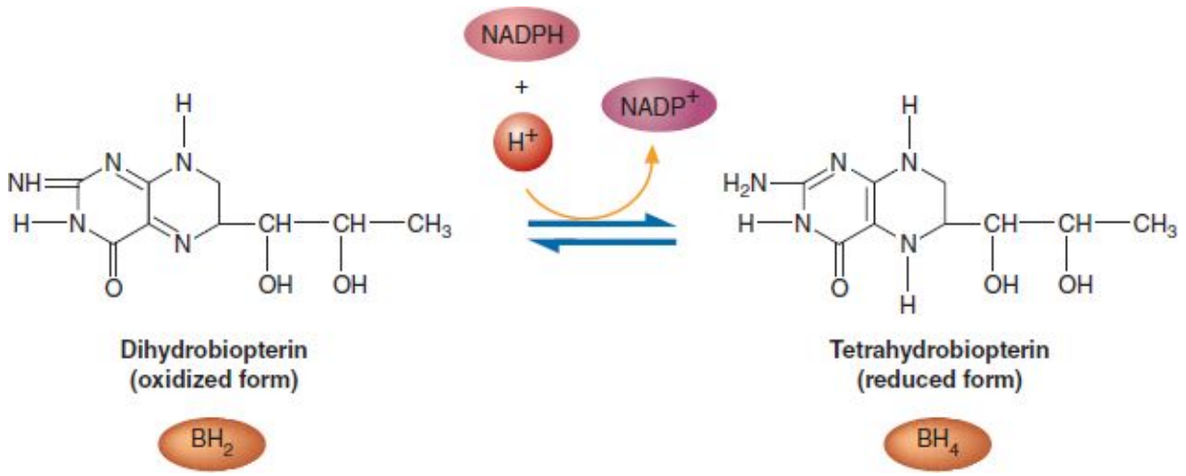


FIGURE 15.9

Tetrahydrobiopterin

Tetrahydrobiopterin (BH_4), derived from GTP, is an essential cofactor in the hydoxylation of phenylalanine to form tyrosine and in the biosynthesis of several neurotransmitters (the catecholamines and serotonin) as well as melatonin (p. 549) and nitric oxide (p. 552). BH_4 (the reduced form) is regenerated from dihydrobiopterin (BH_2) (the oxidized form) by reduction with NADPH.

10. **Leucine**. Leucine, one of the branched-chain amino acids, is converted to HMG-CoA in a series of reactions that include a transamination, two oxidations, a carboxylation, and a

hydration. HMG-CoA is then converted to acetyl-CoA and acetoacetate by HMG-CoA lyase.

AMINO ACIDS FORMING α-KETOGLUTARATE Five amino acids (glutamate, glutamine, arginine, proline, and histidine) are degraded to α-ketoglutarate. An outline of their catabolism is illustrated in **Figure 15.10**. Each pathway is briefly described.

1. **Glutamate and glutamine**. Glutamine is converted to glutamate and NH_4^+ by glutaminase. As described previously, glutamate is converted to α-ketoglutarate by glutamate dehydrogenase or by transamination.
2. **Arginine**. Recall that arginine is cleaved by arginase to form ornithine and urea. In a subsequent transamination reaction, ornithine is converted to glutamate-γ-semialdehyde. Glutamate is then produced as glutamate-γ-semialdehyde is hydrated and oxidized. α-Ketoglutarate is produced by a transamination reaction or by oxidative deamination.

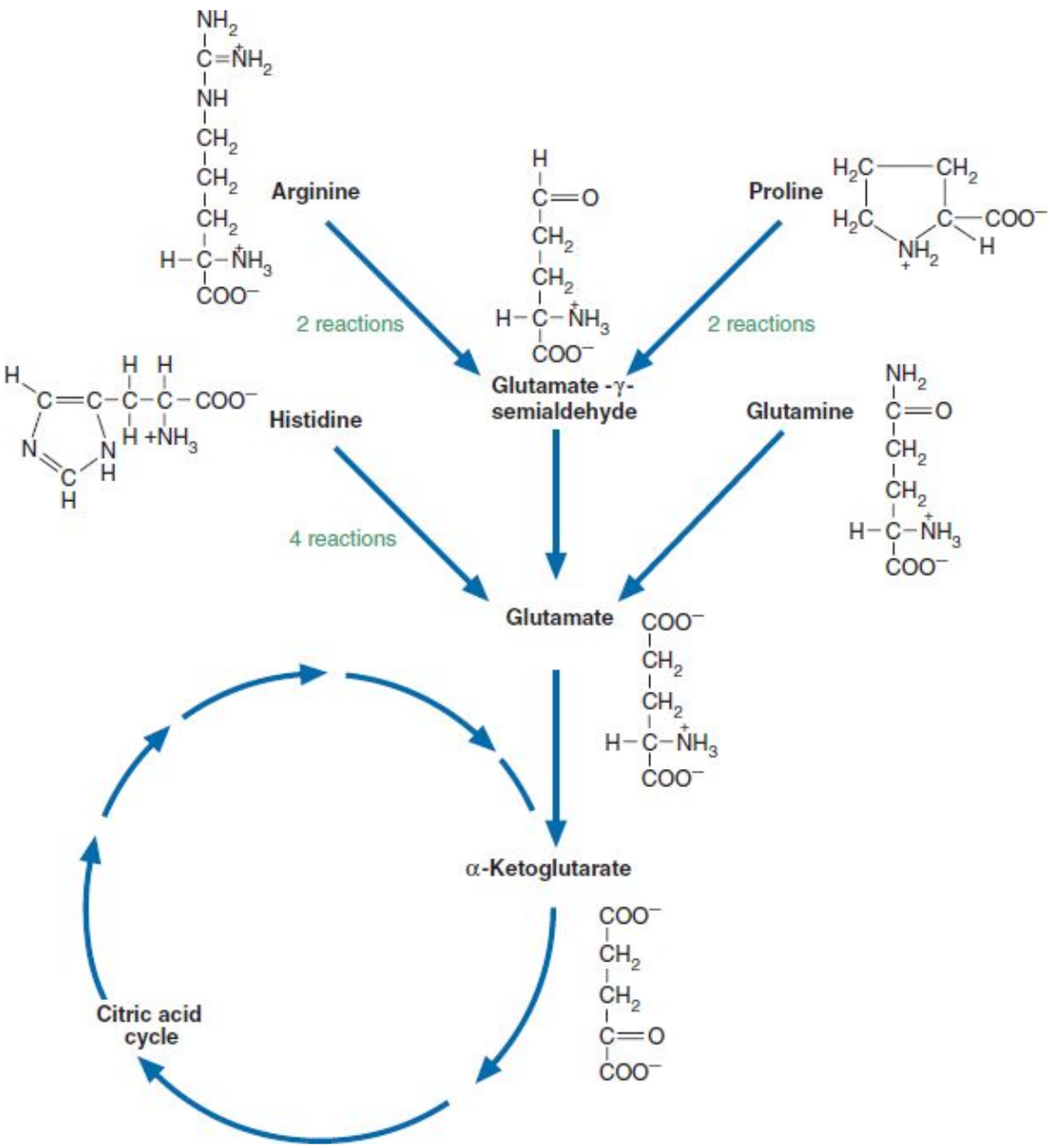


FIGURE 15.10

The Catabolic Pathways of Glutamate, Glutamine, Arginine, Proline, and Histidine

All these amino acids are eventually converted to α-ketoglutarate.

3. **Proline**. Proline catabolism begins with an oxidation reaction that produces Δ^1-pyrroline. The latter molecule is converted to glutamate-γ-semialdehyde by a hydration reaction. Glutamate is then formed by another oxidation reaction.
4. **Histidine**. Histidine is converted to glutamate in four reactions: a nonoxidative deamination, two hydrations, and the removal of a formamino group (NH═CH—) by THF.

AMINO ACIDS FORMING SUCCINYL-CoA Succinyl-CoA is formed from the carbon skeletons of methionine, isoleucine, valine, and threonine (by one of its degradative pathways as already discussed). **Figure 15.11** outlines the reactions that degrade the first three of these amino acids.

1. **Methionine**. Methionine degradation begins with the formation of SAM, which is followed by a demethylation reaction, as described (**Figure 14.17**). The product SAH is hydrolyzed to adenosine and homocysteine. Homocysteine is metabolized to yield α-ketobutyrate, cysteine, and NH_4^+. α-Ketobutyrate is then converted to propionyl-CoA by α-keto acid dehydrogenase. Propionyl-CoA is converted to succinyl-CoA in three steps (**Figure 12.12**). The conversion of methionine to cysteine is sometimes referred to as the **transsulfuration pathway** (**Figure 15.12**). A substantial amount of the sulfate produced from cysteine degradation is excreted in urine. Sulfate in the form of 3′-phosphoadenosine–5′phosphosulfate (PAPS) is also used in the synthesis of sulfatides (p. 413) and proteoglycans (p. 266). Additionally, molecules such as steroids and certain drugs are excreted as sulfate esters. Also recall that the gasotransmitter H_2S is synthesized from cysteine in reactions catalyzed by CBS or CSE (p. 553).

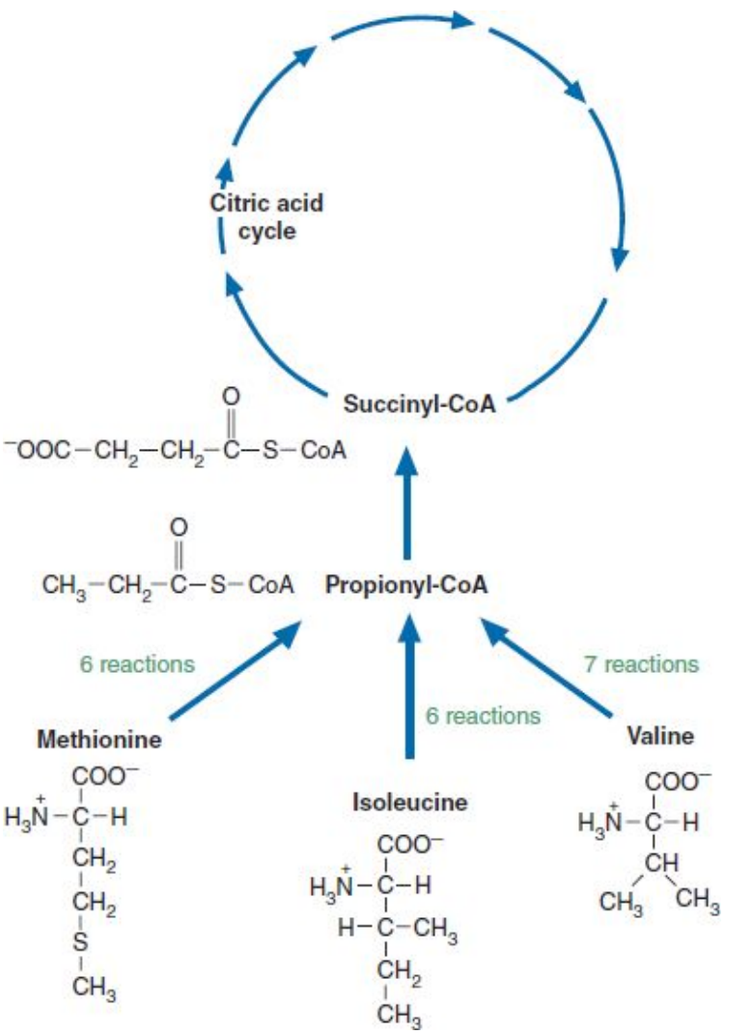


FIGURE 15.11

The Catabolic Pathways of Methionine, Isoleucine, and Valine

Propionyl-CoA is a common intermediate in the degradation of methionine, isoleucine, and valine. Methionine is first converted to homocysteine (**Figure 14.18**), which in turn yields cysteine and α-ketobutyrate (**Figure 14.11**). α-Ketobutyrate is decarboxylated to generate propionyl-CoA. Products of the conversion of isoleucine to propionyl-CoA include acetyl-CoA, three NADH, and one CO_2. Valine degradation products include three NADH, one $FADH_2$, and two CO_2. The conversion of propionyl-CoA to succinyl-CoA, a citric acid cycle intermediate, is outlined in **Figure 12.12**. Note that threonine is also degraded via the propionyl-CoA/succinyl-CoA pathway (see **Figure 15.6**).

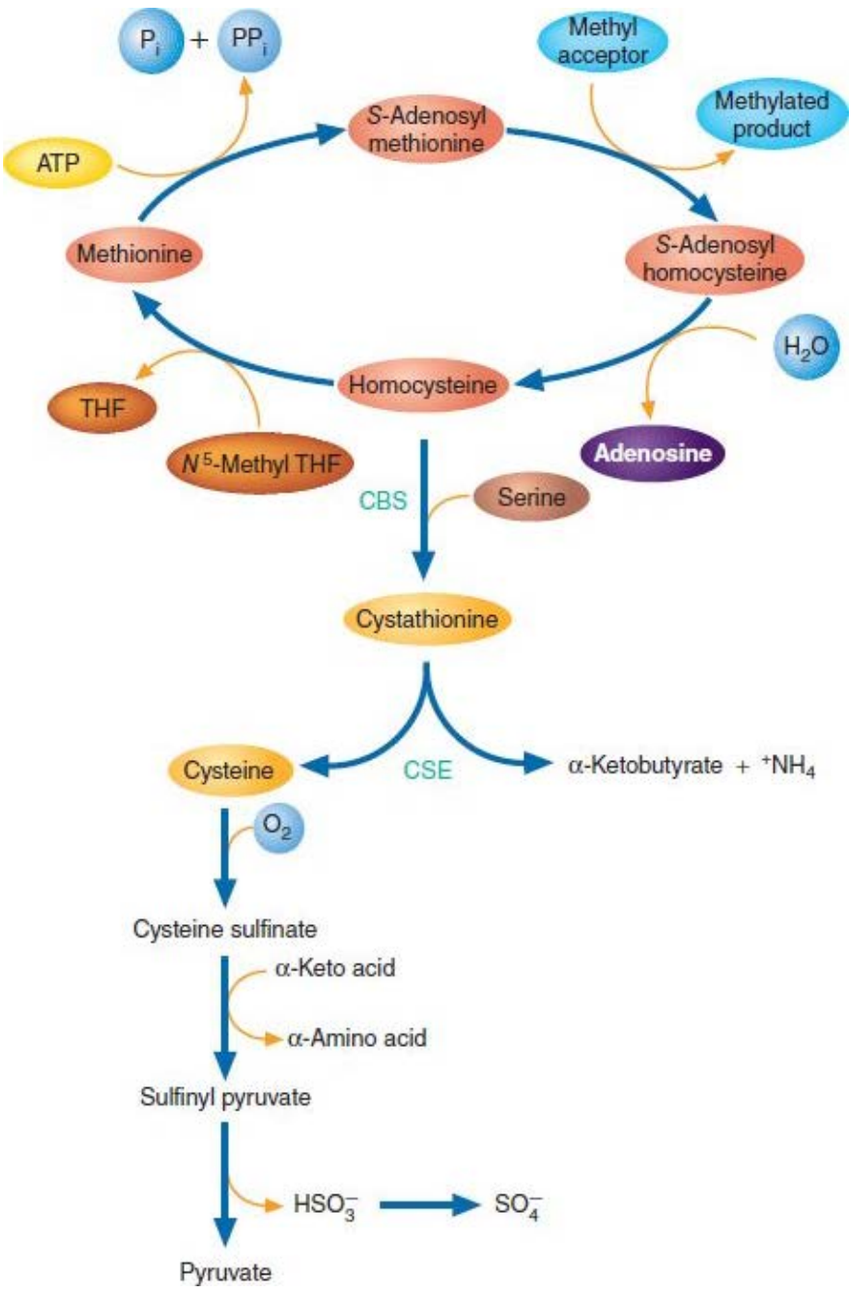


FIGURE 15.12

The Transsulfuration Pathway

The sulfur atom of methionine becomes the sulfur atom of cysteine in two reactions. Cystathionine β-synthase (CBS) converts homocysteine and serine into cystathionine. (CBS activity is depressed by cysteine when the latter molecule's cellular concentration is high.) Cystathionine is converted to cysteine, α-ketobutyrate, and NH_4^+ by γ-cystathionase (CSE). Cysteine may then be incorporated into glutathione (the chief regulator of cellular redox homeostasis), coenzyme A, or proteins; or it may be oxidized by cysteine dioxygenase to form cysteine sulfinate. Cysteine sulfinate can undergo a transamination reaction

followed by desulfuration to yield pyruvate and sulfite (HSO_3^-). Sulfite is subsequently converted to sulfate (SO_4^-) by sulfite oxidase. The sulfate generated in cysteine catabolism is excreted or used in several biosynthetic or catabolic pathways. Note that the transsulfuration and methylation pathways are intimately related.

2. **Isoleucine and valine**. The first four reactions in the degradation of isoleucine and valine are catalyzed by the same four enzymes (**Figure 15.13**). Several reactions in both pathways are similar to the β-oxidation reactions that yield NADH and $FADH_2$. The products of the isoleucine pathway are acetyl-CoA and propionyl-CoA, which is subsequently converted to succinyl-CoA. Hence, isoleucine is both a ketogenic and a glucogenic amino acid. The degradative pathway of valine is similar but yields only succinyl-CoA. Valine is, therefore, a glucogenic amino acid. Numerous tissues can use valine, isoleucine, and leucine, the branched-chain amino acids (BCAA; p. 532), to generate energy. Most BCAA oxidation, however, occurs in skeletal muscle during exercise. BCAA are an important source of energy because, in addition to their high concentration in muscle protein, BCAA degradation generates NADH and $FADH_2$ and their end products (acetyl-CoA, succinyl-CoA, and acetoacetate) are oxidized by the citric acid cycle.

AMINO ACIDS FORMING OXALOACETATE Both aspartate and asparagine are degraded to form OAA. Aspartate is converted to OAA with a single transamination reaction. Asparagine is initially hydrolyzed to yield aspartate and NH_4^+ by asparaginase.

KEY CONCEPTS

Amino acid carbon skeletons can be degraded into one or more of several metabolites. These include acetyl-CoA, acetoacetyl-CoA, α-ketoglutarate, succinyl-CoA, and oxaloacetate.

QUESTION 15.3

Taurine (2-aminoethane sulfonic acid) is a sulfur-containing amine synthesized from cysteine. Although taurine is present in high concentrations in mammalian cells, except for its incorporation in bile salts, the physiological role of this amine is still poorly understood. However, several pieces of information suggest that taurine is an important metabolite. For example, taurine is found in brain tissue in large amounts. In addition, domestic cats have been observed to develop congestive heart failure if fed a taurine-free diet. (Cats cannot synthesize taurine. For this reason, they must consume meat in their diet. Cats that are fed vegetarian diets soon become listless and will die prematurely.) In most animals, taurine is synthesized from cysteine sulfinate (the oxidation product of cysteine) in two reactions: a decarboxylation followed by an oxidation of the sulfinate group (SO_2^-) to form sulfonate (SO_3^-). With this information, determine the biosynthetic pathway for taurine. [*Hint*: Refer to **Figure 15.12**.]

QUESTION 15.4

Taurine is the most abundant amino acid in white blood cells, where it reacts with HOCl, which is synthesized during the respiratory burst in a reaction catalyzed by myeloperoxidase. The product of this reaction, taurine monochloramine (Tau-Cl), is relatively nontoxic and stable compared with HOCl. Tau-Cl modulates the inflammation process by downregulating the production of $NO^\bullet$ and

proinflammatory proteins such as tumor necrosis factor α (TNF-α). Provide the reactions in which HOCl and Tau-Cl are produced. [*Hint*: Review the respiratory burst in **Figure 10.24**.]

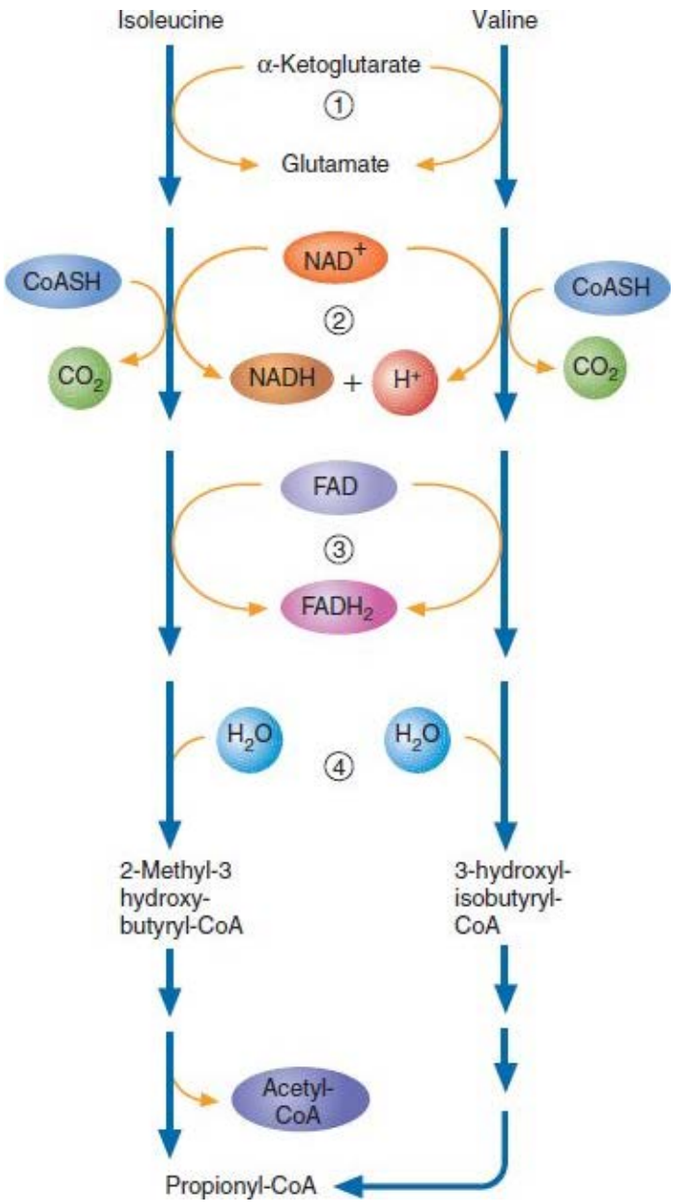


FIGURE 15.13

Degradation of Isoleucine and Valine

The degradation of both isoleucine and valine begins with the same four reactions: a transamination catalyzed by *branched-chain amino acid transaminase* (enzyme 1), an oxidative decarboxylation catalyzed by *branched-chain α-keto acid dehydrogenase* (enzyme 2), an oxidation catalyzed by an FAD-requiring acyl-CoA dehydrogenase (enzyme 3), and a hydration catalyzed by *enoyl-CoA hydratase* (enzyme 4). Isoleucine degradation continues in three reactions to yield acetyl-CoA and propionyl-CoA. The last-named molecule is then converted to succinyl-CoA. Valine degradation continues in four reactions to yield propionyl-CoA.

Biochemistry IN PERSPECTIVE

Disorders of Amino Acid Catabolism

What are the effects on human health of deficiency of a single enzyme in amino acid metabolism?

Defects in amino acid catabolism were among the first genetic diseases to be recognized and investigated by medical scientists. These "inborn errors of metabolism" result from **mutations** (permanent changes in genetic information, i.e., DNA structure). Most commonly, in the genetic diseases related to amino acid metabolism, the defective gene codes for an enzyme. The metabolic blockage that results from such a deficit disrupts what are ordinarily highly coordinated cellular and organismal processes, producing abnormal amounts and/or types of metabolites. Because these metabolites (or their heightened concentrations) are often toxic, permanent damage or death ensues. Several of the most commonly observed inborn errors of amino acid metabolism are discussed below.

Alkaptonuria, the first disease to be linked to genetic inheritance involving a single enzyme, is caused by a deficiency of homogentisate oxidase, an enzyme required for the catabolism of the aromatic ring of phenylalanine and tyrosine. In 1902, Archibald Garrod proposed that a single inheritable unit (later called a gene) was responsible for the urine in alkaptonuric patients turning black. Large quantities of homogentisate, the substrate for the defective enzyme, are excreted in urine. Homogentisate turns black when it is oxidized as the urine is exposed to air. Although black urine appears to be an essentially benign (if somewhat disconcerting) condition, alkaptonuria is not innocuous because alkaptonuric patients develop arthritis in later life. In addition, pigment accumulates gradually and unevenly darkens the skin.

In *albinism*, the enzyme tyrosinase is deficient. Consequently, *melanin*, a black pigment found in skin, hair, and eyes and formed from tyrosine in several cell types, is not produced. In such cells, tyrosinase converts tyrosine to L-DOPA (refer to the online reading Synthesis of Catecholamines) and L-DOPA to dopaquinone. A large number of dopaquinone molecules, which are highly reactive, condense to form melanin. As a result of a lack of pigment, affected individuals (called albinos) are extremely sensitive to sunlight. In addition to their susceptibility to skin cancer and sunburn, they often have poor eyesight.

Phenylketonuria (*PKU*), caused by a deficiency of phenylalanine hydroxylase, is one of the most common genetic diseases associated with amino acid metabolism. If this condition is not identified and treated immediately after birth, mental retardation and other forms of irreversible brain damage occur. This damage results mostly from the accumulation of phenylalanine. High phenylalanine blood levels result in the saturation of the transport mechanism for large neutral amino acids across the blood–brain barrier. Brain damage results from decreased levels of protein and neurotransmitter synthesis. When present in excess, phenylalanine undergoes transamination to form phenylpyruvate, which is then converted to phenyllactate and phenylacetate. Large amounts of these molecules are excreted in the urine. Phenylacetate gives the urine its characteristic musty odor. PKU is treated with a low-phenylalanine diet.

Maple syrup urine disease, also called *branched-chain ketoaciduria*, is a disorder in which the α-keto acids derived from the BCAA leucine, isoleucine, and valine accumulate in large quantities in blood. Their presence in urine imparts a characteristic odor that gives the malady its name. All three α-keto acids accumulate because of mutations in any of four genes that code for the subunits of branched-chain α-keto acid dehydrogenase complex. (This enzymatic activity is responsible for the conversion of the α-keto acids to their acyl-CoA derivatives.) If left untreated, affected individuals experience vomiting, convulsions, severe brain damage, and mental retardation. They often die before 1 year of age. As with phenylketonuria, treatment consists of rigid dietary control.

Deficiency of methylmalonyl-CoA mutase (an enzyme involved in the conversion of propionyl-CoA into succinyl-CoA) results in *methylmalonic acidemia*, a condition in which methylmalonate accumulates in blood. The symptoms are similar to those of maple syrup urine disease.

Methylmalonate may also accumulate because of a deficiency of adenosylcobalamin or weak binding of this coenzyme by a defective enzyme. Some affected individuals respond to injections of large daily doses of vitamin B_{12}.

SUMMARY The deficiency in humans of a single enzyme in amino acid metabolism has widespread effects that typically include brain damage.

15.3 NUCLEOTIDE DEGRADATION

In most living organisms, purine and pyrimidine nucleotides are constantly degraded and/or recycled. During digestion, nucleic acids are hydrolyzed to oligonucleotides by enzymes called **nucleases**. (Short nucleic acid segments containing fewer than 50 nucleotides are called **oligonucleotides**.) Enzymes that are specific for breaking internucleotide bonds in DNA are called *deoxyribonucleases* (DNases); those that degrade RNA are called *ribonucleases* (RNases). Once formed, oligonucleotides are further hydrolyzed by various *phosphodiesterases* in a process that produces a mixture of mononucleotides. *Nucleotidases* remove phosphate groups from nucleotides, yielding nucleosides. These latter molecules are hydrolyzed by *nucleosidases* to free bases and ribose or deoxyribose, which are then absorbed.

Generally speaking, dietary purine and pyrimidine bases are not used in significant amounts to synthesize cellular nucleic acids. Instead, they are degraded within enterocytes. Purines are degraded to uric acid in humans and birds. Pyrimidines are degraded to *β*-alanine or *β*-aminoisobutyric acid, as well as NH_3 and CO_2. In contrast to the catabolic processes for other major classes of biomolecules (e.g., sugars, fatty acids, and amino acids), purine and pyrimidine catabolism does not result in ATP synthesis. The major pathways for the degradation of purine and pyrimidine bases are described next.

Purine Catabolism

Purine nucleotide catabolism is outlined in **Figure 15.14**. There is some variation in the specific pathways used by different organisms or tissues to degrade AMP. In most tissues, AMP is hydrolyzed by 5′-nucleotidase to form adenosine. Adenosine is then deaminated by adenosine deaminase (also called adenosine aminohydrolase) to form inosine. In muscle, AMP is initially converted to IMP by AMP deaminase (also referred to as adenylate aminohydrolase). IMP is subsequently hydrolyzed to inosine by 5′-nucleotidase. The AMP deaminase reaction is also a component of the purine nucleotide cycle (**Figure 15.15**). In this pathway, IMP reacts with aspartate to yield adenylosuccinate. This GTP-requiring reaction is catalyzed by adenylosuccinate synthetase. Adenylosuccinate is then converted by adenylosuccinase to AMP and fumarate. The purine nucleotide cycle is a means of converting amino acids (via aspartate) to citric acid cycle intermediates (via fumarate). In skeletal muscle, AMP deaminase activity is exceptionally high. Activation of muscle AMP deaminase (called myoadenylate deaminase) and increased flux through the purine nucleotide cycle occur during intense exercise.

Purine nucleoside phosphorylase converts inosine, guanosine, and xanthosine to hypoxanthine, guanine, and xanthine, respectively. (The ribose-1-phosphate formed during these reactions is reconverted to PRPP (**Figure 14.26**) by ribose-5-phosphate pyrophosphokinase.) Hypoxanthine is oxidized to xanthine by xanthine oxidase, an enzyme that contains molybdenum, FAD, and two different Fe-S clusters. (Xanthine oxidase–catalyzed reactions produce $O_2^{\cdot-}$ in addition to forming H_2O_2.) Guanine is deaminated to xanthine by guanine deaminase (also called guanine aminohydrolase). Xanthine molecules are further oxidized to uric acid by xanthine oxidase. Xanthine

oxidase is inhibited by the drug molecule allopurinol, a structural analogue of hypoxanthine.

Many animals degrade uric acid further (**Figure 15.16**). Urate oxidase converts uric acid to allantoin, an excretory product in many mammals. Allantoinase catalyzes the hydration of allantoin to form allantoate, which is excreted by bony fish. Other fish, as well as amphibians, produce allantoicase, which splits allantoic acid into glyoxylate and urea. Finally, marine invertebrates degrade urea to NH_4^+ and CO_2 in a reaction catalyzed by urease.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on gout.

DISEASES LINKED TO PURINE CATABOLISM Several diseases result from defects in purine catabolic pathways. *Gout*, which is often characterized by high blood levels of uric acid and recurrent attacks of arthritis, is caused by several metabolic abnormalities and treated with allopurinol. Two different immunodeficiency diseases are now known to result from defects in purine catabolic reactions. *Adenosine deaminase deficiency* results in high levels of deoxyadenosine, which is toxic, especially in the T and B lymphocytes (or **T cells** and **B cells**, respectively). Children with adenosine deaminase deficiency usually die before the age of 2 because of massive infections. In *purine nucleoside phosphorylase deficiency*, levels of purine nucleotides are high, and synthesis of uric acid decreases. High levels of dGTP are apparently responsible for the impairment of T cells that is characteristic of this malady. Individuals with *myoadenylate deaminase deficiency* exhibit exercise-induced muscle fatigue.

Allopurinol and Gout

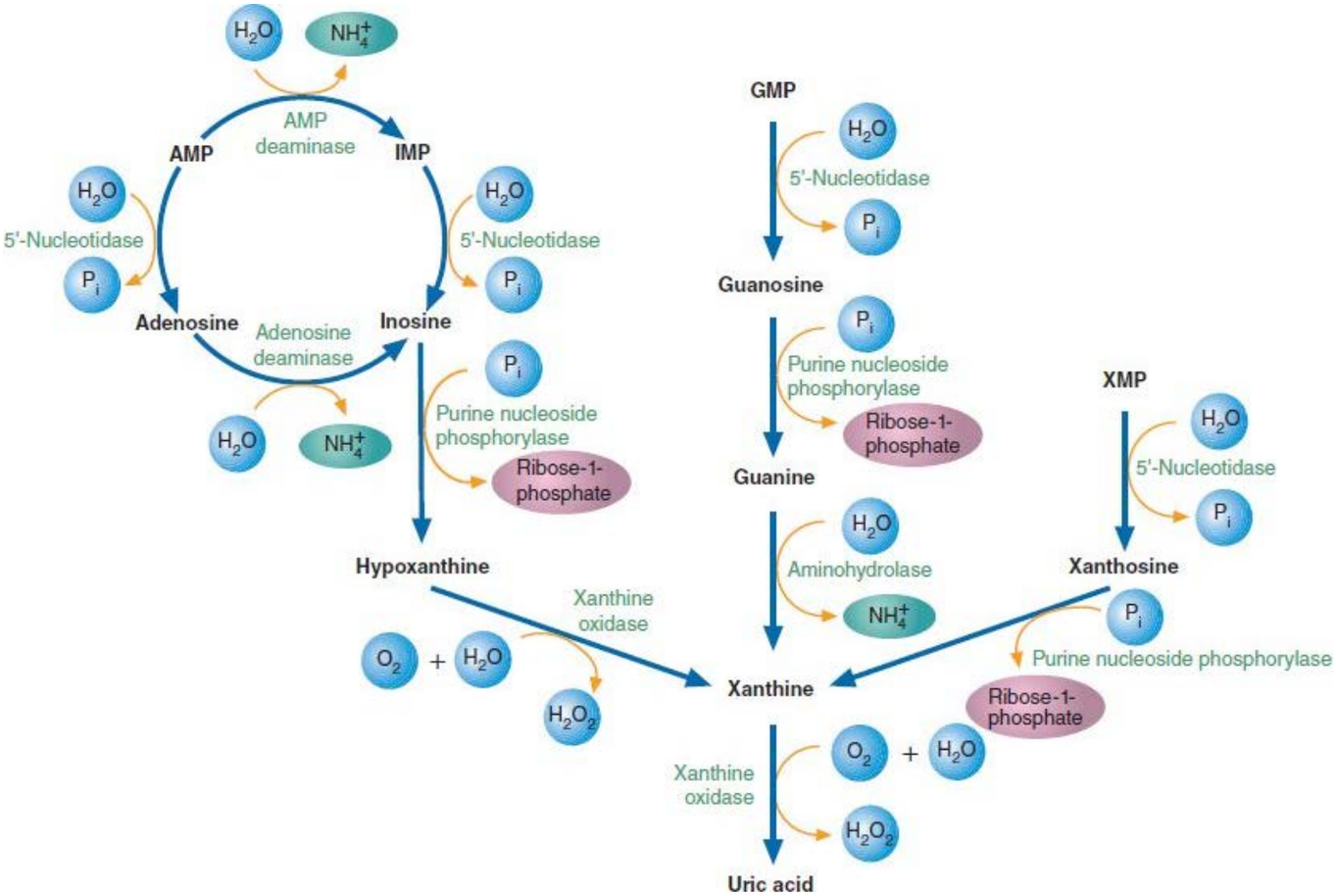


FIGURE 15.14

Purine Nucleotide Catabolism

Ribose-1-phosphate is released in AMP, GMP, and xanthosine monophosphate (XMP) catabolism. Xanthine oxidase-catalyzed reactions generate O_2^-, as well as H_2O_2.

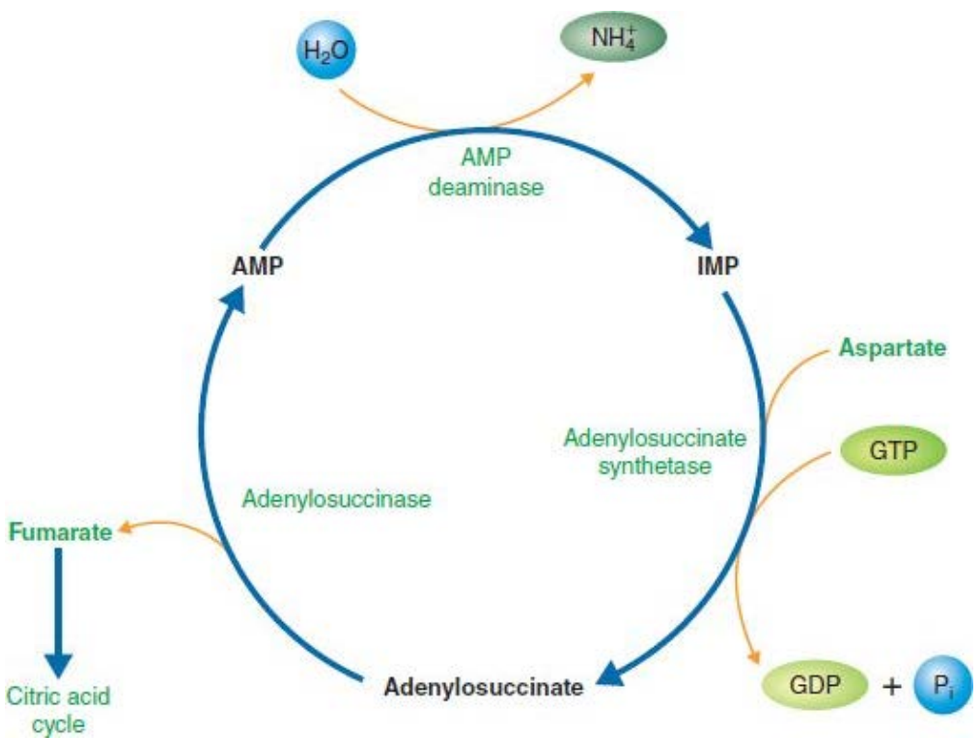


FIGURE 15.15

The Purine Nucleotide Cycle

In skeletal muscle, the purine nucleotide cycle is an anaplerotic process that replenishes citric acid cycle intermediates by producing fumarate from aspartate.

QUESTION 15.5

Unlike primates and birds, many animals possess the enzyme urate oxidase. Suggest a reason why these organisms do not suffer from gout.

Pyrimidine Catabolism

In humans, the purine ring cannot be degraded. This is not true for the pyrimidine ring. An outline of the pathway for pyrimidine nucleotide catabolism is illustrated in **Figure 15.17**.

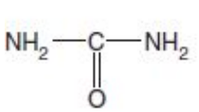
Before they can be degraded, cytidine and deoxycytidine are converted to uridine and deoxyuridine, respectively, by deamination reactions catalyzed by cytidine deaminase. Similarly, deoxycytidylate (dCMP) is deaminated to form deoxyuridylate (dUMP). The latter molecule is then converted to deoxyuridine by 5′-nucleotidase. Uridine and deoxyuridine are then further degraded by nucleoside phosphorylase to form uracil. Thymine is formed from thymidylate (dTMP) by the sequential actions of thymidine kinase and thymidine phosphorylase.

Uracil and thymine are converted to their end products, *β*-alanine and *β*-aminoisobutyrate, respectively, in parallel pathways. In the first step, uracil and thymine are reduced by dihydropyrimidine dehydrogenase to their corresponding dihydro derivatives. As these latter molecules are hydrolyzed, the rings open, yielding *β*-ureidopropionate and *β*-ureidoisobutyrate, respectively. Finally, *β*-alanine and *β*-aminoisobutyrate are produced in deamination reactions catalyzed by *β*-ureidopropionase.

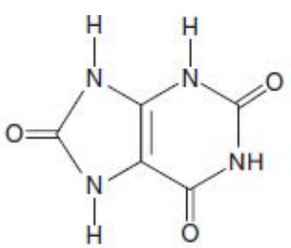
In several conditions, *β*-aminoisobutyrate is produced in such large quantities that it appears in urine. Among these conditions is a genetic predisposition for slow *β*-aminoisobutyrate conversion to succinyl-CoA and diseases that cause massive cell destruction, such as leukemia. Because it is soluble, excess *β*-aminoisobutyrate does not cause problems comparable to those observed in gout.

QUESTION 15.6

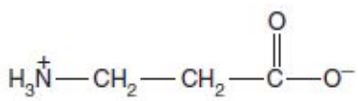
Identify each of the following biomolecules. Explain how they are produced.



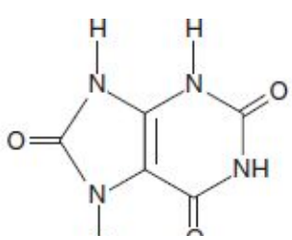
(a)



(b)



(c)



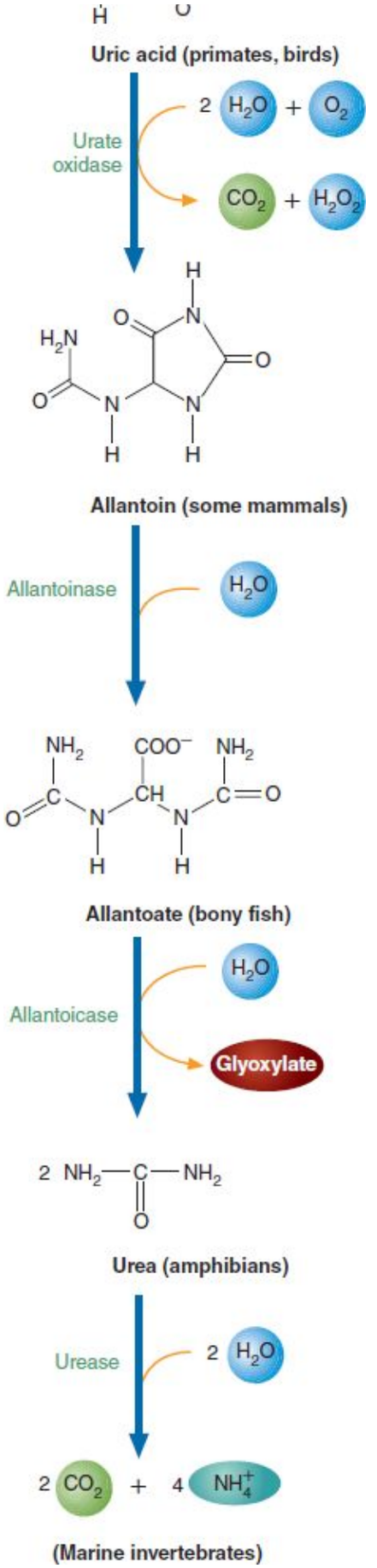


FIGURE 15.16

Uric Acid Catabolism

Many animals possess enzymes that allow them to convert uric acid to other excretory products. The final excretory products of specific animal groups are indicated.

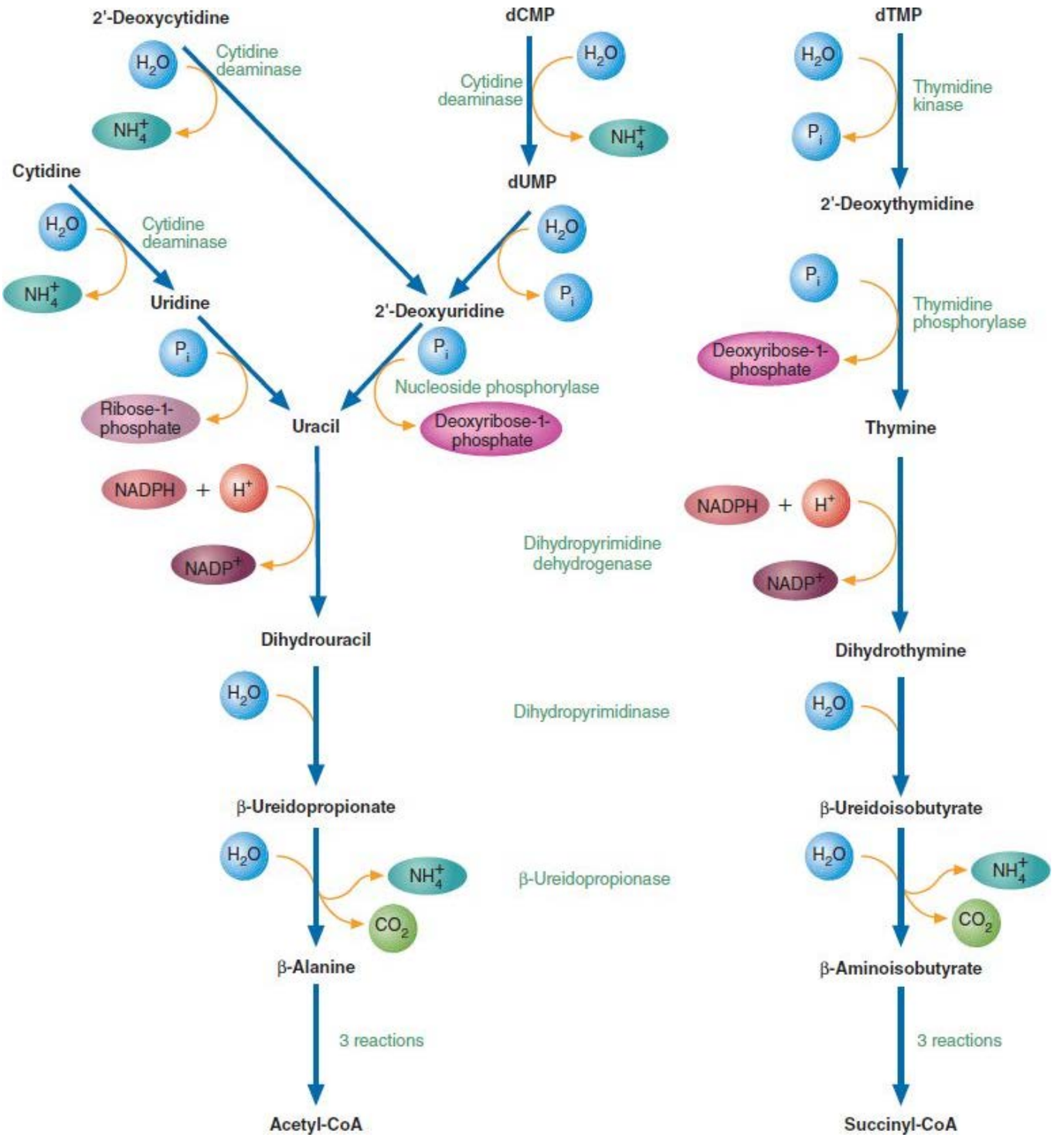


FIGURE 15.17

Degradation of Pyrimidine Bases

Uracil and thymine are degraded to *β*-alanine and *β*-aminoisobutyrate, respectively, in parallel pathways in mammalian liver. The ammonia molecules released by the reactions catalyzed by *β*-ureidopropionase are converted to urea.

QUESTION 15.7

The products of pyrimidine base catabolism, β-alanine and β-aminoisobutyrate, can be further degraded to acetyl-CoA and succinyl-CoA, respectively. Can you suggest the types of reaction required to accomplish these transformations?

KEY CONCEPTS

- Several classes of enzyme degrade nucleic acids: nucleases, phosphodiesterases, nucleotidases, nucleoside phosphorylases, and nucleosidases.
- The bases of purine nucleotides are degraded to form the nitrogenous waste product uric acid.
- β-Alanine and β-aminoisobutyrate are the nitrogenous waste products of pyrimidine base catabolism.

Chapter Summary

1. Animals are constantly synthesizing and degrading nitrogen-containing molecules such as proteins and nucleic acids. Protein turnover is believed to provide cells with metabolic flexibility, protection from accumulations of abnormal proteins, and the timely destruction of proteins during developmental processes. Most short-lived cellular proteins are degraded by the ubiquitin proteasomal system. The process begins with the covalent modification of target proteins called ubiquitination. Ubiquitin, a small, highly conserved protein, is linked to worn-out or damaged proteins or short-lived regulatory proteins. The autophagy-lysosomal system degrades long-lived proteins and organelles.
2. In general, amino acid degradation begins with deamination. Most deaminations are accomplished by transamination reactions, which are followed by oxidative deaminations that produce ammonia. Although most deaminations are catalyzed by glutamate dehydrogenase, other enzymes also contribute to ammonia formation. Ammonia is prepared for excretion by the enzymes of the urea cycle. Aspartate and CO_2 also contribute atoms to urea.
3. Amino acids are classified as ketogenic or glucogenic on the basis of whether their carbon skeletons are converted to fatty acids or to glucose. Several amino acids can be classified as both ketogenic and glucogenic because their carbon skeletons are precursors for both fat and carbohydrates.
4. Nucleic acid turnover is accomplished by several types of enzymes. The nucleases degrade the nucleic acids to oligonucleotides. (The deoxyribonucleases degrade DNA; the ribonucleases degrade RNA.) The phosphodiesterases convert the oligonucleotides to mononucleotides. By removing phosphate groups, the nucleotidases convert nucleotides to nucleosides. The nucleosidases hydrolyze nucleosides to form free bases and ribose or deoxyribose. The nucleoside phosphorylases convert ribonucleosides to free bases and ribose-1-phosphate. Dietary nucleic acids are generally degraded in the intestine and are not used in salvage pathways. Cellular purines are converted to uric acid. Many animals degrade uric acid further because they produce enzymes that are not present in primates. Pyrimidine bases are degraded to either β-alanine (UMP, CMP, dCMP) or β-aminoisobutyrate (dTMP).

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on degradation of nitrogen-containing biomolecules to help you prepare for exams.

Chapter 15 Review Quiz

Suggested Readings

Carter RN, Morton NM. 2016. Cysteine and hydrogen sulphide in the regulation of metabolism: insight from genetics and pharmacology. J Pathol 238:321–2.

Colacurcio DJ. 2018. Dysfunction of autophagy and endosomal–lysosomal pathways: roles in pathogenesis of Down syndrome and Alzheimer's disease. Free Radical Biol Med 114:40–51.

Dupont N, Codogno P. 2016. Autophagy transduces physical constraints into biological responses. Int J Biochem Cell Biol 79:419–26.

Finley D, et al. 2016. Gates, channels and switches: elements of the proteasome machine. Trends Biochem Sci 41(1):77–93.

Kaur J, Debnath J. 2015. Autophagy at the crossroads of catabolism and anabolism. Nat Rev Mol Cell Biol 16:461–72.

Korb VC, et al. 2016. *Mycobacterium tuberculosis*: manipulator of protective immunity. Int J Mol Sci 17(3):131–77.

Livneh I, et al. 2016. The life cycle of the 26S proteasome: from birth through regulation and function and onto its death. Cell Res 26:869–85.

Longo V, Panda S. 2016. Fasting, circadian rhythms and time-restricted feeding in healthy lifespan. Cell Metab 23:1048–59.

Key Words

autophagy, *574*
B cell, *592*
chaperone-mediated autophagy, *574*
glucogenic, *575*
hyperammonemia, *579*
ketogenic, *575*
Krebs bicycle, *579*
Krebs urea cycle, *576*
macroautophagy, *574*
microautophagy, *574*
mutation, *589*
nuclease, *590*
oligonucleotide, *590*
protein turnover, *571*
proteasome, *572*
T cell, *592*
transsulfuration pathway, *586*
ubiquitin, *572*
ubiquitination, *572*
ubiquitin proteasomal system, *572*
urea cycle, *576*

Review Questions

SECTION 15.1

Comprehension Questions

1. Define the following terms:
   a. ammonotelic organism
   b. uricotelic organism
   c. protein turnover
   d. ubiquitin
   e. ubiquitin proteasomal system
2. Define the following terms:
   a. proteasome
   b. ubiquitination
   c. ubiquitin-conjugating system
   d. ubiquitin ligase
   e. autophagy
3. Define the following terms:
   a. chaperone-mediated autophagy
   b. microautophagy
   c. macroautophagy
   d. autophagosome
   e. isolation membrane
4. Define the following terms:
   a. caspase-like activity
   b. lipofuscin
   c. proteasomal lid
   d. proteasomal base
   e. 20S core particle

Fill in the Blanks

5. The process of continuous synthesis and degradation of protein is referred to as _____.
6. The _____ is the term used for polypeptide destruction motifs.
7. _____ is a bacterial molecule used to prevent organ rejection in transplant patients.
8. The three forms of autophagy are _____, _____, and _____.
9. The autophagosome fuses with an endosome to form a(n) _____.
10. _____ are degrons with sequences containing proline, glutamate, serine, and threonine residues.

Short-Answer Questions

11. List three purposes served by protein turnover.
12. What structural features of proteins mark them for destruction?
13. Describe how a protein is targeted for degradation in a proteasome.
14. What are Svedberg units, and why are they not additive?
15. What are the functions of the proteasomal lid?
16. What enzyme activities does the 20S proteasomal core particle have?
17. How does *Mycobacterium tuberculosis* defeat the body's immune system?

18. Compare ammonotelic and uricotelic organisms.

Critical-Thinking Questions

19. Parkinson's disease is a devastating, progressive disorder of the central nervous system. Damage to dopaminergic nerve tracts in the brain causes tremors and impaired muscle coordination. An early-onset (inherited) form of the disease has been linked to a defective protein called Parkin. Parkin has E3 (ubiquitin ligase) activity. Speculate, in general terms, about how neurons with this defect are damaged.
20. Describe the connections between the endosomal lysosomal system and patients with diseases such as Alzheimer's disease, Niemann–Pick disease, and tuberculosis.
21. Protein aggregates form in aging cells, thereby compromising cell function. A significant feature of aggregate formation is proteasome inhibition. Explain, in general terms, how this inhibition could occur.

SECTION 15.2

Comprehension Questions

22. Define the following terms:

    a. ketogenic amino acid
    b. glucogenic amino acid
    c. L-amino acid oxidase
    d. Krebs urea cycle
    e. CPSI

23. Define the following terms:

    a. arginosuccinate
    b. Krebs bicycle
    c. *N*-acetylglutamate
    d. BH_4
    e. taurine

24. Define the following terms:

    a. CBS
    b. CSE
    c. transsulfuration pathway
    d. bacterial urease
    e. serine dehydratase

25. Define the following terms:

    a. albinism
    b. maple syrup urine disease
    c. alkaptonuria
    d. methylmalonic acidemia
    e. phenylketonuria

Fill in the Blanks

26. Amino acids that degrade yielding acetyl-CoA or acetoacetyl-CoA are referred to as _____.
27. _____ amino acids are degraded to yield pyruvate or citric acid cycle intermediates.
28. Urea synthesis begins with the formation of _____.
29. Urea is eliminated from the body by the _____.
30. _____ is a potentially fatal condition in which there are high blood levels of NH_4^+.

Short-Answer Questions

31. Explain the significance of aspartate in the following processes: citric acid cycle, urea cycle, and amino acid degradation.
32. What are the seven metabolic products produced by the degradation of the standard amino acids?
33. Indicate which of the following amino acids are ketogenic and which are glucogenic:

    a. tyrosine
    b. lysine
    c. glycine
    d. alanine
    e. valine
    f. threonine

34. Describe how glutamate is degraded.
35. The urea cycle occurs partially in the cytoplasm and partially in mitochondria. Describe the urea cycle reactions in reference to their cellular locations.
36. Describe the role of the glucose–alanine cycle in ammonia transfer to the liver.
37. In individuals with PKU, tyrosine is an essential amino acid. Explain.
38. Describe the Krebs bicycle. What molecule links the two cycles?
39. Most amino acids are degraded in the liver. This is not true of the BCAA, most of which are degraded in extrahepatic tissues with high protein turnover. Suggest some examples of these tissues.
40. Describe lysine degradation.
41. Explain why feeding domestic cats with a vegetarian diet threatens their health.
42. Trace the origin of the carbon and nitrogen atoms in the urea molecule.
43. Explain why the carbon skeletons of ketogenic amino acids yielding acetyl-CoA only cannot be converted into glucose.

Critical-Thinking Questions

44. Mammals excrete most nitrogen atoms as urea. The urea cycle itself requires considerable amounts of ATP. What mechanism does the liver cell have to compensate for that energy input?
45. Describe how increasing concentrations of ammonia stimulate the formation of *N*-acetylglutamate and turn on the urea cycle.
46. Phenylketonuria can be caused by deficiencies in phenylalanine hydroxylase and by enzymes catalyzing the formation and regeneration of 5,6,7,8-tetrahydrobiopterin. How can this second defect cause the symptoms of PKU?
47. In their in vitro studies using liver slices, Krebs and Henseleit observed that urea formation was stimulated by the addition of ornithine, citrulline, and arginine. Other amino acids had no effect. Explain these observations.
48. Specify the type of carbon unit that is transferred in one-carbon metabolism by each of the following molecules:

a. N , N -methylene THF
b. serine
c. choline
d. S-adenosylmethionine

49. Some animals living in water excrete nitrogen as ammonia. Land animals, which must conserve water, excrete urea or uric acid. How does excretion of these molecules aid in water conservation?
50. Diabetes mellitus is a set of metabolic diseases with the common symptom of an inability to transport glucose into target cells (muscle cells and adipocytes). The body compensates in part by degrading muscle protein to generate energy. Explain how this process works.
51. Individuals with hyperammonemia are given α-keto acids as a treatment. Explain. (Visit the online website to read the essay on hyperammonemia.)
52. Review the Krebs bicycle and calculate the ATPs degraded or synthesized in this pathway. Is urea biosynthesis a net energy-requiring or energy-generating process?

SECTION 15.3

Comprehension Questions

53. Define the following terms:

a. oligonucleotide
b. nuclease
c. phosphodiesterase
d. nucleosidase
e. nucleotidase

54. Define the following terms:

a. uric acid
b. allantoin
c. allantoate
d. urate oxidase
e. allantoicase

55. Define the following terms:

a. T cells
b. B cells
c. phosphodiesterase
d. AMP deaminase
e. myoadenylate deaminase deficiency.

Fill in the Blanks

56. _____ are short nucleic acid segments containing 50 nucleotides or less.
57. Uracil and thymine are converted to their end prod-ucts _____ and _____, respectively.
58. The purine nucleotide cycle produces fumarate, _____, and _____.
59. _____ is a potentially toxic product of reactions catalyzed by xanthine oxidase.
60. β-Alanine is one of the waste products of _____ metabolism.

Short-Answer Questions

61. Identify which of the following molecules yields uric acid when degraded:

    a. DNA
    b. FAD
    c. CTP
    d. PRPP
    e. β-alanine
    f. urea
    g. NAD^+

62. Foods high in purines can trigger a gout attack. Provide several examples of these foods and explain how they contribute to this malady. (Read the online essay on gout.)
63. Purine nucleoside phosphorylase (PNP) deficiency disease is an autosomal recessive disease that results in severe immunodeficiency. Explain.
64. What is the end product of catabolism of the pyrimidine base thymine? What circumstances cause excess amounts of this end product, and why doesn't this molecule (unlike uric acid, the end product of purine catabolism) cause a gout-like illness?

Critical-Thinking Questions

65. Create a diagram that illustrates how the purine nucleotide cycle contributes to energy metabolism in skeletal muscle. Include glycolysis and the citric acid cycle in your drawing.
66. Caffeine, a methylated xanthine found in chocolate, coffee, and tea, is excreted as uric acid. Use your knowledge of the metabolism of other purine compounds to suggest how caffeine is metabolized.
67. Dihydrouracil and β-ureidopropionate (N-carbamoyl–β-alanine) are intermediates in the conversion of uracil to β-alanine. Provide the structures of the molecules in this pathway.
68. Refer to Question 67 and outline the pathway by which the nitrogen of uracil is used in the synthesis of urea.
69. Why do primates suffer from gout, whereas most other animals do not?
70. Why can't humans degrade purine rings?

MCAT Study Questions

71. Which of the following is a substrate in the urea cycle?

    a. fumarate
    b. glutamate
    c. aspartate
    d. α-ketoglutarate

72. Which of the following biochemical pathways delivers amino nitrogen from muscle cells to the liver to be disposed of in urea molecules?

    a. Krebs bicycle
    b. Cori cycle
    c. purine cycle
    d. glucose–alanine cycle

73. Products of methionine degradation include all of the following except

    a. succinyl-CoA

b. glutamate-γ-semialdehyde
c. homocysteine
d. *S*-adenosylmethionine

74. Which of the following is a degradation product of both cytidine and deoxycytidine?

a. β-alanine
b. β-aminoisobutyrate
c. succinyl-CoA
d. uric acid

75. Urea is the product of which of the following enzymes?

a. urate oxidase
b. aminohydrolase
c. urease
d. xanthine oxidase

CHAPTER 16 Integration of Metabolism

Integration of Metabolism Swans fly at speeds of 25 to 50 mph as a result of metabolic processes, most notably the conversion of food molecules derived from aquatic plants into skeletal muscle contraction.

OUTLINE

HYPERTENSION AND URIC ACID: A DIET CONNECTION?

16.1 OVERVIEW OF METABOLISM

16.2 INTERCELLULAR COMMUNICATION

Polypeptide Signal Molecules

Cell-Surface Receptors
Enzyme-Linked Receptors
Nuclear Receptors

16.3 METABOLISM IN THE MAMMALIAN BODY: THE FEEDING–FASTING CYCLE
The Feeding Phase
The Fasting Phase
Feeding Behavior

Biochemistry in Perspective
Diabetes Mellitus

Biochemistry in Perspective
Obesity and the Metabolic Syndrome

AVAILABLE ONLINE
Biochemistry in Perspective
Mammalian Hormones and the Hormone Cascade System
Biochemistry in Perspective
Metabolism in the Mammalian Body: Division of Labor

Hypertension and Uric Acid: A Diet Connection?

Jake is a 15-year-old with a problem. He is an overweight teenager newly diagnosed with hypertension at the hypertension clinic at a local university medical center. Jake was one of several dozen overweight teens selected for a clinical trial because they had recently been diagnosed with mild untreated hypertension (blood pressures of 135/80 [systolic/ diastolic], compared with normal values of 110–120/60–70). They also had serum uric acid (SUA) levels of 6 mg/dL (normal values are 5.5 mg/dL or lower) and no kidney damage or other diseases related to hypertension.

Jake's clinical trial was designed to ascertain the relationship of high blood uric acid levels (hyperuricemia) to hypertension. The experimental drug used in the trial, allopurinol, is an inhibitor of xanthine oxidase (p. 590), the enzyme that converts the purine derivatives hypoxanthine and xanthine to uric acid, a nitrogenous waste product in humans. Although hypertension has long been linked to hyperuricemia (and assigned a causative role by early researchers), modern biomedical scientists have regarded high SUA levels as a consequence of hypertension (decreased uric acid excretion caused by hypertensive kidney damage). Recent research by epidemiologists (public health researchers who investigate disease risk factors) and experimental studies using lab animals has challenged this view.

In Jake's clinical trial, allopurinol lowered SUA levels to 5 mg/dL or less in two-thirds of the patients. Lower SUA levels in these patients were associated with statistically significant reductions in blood pressure. In other investigations, vascular damage has been demonstrated in hyperuricemic lab animals. Together, these results support the hypothesis that hypertension is the result and not the cause of high SUA levels. In other words, hypertension-related damage to renal blood vessels is caused by toxic amounts of uric acid.

High serum uric acid levels are statistically linked not only to hypertension and kidney disease but also to obesity, diabetes, and cardiovascular diseases (pp. 626–27). SUA values have been steadily rising over the past century. Long thought to be the result of the sedentary lifestyles typical of industrialized societies, these disease conditions have recently been associated by epidemiologists with excessive fructose consumption. Major increases in hypertension and obesity beginning about three decades ago parallel the introduction of increased quantities of fructose (via sucrose and high-fructose corn syrup) into

processed foods and beverages. Indeed, administering fructose-rich food to humans in clinical trials results in pronounced increases in blood pressure. Recall that fructose metabolism bypasses the key regulatory step in glycolysis, catalyzed by PFK-1 (p. 298). After a high-fructose meal, fructose, metabolized primarily in the liver, is rapidly converted into phosphorylated trioses that then enter pathways involved in lipogenesis and VLDL synthesis and secretion. This unregulated process depletes hepatic ATP levels and increases AMP, the substrate for uric acid synthesis. High SUA levels are believed to damage blood vessels, including those of the kidney, by inhibiting endothelial nitric oxide synthase (p. 552), the enzyme that synthesizes nitric oxide, a molecule that dilates arteries.

Overview

PREVIOUS CHAPTERS DEAL WITH THE METABOLISM OF CARBOHYDRATES, LIPIDS, AND OTHER MOLECULES. HOWEVER, THE WHOLE IS NOT JUST THE sum of its parts. Multicellular organisms are extraordinarily more complex than a list of their components would suggest. Chapter 16 takes a wider view of functioning of the mammalian body. A review of the mechanisms of action of hormones and other signal molecules that make sophisticated regulation possible is followed by a description of the feeding–fasting cycle, a physiological process that ensures that adequate energy and nutrient resources are available. The contributions of the major organs in mammals to this process are described in an online reading, Metabolism in the Mammalian Body: Division of Labor.

The most distinctive characteristic of living organisms is their capacity to sustain adequate, if not always optimal, operating conditions, despite changes in their internal and external environments. And if this feat is not amazing enough, consider that they must also simultaneously repair damaged components and, when possible, undergo cell divisions and other forms of growth. To accomplish these functions, the anabolic and catabolic reaction pathways that use carbohydrates, lipids, and proteins as energy sources and biosynthetic precursors must be precisely regulated. For multicellular organisms such as animals, this endeavor is astonishingly complicated. The operation of such a complex system as the body is maintained by a continuous flow of information among its parts. A simple system for information transfer is composed of a primary signal (e.g., a hormone), a target (a specific cell receptor), and a transducer system (that converts the signal to a cellular response). In view of the complexities of multicellular organisms, the need for a large number of primary signals, specific receptors, and transducer systems is not surprising. In the mammalian body, information transfer is accomplished by hormones and other types of signaling molecules that affect cells in other parts of the body. Together these molecules allow for a high degree of sophisticated regulation.

Chapter 16 focuses on the integration of the major metabolic processes in mammals. The chapter begins with an overview of metabolic processes and descriptions of the most important classes of signal molecules and their mechanisms of action. The chapter ends with an overview of the hormone- and neurotransmitter-regulated feeding–fasting cycle, which has great physiological importance because of its role in energy acquisition. The chapter also describes disorders of metabolic regulation: diabetes, obesity, and the metabolic syndrome.

16.1 OVERVIEW OF METABOLISM

The central metabolic pathways are common to most organisms. Throughout the life of an organism, a precise balance is struck between anabolic (synthetic) and catabolic (degradative) processes. An overview of the principal anabolic and catabolic pathways in heterotrophs such as animals is illustrated in **Figure 16.1**. Except during youth, illness, or pregnancy, the animal's tissues exist in a metabolic steady state throughout the remainder of its life. In a **steady state**, the rate of anabolic processes is approximately equal to that of catabolic processes.

How are animals (or other multicellular organisms) able to maintain a balance between anabolic and catabolic processes as they respond and adapt to changes in their internal and external environments? They do so using various types of intercellular communication. Most intercellular communication occurs by means of biochemical signals. Once released into the extracellular environment, each signal molecule is recognized by specific cells (called **target cells**), which then respond in a specific manner. Most signals are modified amino acids, fatty acid derivatives, peptides, proteins, or steroids.

In animals, the nervous and endocrine systems are primarily responsible for coordinating metabolism. The nervous system provides a rapid and efficient mechanism for acquiring and processing environmental information. Nerve cells, called neurons, release neurotransmitters (p. 551) at the end of long cell extensions called axons into tiny intercellular spaces called synapses. The neurotransmitter molecules bind to nearby cells, evoking specific responses from those cells.

The best-researched example of metabolic regulation is the endocrine system, in which signal molecules called hormones are secreted directly into the blood. The **endocrine system** is composed of specialized cells, many of which are found in glands. After these hormone molecules, referred to as *endocrine* hormones, have been secreted, they travel through the blood until they reach a target cell. Some hormones exert very specific effects on one type of target cell; other hormones act on a variety of target cells. For example, thyroid-stimulating hormone (TSH), released from the pituitary gland in the brain, stimulates follicular cells in the thyroid gland to release T_3 (triiodothyronine) and T_4 (thyroxine) (**Figure 16.2**). In contrast, T_3 (the most active form of thyroid hormone) and T_4 then proceed to stimulate a variety of cellular reactions in numerous cell types (e.g., T_3 stimulates glycogenolysis in liver cells and glucose absorption in the small intestine).

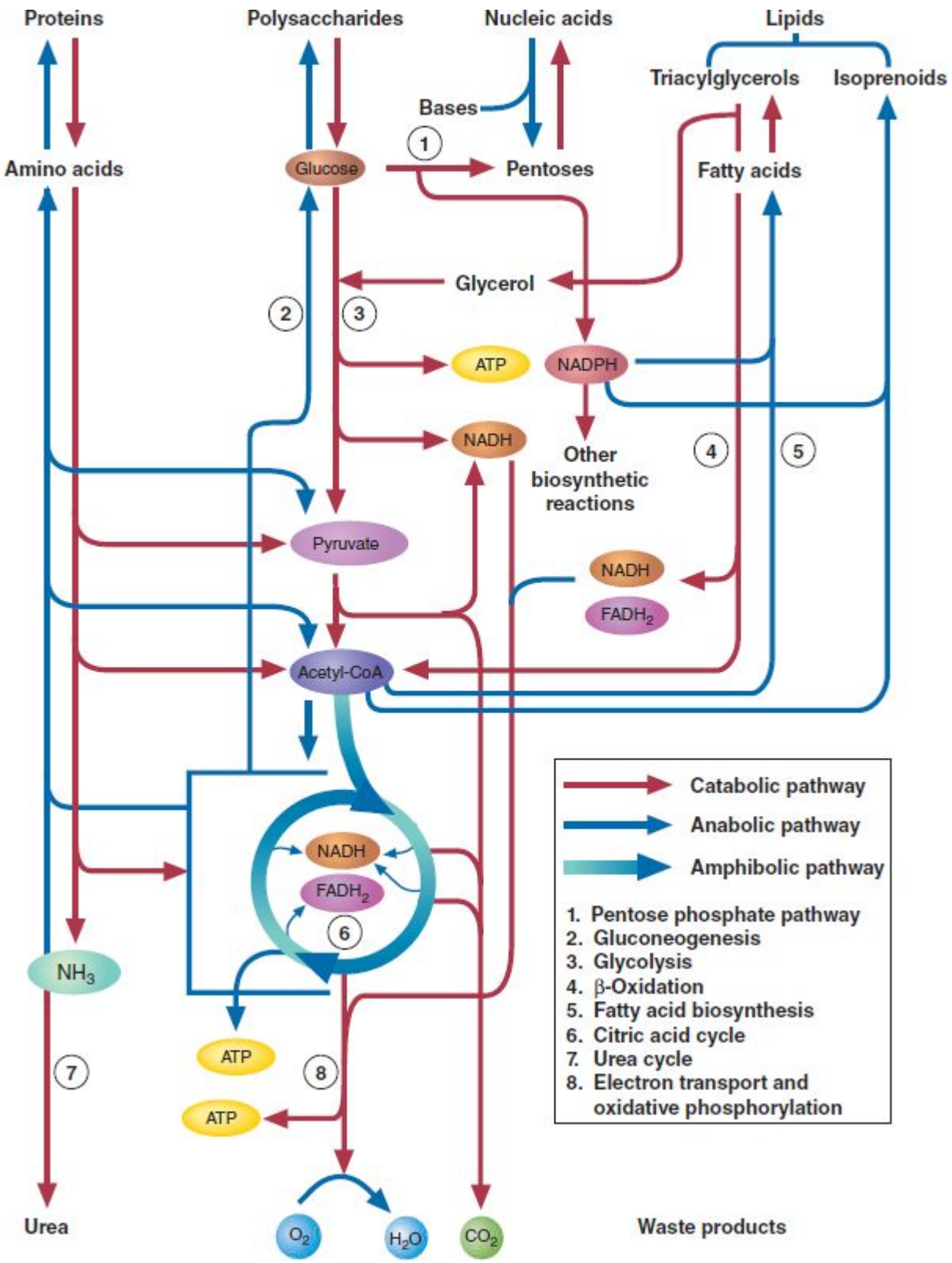
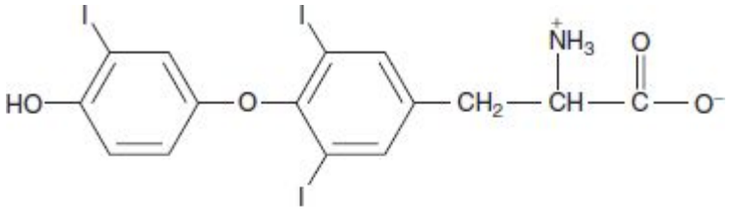


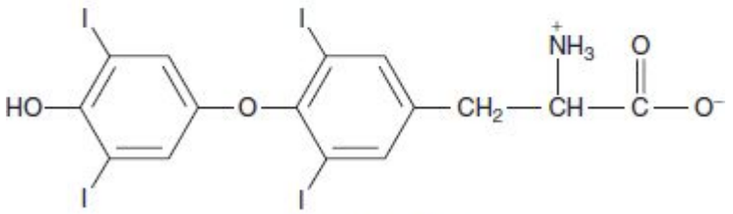
FIGURE 16.1

Overview of Metabolism

This simplified overview of metabolism illustrates the anabolic and catabolic pathways of the major biomolecules in heterotrophs (i.e., the biochemical pathways that synthesize, degrade, or interconvert important biomolecules and generate energy).



Triiodothyronine (T_3)



Thyroxine (T_4)

FIGURE 16.2

Structure of the Thyroid Hormones T_3 and T_4

3D animation of Triiodothyronine

Other signal molecules may also have autocrine or paracrine effects, which reach target cells by diffusing short distances. In *autocrine* signaling, a cell produces and secretes signal molecules via exocytosis; these signal molecules then bind to receptors on the surface of the same cell. *Paracrine* signal molecules released from one cell bind to receptors on neighboring cells.

16.2 INTERCELLULAR COMMUNICATION

A wide variety of biomolecules regulate the body's metabolic pathways by altering the activity, the concentration, or the location of enzymes in processes that are initiated by ligand–receptor binding events. Hormones are classified as *water-soluble* (peptides, polypeptides, and amino acid derivatives) and *lipid-soluble* (steroid and thyroid hormones). Additional water-soluble proteins, referred to as growth factors and cytokines, are also important regulators of cell activities.

Growth factors have traditionally been defined as proteins that stimulate cell growth, proliferation, and differentiation. Examples include epidermal growth factor (EGF), platelet-derived growth factor (PDGF), and insulin-like growth factor (IGF). The term **cytokine** has been used to describe proteins that promote blood and immune system cells to proliferate or differentiate. Examples include interleukins (which stimulate immune responses such as inflammation), interferons (infection-fighting proteins released by host cells in response to pathogens), and erythropoietin (which stimulates red cell production in response to cellular hypoxia). Research efforts in recent years have revealed that both growth factors and cytokines have more diverse roles than previously understood and often have similar or overlapping functions and signal transduction mechanisms. As a result, many life scientists now use the terms *growth factor* and *cytokine* interchangeably. In addition, various growth factors and cytokines can have autocrine, paracrine, or endocrine signaling mechanisms.

Polypeptide Signal Molecules

In animals, the vast majority of water-soluble intercellular signal molecules are peptides or polypeptides. They initiate their actions by binding to receptors on the outer surface of the target cell's plasma membrane. In mammals, the best-researched examples of intercellular signal molecules are the endocrine hormones, polypeptides whose synthesis and secretion are regulated by a complex cascade mechanism that is ultimately controlled by the central nervous system. (Note that an overview of mammalian endocrine hormones is provided in an online Biochemistry in Perspective essay on mammalian hormones and the hormone cascade system.) Sensory signals are received by the hypothalamus, an area in the brain that integrates the nervous and endocrine systems. For example, in response to high blood Na^+ levels, hypothalamic osmoreceptor cells trigger the secretion of the antidiuretic peptide hormone vasopressin (p. 149) into the posterior pituitary, where it is then released into the bloodstream. A major feature of vasopressin's actions in the kidney is increased synthesis and insertion of aquaporin channels into specific epithelial cells. As a result, water molecules move down their concentration gradient, thereby increasing water reabsorption from the urinary filtrate.

The receptors for most water-soluble hormones are located on the surface of target cells. The binding of these hormones to their cognate membrane-bound receptors triggers an intracellular response. These intracellular responses are initiated by a group of molecules referred to as **second messengers**. (The polypeptide hormone is the first messenger.) Several second messengers have been identified. These include the cyclic AMP (cAMP), cyclic GMP (cGMP), calcium ions, and certain phosphatidylinositol derivatives. Most second messengers trigger physiological changes such as cell proliferation, differentiation, or apoptosis by altering the activities or concentrations of enzymes in a process referred to as an enzyme cascade. In an *enzyme cascade* (**Figure 16.3**), a second messenger triggers a sequence of successive reactions, primarily phosphorylations/dephosphorylations, which cause enzymes to undergo conformational transitions that switch them from their inactive forms to their active forms, or vice versa, while simultaneously amplifying the original signal. In addition to rapidly altering the activities of enzymes already present, enzyme cascades alter the expression of specific protein-coding genes via transcription factors.

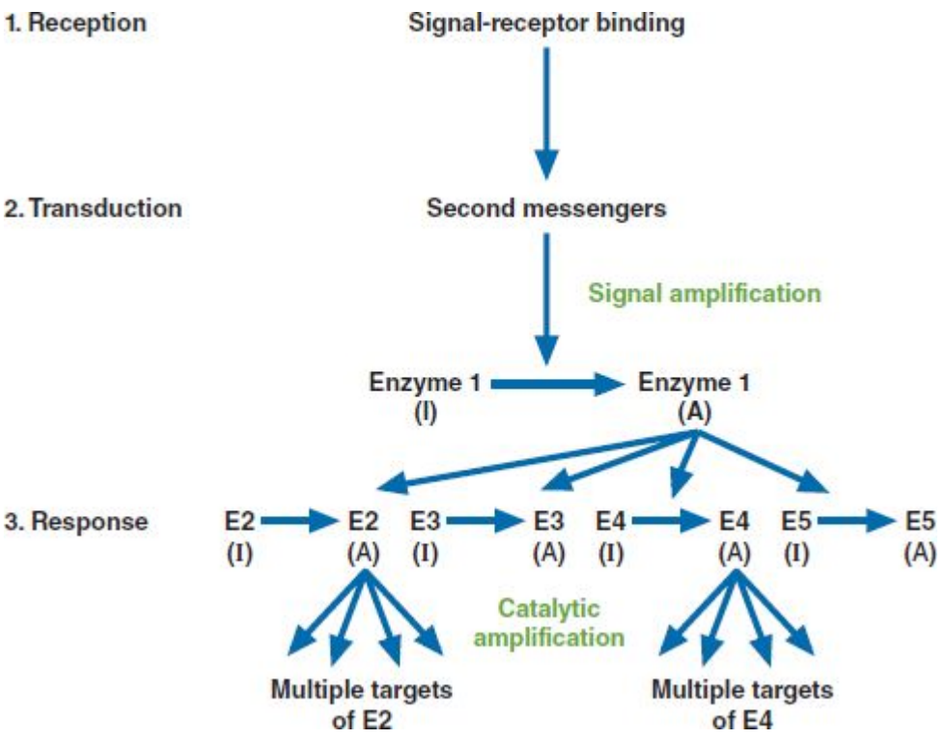


FIGURE 16.3

Signal Transduction

Signal transduction mechanisms, initiated by the binding of ligands to cell-surface receptors, occur in three phases. (1) Reception: a signal molecule binds to its receptor. (2) Transduction: as a result of receptor binding, an enzyme cascade is initiated by second messenger molecules. The enzyme molecules activated by the second messenger molecules modify multiple copies of a number of different target enzymes. Newly activated enzymes may also modify multiple copies of a second set of target proteins. (3) Response: cellular functions are altered by changes in the activities of existing enzymes, rearrangements of the cytoskeleton, and/or altered gene expression. Signal amplification occurs at each step of the cascade. (I = inactive, A = active)

The target cell's response to the binding of the hormone molecule begins when the second messenger binds to a specific enzyme. For example, the binding of cAMP to inactive protein kinase A (PKA) converts it to active PKA, which in turn modifies the activity of many target enzymes through phosphorylation of specific serine or threonine residues. The original signal thereby generates an amplified and diversified response.

Animals use several mechanisms to prevent excessive hormone synthesis and release. The most prominent of these is feedback inhibition (**Figure 1.22**). For example, the brain's hypothalamus and anterior pituitary are controlled by the target cells they regulate. When blood levels of T3 and T4 rise above a specific threshold, TSH release by the anterior pituitary is inhibited. The thyroid hormones also inhibit the responsiveness of TSH synthesizing cells to thyrotropin-releasing hormone (TRH), the hypothalamic hormone that stimulates TSH release from the anterior pituitary. (A *tropic hormone* stimulates an endocrine gland to secrete its hormones.)

s

Target cells also possess mechanisms that protect against overstimulation by hormones. In a process referred to as **desensitization**, target cells adjust to changes in stimulation levels by decreasing the number of cell-surface receptors or by inactivating those receptors. The reduction in cell-surface receptors in response to stimulation by specific hormone molecules is called **downregulation**. In downregulation, receptors are internalized by endocytosis. Depending on cell type and several metabolic factors, the receptors may eventually be recycled to the cell surface or be degraded. If degraded, new receptor proteins must be synthesized to replace old receptors.

Some disease states are caused by or associated with target cell insensitivity to specific hormones. For example, some cases of diabetes mellitus are associated with **insulin resistance**, caused by a decrease in cell-surface insulin receptors. (Diabetes is discussed in a Biochemistry in Perspective essay later in this chapter: pp. 615–17.)

KEY CONCEPTS

Many of the hormones in the mammalian body are controlled by a complex cascade mechanism and ultimately regulated by the central nervous system.

QUESTION 16.1

Review the epinephrine-stimulated activation of glycogen breakdown presented earlier (**Figure 8.21**). Identify the following signal transduction components in this biochemical process: primary signal, receptor, transducer, and response.

Cell-Surface Receptors

There are two major classes of cell-surface receptor: G protein-coupled receptors and enzyme-linked receptors. Intracellular receptors, referred to as nuclear receptors, are described on pp. 614–18.

G PROTEIN-COUPLED RECEPTORS **G protein-coupled receptors (GPCRs)** are the largest known protein receptor family. (Humans have more than 900 GPCR genes, approximately 4% of the human protein-coding genome.) GPCR proteins are composed of seven membrane-spanning helices (**Figure 16.4**) that are arranged into a three-dimensional barrel-like shape. The seven transmembrane segments are linked by three extracellular loops and three intracellular loops. An extracellular N-terminal segment forms part of the ligand-binding site, and an intracellular C-terminal segment interacts with G proteins, also known as heterotrimeric guanosine nucleotide-binding proteins. GPCRs transduce a wide variety of stimuli into intracellular signals. In addition to hormones such as glucagon, TSH, the catecholamines, and the endocannabinoids (e.g., the arachidonic acid derivative anandamide), GPCRs also respond to neurotransmitters (e.g., acetylcholine, glutamate, dopamine, and GABA), neuropeptides (e.g., vasopressin and oxytocin), odorants and tastants (molecules that stimulate the senses of odor and taste, respectively), and light (rhodopsin).

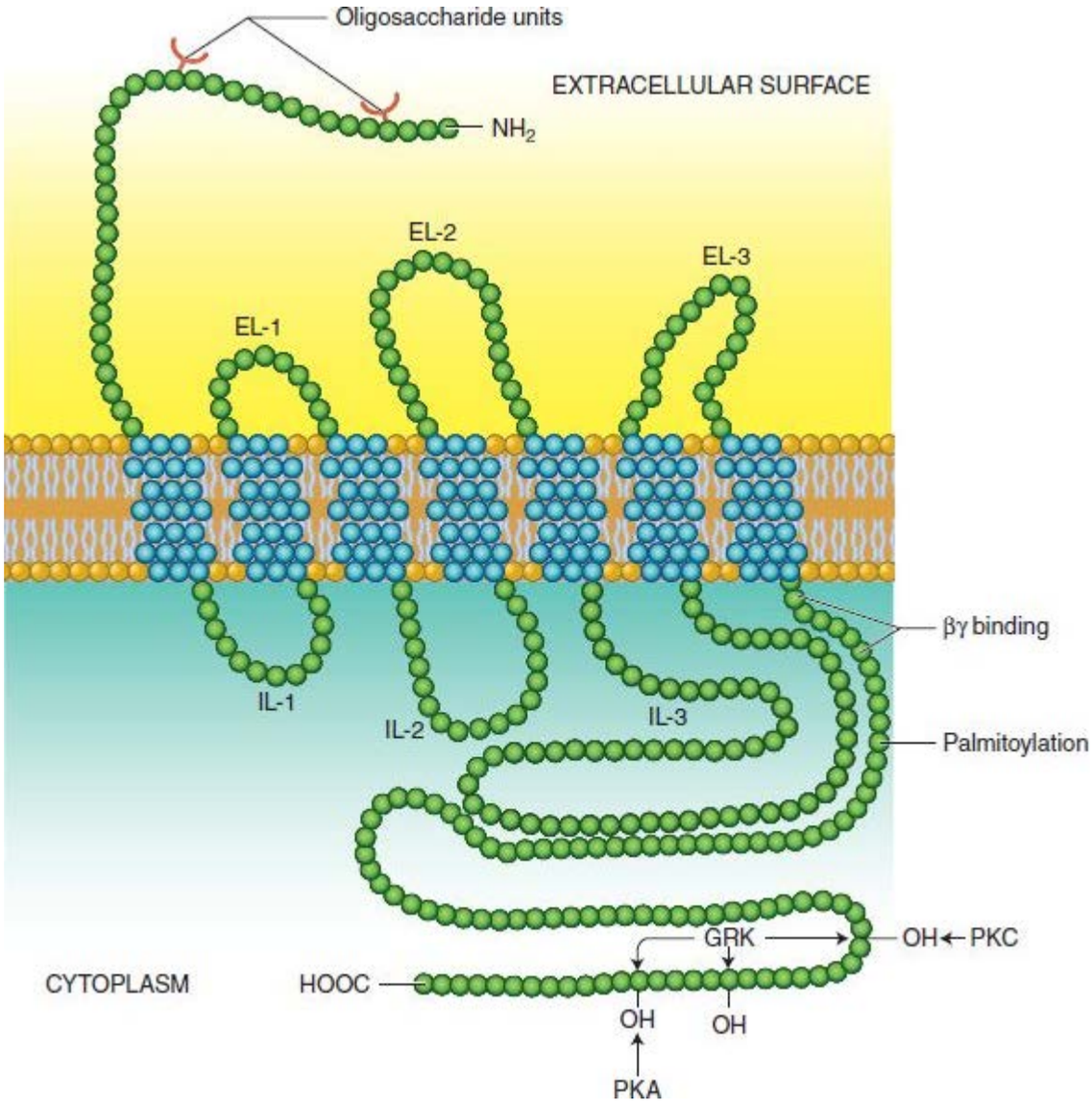


FIGURE 16.4

G protein-Coupled Receptor Structure

GPCRs are single-polypeptide chains characterized by seven transmembrane α-helices (forming into a barrel), extracellular loops (EL-1, EL-2, and EL-3) and intracellular loops (IL-1, IL-2, and IL-3), and variable-length extracellular N-terminal and intracellular C-terminal segments. GPCRs bind their ligands within their transmembrane domain (e.g., small molecules such as epinephrine) and/or extracellular (N-terminal) domain (e.g., polypeptides such as glucagon). N-terminal segments and extracellular loops are often glycosylated (N-linked oligosaccharides). Certain serine or threonine residues on the internal segment can be phosphorylated by PKA, PKC, and GRK (p. 610). This generalized illustration of a GPCR depicts the approximate locations of the $\beta\gamma$-binding segment (for G protein binding), one (or more) cysteine residues to undergo palmitoylation (allowing penetration into the plasma membrane), and the serine or threonine residues whose phosphorylation is required to initiate desensitization (refer to p. 605).

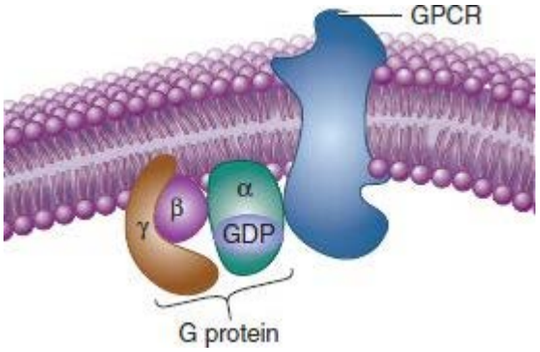


FIGURE 16.5

GPCR Structure and G protein

The binding of a ligand to a G protein-coupled receptor initiates a signaling mechanism that begins with the binding of the GPCR to a G protein. Each G protein consists of three subunits: α, β, and γ. The α subunit binds GDP and GTP. Before a G protein is activated the α-subunit binds GDP.

G proteins are the molecular switches that transduce ligand binding to GPCRs into intracellular signals (**Figure 16.5**). G proteins are composed of α, β, and γ subunits. The α subunit binds GTP and GDP. The transmembrane region of the ligand-bound GPCR undergoes a conformational change that causes it to act as a **guanine nucleotide exchange factor (GEF)**, resulting in the exchange of GDP bound to the α subunit for GTP. The GTP-Gα subunit then separates from the $\beta\gamma$ dimer. There are four subgroups of α subunits: Gα_s (stimulates adenylate cyclase), Gα_i (inhibits adenylate cyclase), Gα_q (activates phospholipase C), and Gα_{12} (activates rho, a type of monomeric G protein). The best-researched examples of α subunits are Gα_s and Gα_i.

Termination of the GPCR signal involves the hydrolysis of the Gα-subunit's GTP. The binding of a **GTPase-activating protein (GAP)**, a protein that stimulates the GTPase activity of G proteins, results in GTP hydrolysis to yield a GDP-Gα subunit.

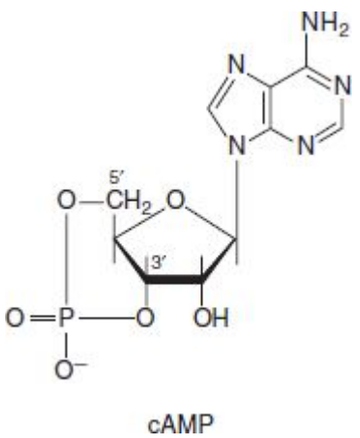
The $\beta\gamma$ complex, composed of β and γ subunits (five and twelve different types, respectively), binds to and inhibits the GDP-Gα subunit. Myristoyl or palmitoyl groups attached to the α subunits and prenyl groups attached to γ subunits anchor G proteins to the plasma membrane. The $\beta\gamma$ dimer promotes the association of the α subunit to the GPCR and, in the absence of receptor activation, prevents GDP/GTP exchange. It also facilitates the anchoring of the α subunit to the membrane and plays a role in effector signaling downstream (postreceptor activation).

Examples of ligands that bind to GPCRs linked to Gα_s proteins (transiently raising local cAMP levels) include glucagon (p. 299), melanocyte stimulating hormone, and TSH, as well as numerous odorants and tastants. Ligand binding to GPCRs associated with Gα_i results in inhibition of adenylate cyclase that suppresses PKA-catalyzed target protein phosphorylation. For example, lipolysis in adipocytes is inhibited when epinephrine (a catecholamine neurotransmitter and hormone, p. 551) binds to the α_2 adrenergic GPCR, resulting in the activation of Gα_i.

When ligand binds to its cognate GPCR, GTP-α_s moves over the cytoplasmic surface of the plasma membrane to bind to and activate its target protein, the second messenger–generating enzyme adenylate cyclase. As a result, there is a localized increase in cAMP synthesis, which then triggers a signal transduction cascade. G$\beta\gamma$ dimers also have roles in GPCR signal transduction other than as inactivators of Gα subunits. Examples include activation of a cardiac inwardly rectifying potassium channel (causing membrane hyperpolarization that slows the heart rate) and regulation of phosphoinositide metabolism (p. 609).

Second messenger molecules other than cAMP are also involved in GPCR signal transduction. These include components of the phosphatidylinositol cycle and calcium ions.

cAMP cAMP (**Figure 16.6**) is generated from ATP by adenylate cyclase when hormones such as glucagon, TSH, and epinephrine bind to their receptors. G_S bound to an occupied receptor (signal initiation) undergoes GDP/GTP exchange, and the GTP-α_S subunit dissociates and activates adenylate cyclase (**Figure 16.7**). GTP hydrolysis terminates this association, and GDP-α_S recombines with GPCR-$\beta\gamma$ (primary signal termination). (In effect, α_S-GTP is a timing device: The rate at which the α subunit hydrolyzes GTP determines the signal's duration.) The activated adenylate cyclase synthesizes a number of cAMP molecules (signal amplification), which diffuse into the cytoplasm, where they bind to and activate the regulatory subunits of the heterotetramer cAMP-dependent protein kinase (PKA). The now-active PKA subunits then phosphorylate and thereby alter the catalytic activity of key regulatory enzymes. The cAMP is quickly hydrolyzed by phosphodiesterase (secondary signal termination). The activated PKA subunits migrate into the nucleus where they phosphorylate CREB (cAMP response element binding protein, p. 165), thereby creating the unstructured segment that allows binding to a coactivator called CBP (CREB-binding protein). CRB promotes transcription in part by acetylating histones, which results in increasing the accessibility of DNA to transcription factors. Examples of metabolic enzymes synthesized in response to activated CREB include the gluconeogenic enzymes PEPCK and glucose-6-phosphatase, as well as tyrosine hydroxylase, the enzyme that catalyzes the first reaction in catecholamine synthesis. (Refer to the online Chapter 14 reading Catecholamine Biosynthesis.)



cAMP

FIGURE 16.6

Structure of the Second Messenger Molecule cAMP

Note that cAMP, also referred to as adenosine 3′,5′-phosphate, is a diester that links the 3′ and 5′ carbons.

3D animation of cAMP

The target proteins affected by cAMP depend on the cell type. In addition, several hormones may activate the same G protein. Therefore, different hormones may elicit the same or similar effects. For example, glycogen degradation in liver cells is initiated by both epinephrine and glucagon.

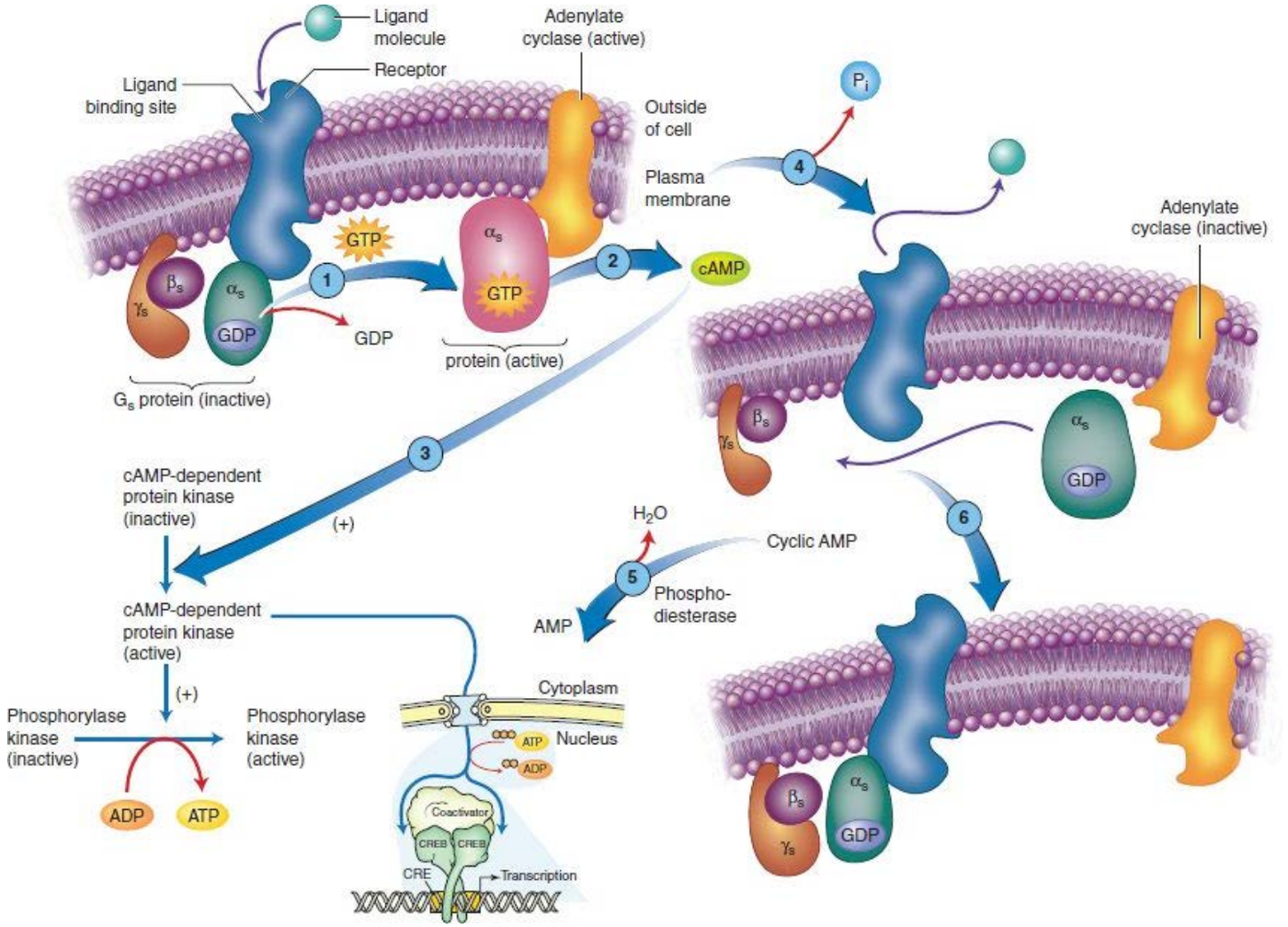


FIGURE 16.7

The Adenylate Cyclase Second Messenger System That Controls Glycogenolysis

When the receptor is unoccupied, the G_S protein α_S subunit has GDP bound and is complexed with the $\beta\gamma$ dimer. The binding of hormone (1) activates the receptor and leads to replacement of GDP with GTP by a GEF activity (not shown). The activated α-subunit interacts with and activates adenylate cyclase, which then catalyzes the conversion of ATP to cAMP (2). The second messenger molecule cAMP binds to and activates cAMP-dependent protein kinase (PKA) (3). PKA proceeds to stimulate glycogen breakdown by activating (via phosphorylation) the glycogen-degrading enzyme glycogen phosphorylase. The active subunits of PKA also move into the nucleus, where they activate the transcription factor CREB, allowing it to bind to DNA sequences referred to as CREs (cAMP-response elements) in combination with the coactivator CBP. As a result, cAMP-inducible genes are transcribed. (4) Signal transduction ends when the ligand leaves the receptor, the bound GTP is hydrolyzed to GDP by the GTPase activity within the α_S subunit, and the α_S subunit dissociates from adenylate cyclase. (5) Cyclic AMP is deactivated by hydrolysis to AMP, a reaction catalyzed by phosphodiesterase. (6) The α_S subunit then reassociates with the $\beta\gamma$ dimer.

Some hormones inhibit adenylate cyclase activity. Such molecules depress cellular protein phosphorylation reactions because their receptors interact with G_i protein. When G_i is activated, its α_i subunit dissociates from the $\beta\gamma$ dimer and prevents the activation of adenylate cyclase. For example, because its receptors in adipocytes are associated with G_i, PGE_1 depresses lipolysis. Note that the effect of a hormone often depends on which GPCR receptor it is bound to. For example, lipolysis (p. 448) is stimulated by epinephrine when it binds to the β-adrenergic receptor (activates adenylate cyclase), and is depressed when epinephrine binds to the $\alpha 2$ adrenergic receptor (inactivates adenylate cyclase).

THE PHOSPHATIDYLINOSITOL CYCLE, IP_3, DAG, AND CALCIUM The phosphatidylinositol cycle (**Figure 16.8**) mediates, primarily via G_q proteins, the actions of numerous signaling molecules that include neurotransmitters, hormones and growth factors. Examples include acetylcholine (e.g., insulin secretion in pancreatic cells), epinephrine (α_i receptors), and vasopressin. Phosphatidylinositol-4,5-bisphosphate (PIP_2), a minor membrane phospholipid, is cleaved by phospholipase C to form the second messengers **DAG (diacylglycerol)** and **IP_3 (inositol-1,4,5-trisphosphate)**. A hormone–receptor complex induces the activation of a G protein, thereby releasing $G_q\alpha$, which in turn activates phospholipase C.

The DAG product of the phospholipase C–catalyzed reaction activates protein kinase C (PKC). Several PKC activities have been identified. Examples include acetylcholine-induced gastric acid secretion and serotonin-triggered platelet aggregation. Depending on the cell, activated PKC phosphorylates specific regulatory enzymes, thereby activating or inactivating them. DAG is also a source of arachidonic acid, a precursor in the synthesis of prostaglandins (p. 406) and the endocannabinoid anandamide.

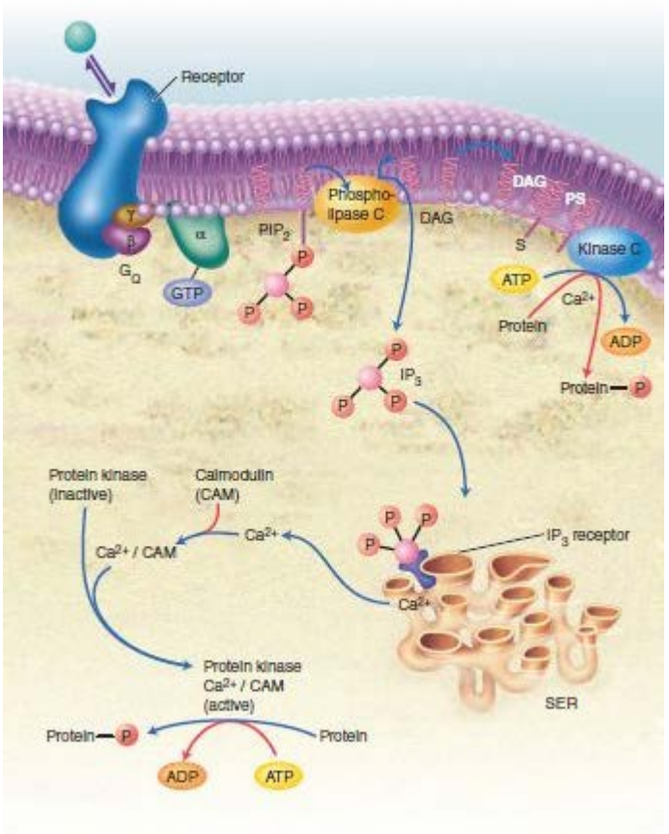


FIGURE 16.8

The Phosphatidylinositol Pathway

The binding of certain hormones to their receptor activates the α subunit of a G_q protein. The α subunit then activates phospholipase C, which cleaves IP_3 from PIP_2, leaving DAG in the membrane. DAG, acting with phosphatidylserine (PS) and Ca^{2+}, activates protein kinase C (PKC), which subsequently phosphorylates key cell signal regulators. IP_3 binds to receptors on the SER, opening Ca^{2+} channels. Then Ca^{2+} moves into the cytoplasm and modulates downstream signaling proteins.

Once generated, IP_3 diffuses to the calcisome (SER), where it binds to the IP_3 receptor (a

calcium channel). Cytoplasmic calcium levels then rise as calcium ions flow through the activated open channel. Calcium ions are involved in the regulation of a large number of cellular processes, including contributing to the activation of plasma membrane-associated PKC. Because calcium levels are still relatively low (approximately 10^{-6} M) even when the calcium release mechanism has been activated, the calcium-binding sites on calcium-regulated proteins must have a high affinity for the ion. Several calcium-binding proteins modulate the activity of other proteins in the presence of calcium. Calmodulin, a type of calcium-binding protein, mediates many calcium-regulated reactions. In fact, calmodulin is a regulatory subunit for some enzymes (e.g., phosphorylase kinase, which converts phosphorylase b to phosphorylase a in glycogen metabolism).

GPCR REGULATION GPCRs are complex and powerful signaling machines that are largely responsible for the cell specialization required by mammals and other multicellular organisms. Rigorous GPCR regulation protects diverse developmental and survival processes and, in the absence of mutations, prevents a wide variety of diseases, most notably malignant transformations (cancers).

The basic GPCR regulatory mechanism is terminating the activity of Gα subunits by GTP hydrolysis facilitated by GAPs. The now GDP-bound Gα subunit rebinds to the $\beta\gamma$ dimer, momentarily ending the GPCR signal. The next phase of GPCR regulation is GPCR desensitization. This process is initiated by the phosphorylation of certain serine and threonine residues in the protein's cytoplasmic segments, in most cases the carboxy-terminal tail and internal loop 3 (IL-3). The second messenger-dependent kinases such as PKA and PKC may phosphorylate some of these residues. However, a class of enzymes referred to as G protein-coupled receptor kinases (GRKs) catalyze the majority of the phosphorylation reactions that facilitate the recruitment and binding of β-arrestin proteins to the receptor. *β-Arrestins* uncouple GPCRs from their G proteins (steric inhibition) and proceed to initiate receptor endocytosis (desensitization) by recruiting proteins that facilitate receptor clustering into clathrin-coated pits (see **Figure 2.18**).

GPCRs undergoing endocytosis can have different cellular fates (resensitization or lysosomal degradation) that depend on receptor–β-arrestin binding strength. Low-affinity binding GPCRs include α- and β-adrenergic, cannabinoid, and chemokine receptors. For these proteins, after β-arrestin release, serine and threonine dephosphorylation and the presence of a carboxy-terminal tail recycling signal sequence, there is a fairly rapid receptor recycling back to the plasma membrane, where they can again bind ligands. Depending on signal sequences and other factors, stable receptor–β-arrestin binding GPCRs such as those for glucagon, glucagon-like peptide, and growth hormone releasing-hormone are slowly recycled from an endosome to the plasma membrane, or they may be diverted to lysosomes for degradation.

Enzyme-Linked Receptors

Enzyme-linked cell-surface receptors initiate signal transduction via intrinsic enzymatic activity or recruitment of intracellular enzymes. Two prominent examples are guanylate cyclase and tyrosine kinase receptors.

Guanylate CYCLASE RECEPTORS Guanylate cyclase proteins occur in two forms: membrane-bound and soluble (cytoplasmic). When activated, both convert GTP to the second messenger molecule *cyclic GMP* (cGMP).

Two types of molecule are now known to activate membrane-bound guanylate cyclase A (GC-A): atrial natriuretic peptide and the guanylin peptides. *Atrial natriuretic factor* (ANF), a peptide

that is released from heart atrial cells in response to increased blood volume, causes diuresis (increased urine production) and natriuresis (Na^+ excretion in urine) and lowers blood pressure. ANF activates GC-A in several cell types, resulting in the activation of a cGMP-dependent protein kinase (PKG). In the kidney's collecting tubules, ANF-stimulated PKG activation results in the inhibition of Na^+ channels and Na^+K^+-ATPase in tubule cell membranes, thereby depressing sodium reabsorption from the urinary filtrate. ANF also decreases blood pressure by activating cGMP-dependent PKG in vascular smooth muscle cells that results in vasodilation.

Guanylin and a closely related peptide called uroguanylin are intestinal hormones released in response to high salt intake. Specialized intestinal epithelial cells regulate local as well as systemic fluid and electrolyte homeostasis. When guanylin peptides bind to GC-C receptors on the lumenal surface of these cells, cGMP concentrations increase and PKG is activated. Subsequent phosphorylation of the CFTR by PKG results in Cl^- and HCO_3^- secretion into the intestinal lumen. cGMP-mediated inhibition of a cAMP hydrolyzing phosphodiesterase results in cAMP accumulation. As a result of cAMP activation of PKA, additional Cl^- is released through the CFTR and Na^+ absorption via the Na/H^+ exchange channel is inhibited. Electrolyte secretion is accompanied by water release.

Enterotoxin (produced by several bacterial species) causes diarrhea by binding to the GC-C receptors of intestinal cells. For example, one form of traveler's diarrhea is caused by a strain of *E. coli* that produces *heat-stable enterotoxin*. When this toxin binds to a GC-C receptor, there is excessive secretion of electrolytes and water into the lumen of the small intestine.

Traveler's Disease

Soluble guanylate cyclase (sGC) is a heterodimer that is activated when nitric oxide (NO•) binds to its heme prosthetic groups (one heme per dimer). Once it is activated, sGC regulates cGMP-dependent protein kinases and ion-gated channels. For example, blood pressure decreases as a result of •NO-triggered cGMP synthesis in arterial smooth muscle cells. Vasodilation, as a result of sGC activation, is caused by a cGMP-mediated reduction in intracellular calcium.

RECEPTOR TYROSINE KINASES **Receptor tyrosine kinases** (**RTKs**) are a family of transmembrane receptors that bind ligands such as insulin, EGF, platelet-derived growth factor (PDGF), and insulin-like growth factor I (IGF-I). Although there are several structural differences among members of this group, they do possess features in common: an external domain that binds specific extracellular ligands, a transmembrane segment, and a cytoplasmic catalytic domain with tyrosine kinase activity. When a ligand binds to the external domain, a conformational change in the receptor protein activates the tyrosine kinase domain. The tyrosine kinase activity initiates a phosphorylation cascade that begins with an autophosphorylation of the tyrosine kinase domain. Most research efforts have been devoted to the insulin receptor.

The *insulin receptor* (**Figure 16.9**) is a transmembrane glycoprotein composed of two types of subunit connected by disulfide bridges. Two large α subunits (130 kDa) extend extracellularly, where they form the insulin-binding site. Each of the two β subunits (90 kDa) contains a transmembrane segment and a tyrosine kinase domain.

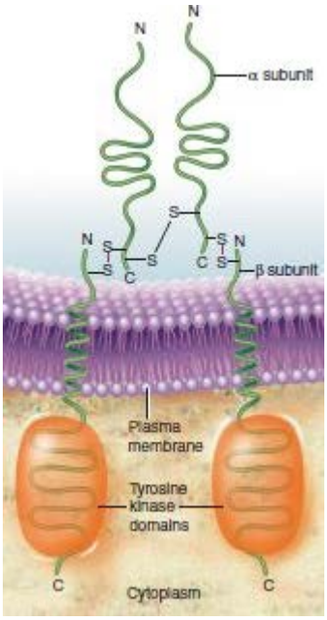


FIGURE 16.9

The Insulin Receptor

The insulin receptor is a tetramer composed of two pairs of α and β subunits. The subunits are connected to each other by disulfide bridges.

The binding of an insulin molecule to the α subunits activates receptor tyrosine kinase activity, which in turn causes several phosphorylation cascades that modulate the activities of numerous intracellular proteins. Insulin also alters the expression of more than 150 genes. Insulin receptor substrate 1 (IRS-1), one of several IRS proteins, is among the best-researched examples of proteins directly phosphorylated (on its tyrosine residues) by the insulin receptor. IRS-1 is a type of *docking protein*, a noncatalytic signaling protein associated with a receptor that binds proteins that initiate signal cascades. Among these proteins is phosphatidylinositol-3-kinase (PI3K) (**Figure 16.10a**). Once it is activated, IRS-1 binds to PI3K, which then proceeds to phosphorylate PIP_2 (a minor plasma membrane phospholipid) to form PIP_3 (phosphatidylinositol 3,4,5-trisphosphate) and DAG. (Do not confuse PIP_3 with IP_3: see p. 609.) Once PIP_3-dependent protein kinase (PDK1) has bound PIP_3, it then activates several kinases. Among the most important of these is PKB (also known as Akt), a serine/threonine kinase that has a central role in cell signal transduction mechanisms (**Figure 16.10b**). Activities of PKB include stimulation of glycogen synthesis (via inhibition of GSK3: see p. 323) and inhibition of lipolysis (via inhibition of PKA activity). PKB also facilitates glucose transport into adipocytes and muscle cells by stimulating the translocation of the GLUT4 transporter (p. 619) to the plasma membrane. The PI3K/PKB pathway is controlled by PTEN, a phosphatase that opposes PI3K by dephosphorylating PIP_3 to yield PIP_2.

The most prominent role of PKB is the activation of mTOR, a component of mTORC1 (mechanistic target of rapamycin complex 1). The mTORC1 complex is a central kinase sensor that integrates hormonal activity, nutrient availability, energy status, and cell responses to various forms of stress (e.g., osmotic, oxidative, and inflammatory). Examples of activated mTORC1-affected gene expression changes include increased ribosome and protein synthesis and decreased autophagy (p. 574). SREBP-1c (p. 299) and PPARγ (p. 471) are two major transcription factors that are activated by mTORC1. SREBP-1c, in combination with ChREBP (carbohydrate response element binding protein) (p. 300), causes the expression of lipogenic genes such as those coding for FAS and ACC (p. 461). SREBP-1c also suppresses gluconeogenic enzyme synthesis. Activated PPARγ (p. 471) stimulates the expression of lipogenic genes.

mTORC1/insulin-driven processes are modulated by at least two mechanisms. First, mTORC1 is part of an autoregulatory pathway that results in inhibitory serine phosphorylations on IRS-1, so as to inhibit PI3K activity. Second, AMPK, activated in response to low energy and nutrient levels or other cell stressors, represses mTORC1 via phosphorylation of an mTOR regulatory protein.

Insulin/RTK signaling promotes cell growth, differentiation, proliferation, and survival via the Ras/MAP kinase pathway (p. 742), which ultimately results in significant cellular gene expression changes. MAP (mitogen-activated protein) kinases are serine/threonine-phosphorylating enzymes that occur in three-tier signaling modules that form an activation cascade in which MAP kinase kinase kinase (MAP3K) phosphorylates the next enzyme, MAP kinase kinase (MAP2K), which then phosphorylates MAP kinase (MAPK). The specific names of the enzymes in the insulin-initiated pathway are Raf (MAP3K), MEK (MAP2K), and ERK1/2 (MAPK). The RAS/MAP kinase pathway is initiated when insulin binding to its RTK results in IRS-1 phosphorylation. The binding of Grb2, an adapter protein, to an IRS-1 domain containing a phosphorylated tyrosine, allows the binding of a GEF that catalyzes the replacement of GDP with GTP on Ras, a small monomeric G protein. Ras-GTP then begins the MAP kinase cascade. Once MAPK is phosphorylated, it enters the nucleus, where it phosphorylates specific proteins that regulate gene expression. (The activation of the Ras/MAP kinase pathway in response to epidermal growth factor binding to its target cell receptor is illustrated in **Figure 18.55** on p. 742.)

KEY CONCEPTS

- There are two major types of cell-surface receptor that bind to hormone molecules: G protein-coupled receptors and receptor tyrosine kinases.
- The G proteins activated by GPCRs utilize one or more second messenger molecules to transduce the original signal into a signal cascade.
- Receptor tyrosine kinases activate phosphorylation cascades when they undergo autophosphorylation triggered by the binding of signal molecules.

QUESTION 16.2

In a hypothetical cAMP-mediated signal transduction cascade, the GTP-α_s/adenylate cyclase interaction following a single hormone–receptor binding event lasts for 2.3 seconds. The catalytic rate (turnover number) for the adenylate cyclase in question is 350 cAMP molecules produced per second. How many cAMP molecules would be produced if five hormone-receptor binding events were to occur before the hormone molecule dissipates in the bloodstream? What is the amplification effect of this step in the signaling pathway?

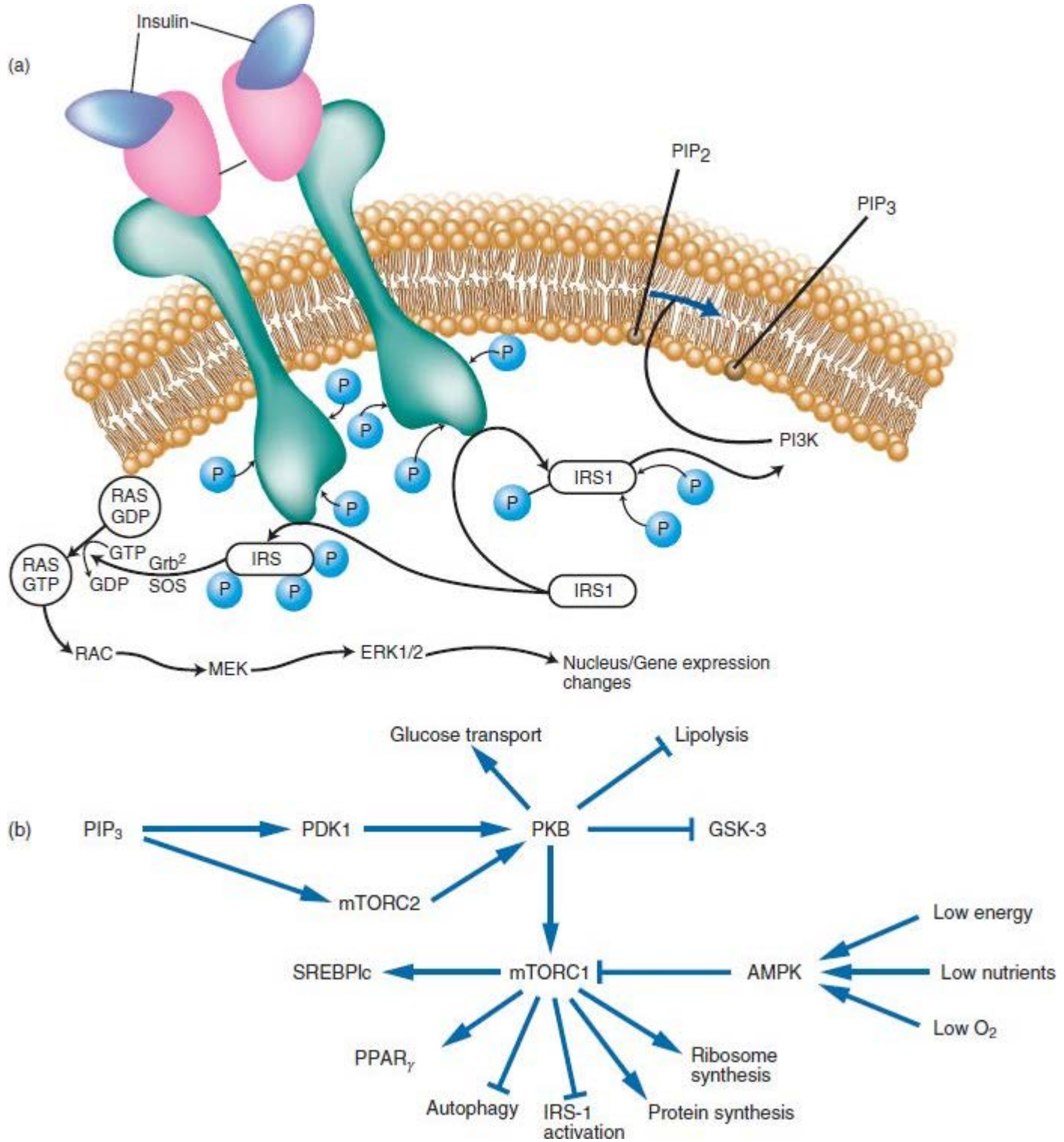


FIGURE 16.10

Simplified Model of Insulin Signaling

(a) The binding of insulin to its tyrosine kinase receptor causes autophosphorylation reactions. The activated insulin receptor subsequently phosphorylates an initial set of substrate molecules. Only one is illustrated in this figure: insulin receptor substrate 1 (IRS1). The newly phosphorylated IRS1 then binds to and activates phosphoinositol-3-kinase (PI3K), which then phosphorylates PIP_2 to form PIP_3. PIP_3 subsequently binds to and activates PIP_3-dependent protein kinase (PDK1), which in turn activates, via phosphorylation, various kinases (e.g., PKB and PKC). These latter molecules continue the signal cascade, ending with alterations in gene expression. The activated insulin receptor also activates the MAP kinase pathway, a cascade of enzymes that regulates the transcription of genes involved in cell growth and differentiation (not shown; refer to **Figure 18.55**). Grb2 is an adaptor protein that binds the GEF protein SOS that subsequently activates the G protein Ras. Ras proceeds to bind to and activate Raf (a MAP kinase kinase kinase). After its activation, Raf

phosphorylates Mek (a MAP kinase kinase). When Mek then phosphorylates ERK (a MAP kinase), the now activated enzyme moves into the nucleus where it regulates transcription factors via phosphorylation reactions. (b) PKB (Akt) stimulates glycogen synthesis by phosphorylating (and thereby inactivating) GSK3, an enzyme that inactivates glycogen synthase. PKB also inhibits lipolysis and activates the transfer of GLUT4 (a glucose transporter protein) to the plasma membrane and the activity of mTOR in the mTORC1 complex, which regulates numerous cellular processes. Examples of mTORC1-affected processes shown in this figure include protein and ribosome synthesis (stimulated) and autophagy (inhibited). mTORC1 is also known to affect the expression of certain genes. For example, mTORC1 promotes, via the transcription factors SREBP1c and PPARγ, the expression of genes that code for proteins involved in lipid and cholesterol synthesis. mTORC1 activity is inhibited by AMPK when the levels of ATP, nutrients, and O_2 are low. When nutrient levels are high, mTORC1 modulates its own activity by activating an IRS1 and PI3K inhibitor, which reduces the activity of the insulin-IRS1-PI3K signal cascade. mTORC2, activated in part by PIP_3, also influences insulin signaling processes. For example, mTORC2 phosphorylation of PKB promotes the enzyme's full activation.

QUESTION 16.3

Explain the sequence of events that occurs when epinephrine triggers the synthesis of cAMP. Once formed, cAMP breaks down rapidly. Why is this an important feature for a second messenger in a signal transduction process?

QUESTION 16.4

The A subunit of cholera toxin causes the cAMP-mediated opening of chloride channels. A massive diarrhea results because the GTP hydrolysis of GTP-α_s is prevented. Describe why this inhibition leads to the diarrhea.

QUESTION 16.5

Cancer often results from a multistage process involving an initiating event (mediated by a viral infection or a carcinogenic chemical), followed by exposure to tumor promoters. Tumor promoters, a group of molecules that stimulate cell proliferation, cannot induce tumor formation by themselves. The phorbol esters, found in croton oil (obtained from the seeds of the croton plant, *Croton tiglium*), are potent tumor promoters. (Other examples of tumor promoters include asbestos and several components of tobacco smoke.) In one of the tumor-promoting actions of the phorbol esters, these molecules mimic the actions of DAG. In contrast to DAG, the phorbol esters are not easily disposed of. Explain the possible biochemical consequences of phorbol esters in an "initiated" cell. What enzyme is activated by both DAG and phorbol esters?

Tumor Promoters

QUESTION 16.6

The term *diabetes*, derived from the Greek word *diabeinein* ("to go to excess"), was first used

by Aretaeus (CE 81–138) to identify a group of symptoms that included intolerable thirst and "a liquefaction of the flesh and limbs into urine." After reviewing the accompanying Biochemistry in Perspective essay (pp. 615–17), explain the physiological and biochemical basis for Aretaeus's findings.

Diabetes

Nuclear Receptors

Nuclear receptors are ligand-activated transcription factors that bind hydrophobic molecules such as steroid and thyroid hormones. These receptors share a common structure with an N-terminal transcription regulation domain, a DNA-binding domain, and a ligand-binding domain. Once they have reached their target cells (by binding to specific proteins), hydrophobic hormone molecules dissociate from their transport proteins, diffuse through the plasma membrane, and bind to their intracellular receptors (**Figure 16.11**). Depending on the type of hormone involved, initial binding to receptors may occur within the cytoplasm (e.g., glucocorticoid) or the nucleus (e.g., estrogen, androgens, and thyroid hormone). In the absence of hormone, several types of receptor have been observed to form complexes with other proteins. For example, unoccupied glucocorticoid receptors are found in the cytoplasm bound to chaperone proteins, such as hsp90 (p. 171). Chaperone proteins block the receptor's DNA-binding site when the hormonal ligand is not present. When the hormone binds to its receptor, the chaperones dissociate and the receptor–ligand complex migrates to the nucleus as a homodimer.

KEY CONCEPT

Hydrophobic hormones such as the steroids and thyroid hormones diffuse across cellular membranes and bind to intracellular receptors.

Within the nucleus, each hormone–receptor complex binds to specific DNA segments called **hormone response elements** (HREs). Depending on the presence of cell-type-specific coregulator proteins (activators or inhibitors), hormone-receptor complexes binding to an HRE either enhance or diminish the transcription of a specific gene. The same hormone-receptor complex can bind to and influence the transcription of as many as 50 to 100 different genes, thereby inducing global changes in cellular function.

Biochemistry IN PERSPECTIVE

Diabetes Mellitus

Why does diabetes mellitus, a disease in which glucose transport is

compromised, damage the entire body? Diabetes mellitus is a group of devastating metabolic diseases caused by insufficient insulin synthesis, increased insulin destruction, or ineffective insulin action. There are two major forms of diabetes: type 1 and type 2. **Type 1 diabetes** (previously referred to as *juvenile-onset* or *insulin-dependent diabetes*) is an autoimmune disease caused by the destruction of the insulin-producing pancreatic *β*-cells. **Type 2 diabetes** (previously referred to as *adult-onset* or *non-insulin-dependent diabetes*) is caused by insensitivity of target tissues to insulin. Once quite rare, diabetes is now a leading cause of death in the United States, where it afflicts at least 10% of the population and is the 7th leading cause of death. The worldwide incidence of diabetes is 8.6% in adults (ages 18–99).

The basic feature of diabetes is dysfunctional fuel metabolism. Insulin deficiency or the insensitivity of insulin's target tissues (muscle, adipose tissue, and liver) results in **hyperglycemia** (high blood glucose levels) and **dyslipidemia** (abnormal blood lipid and lipoprotein levels). Without effective insulin action, blood glucose levels are higher than normal because muscle and adipose tissue fail to absorb it. This circumstance is exacerbated in the liver by gluconeogenesis and glycogenolysis (ordinarily suppressed by insulin), which produce additional glucose for delivery into an already hyperglycemic bloodstream. Without effective insulin action, lipid metabolism also becomes impaired, especially in liver and adipose tissue.

Hyperglycemia is the proximate cause of the acute symptoms seen in all forms of diabetes. Extreme thirst and frequent urination are caused by **glucosuria** (glucose in urine), which leads to **osmotic diuresis** [excessive loss of water and electrolytes (Na^+, K^+, and Cl^-)]. Extreme fatigue results from the inability of cells to generate sufficient energy. The body's glucose-starved cells trigger a hunger response in the brain's appetite center (p. 624), causing extreme hunger (polyphagia). Hyperglycemia also activates several processes that lead to long-term damage to the body. The formation of advanced glycation end products via the glycation of proteins (p. 256) damages the endothelial cells that line blood vessels and other cell types, thereby contributing to atherosclerosis and other degenerative diseases.

High blood glucose also stimulates the sorbitol pathway. In some cells in which glucose uptake is insulin-independent (e.g., peripheral nerves and the lens of the eye), excess glucose molecules are converted into sorbitol (p. 253) by NADPH-requiring *aldose dehydrogenase*. Sorbitol accumulation causes glycation of intracellular proteins as well as osmotic stress. Excess $NADP^+$ can result in lower cellular levels of GSH (p. 549) and NO• (p. 552). The oxidation of some sorbitol molecules is coupled to NAD^+ reduction. In addition to increasing lactate synthesis (from pyruvate), excess NADH also stimulates superoxide production by the mitochondrial electron transport system and activation of the superoxide-producing enzyme *NADH oxidase*. In diabetics, the accumulation of sorbitol and the redox changes are associated with nerve damage and cataract formation. Finally, hyperglycemia is one of several factors that cause a chronic systemic inflammatory process, mediated by proinflammatory cytokines that activate a network of inflammatory signaling pathways.

In the absence of effective insulin action, there is increased lipolysis (Section 12.1) in adipose tissue (caused by the unopposed action of glucagon), which releases large quantities of fatty acids into blood. In the liver, because these molecules are degraded by *β*-oxidation in combination with low concentrations of OAA (caused by excessive gluconeogenesis), large amounts of acetyl-CoA, the substrate for forming ketone bodies, are produced. Fatty acids not used to generate energy or ketone bodies are used in VLDL synthesis. This process causes *hyperlipoproteinemia* (high blood concentrations of lipoproteins) because lipoprotein lipase synthesis is depressed when insulin is lacking.

Type 1 Diabetes

In most cases of type 1 diabetes, autoimmune destruction of the insulin-producing *β*-cells in the pancreas is the result of genetic and environmental factors. The symptoms tend to appear abruptly when almost all insulin-producing capacity has been destroyed, the result of an ongoing inflammatory process over several months or years. As in other inflammatory and autoimmune processes, *β*-cell destruction is initiated when an autoantibody binds to a cell-surface antigen. One of the most common autoantibodies found in type I diabetes is now believed to bind specifically to an antigen with glutamate decarboxylase activity. Autoantibodies to insulin and the tyrosine phosphatase IA-2 have also been detected.

The most serious acute symptom of type I diabetes is **ketoacidosis**, which is the result of unrestrained fatty acid oxidation. Ketone bodies are released in such large amounts that the body's capacity to oxidize them is exceeded. Elevated concentrations of ketones in the blood (**ketosis**) and low blood pH (acidosis) along with hyperglycemia cause excessive water losses. (The odor of acetone on a patient's breath is characteristic of ketoacidosis because the major mechanism for the removal of this volatile molecule is through the lungs.) Ketoacidosis and dehydration, if left untreated, can lead to coma and death. Type 1 patients are treated with injections of insulin synthesized using recombinant DNA technology. Before Frederick Banting and Charles Best discovered and purified insulin in 1922, most type 1 diabetics died within a year after being diagnosed. Although exogenous insulin prolongs life, it is not a cure. Most diabetics have a shortened life span because of the long-term complications of their disease.

Type 2 Diabetes

Because its onset is gradual, type 2 diabetes appears to be milder than type 1 diabetes, yet its long-term effects on the body are just as devastating. Unlike type 1 diabetics, people with type 2 diabetes tend to have elevated blood levels of insulin at the time of diagnosis and, for various reasons, are resistant to insulin. Type 2 diabetics have high blood levels of inflammatory proteins such as C-reactive protein (a liver protein secreted into blood in response to inflammation) and interleukin-1β (IL-1β) and IL-6.

The insensitivity of target tissues to insulin in type 2 diabetics causes blood glucose levels to rise, leading to increased release of insulin from pancreatic *β*-cells. The increased exposure of target cell insulin receptors to insulin promotes receptor internalization and decreased synthesis of the receptor and several downstream signaling proteins. One of the eventual consequences of high blood insulin levels (**hyperinsulinemia**) is compromised function and reduced mass of the insulin-secreting *β*-cells. The formation of amyloid deposits, aggregates of a misfolded protein, in islet cells may be a contributing factor to *β*-cell apoptosis. (Disease-related protein aggregation is described on p. 786.)

The onset of type 2 diabetes occurs when the insulin response falls to a level where fasting blood glucose levels exceed 126 mg/dl (normal = <100 mg/dl). Approximately 85% of type 2 diabetics are obese. Because obesity itself promotes tissue insensitivity to insulin, individuals who are prone to this form of diabetes are at increased risk for the disease when they gain weight.

INSULIN RESISTANCE, TYPE 2 DIABETES, AND OBESITY Insulin resistance is a common feature of obesity, atherosclerosis, and nonalcoholic fatty liver disease, in addition to type 2 diabetes. The proximate cause of insulin resistance is disruption of insulin receptor-catalyzed tyrosine phosphorylation of IRS by several kinases (e.g., PKC and JNK). ER stress and oxidative stress also initiate and promote numerous inflammatory processes.

Obesity contributes to type 2 diabetes because excess lipid accumulation in adipose tissue, combined with cell stresses such as high fat diets, oxidative stress, and ER stress, results in low-grade inflammation. The inflammatory process causes adipocytes to release proinflammatory cytokines such as tumor necrosis factor-α (TNF-α) and IL-6.

TNF-α, IL-6, and similar proteins elicit adipose tissue infiltration by macrophages and other white blood cells, which then also secrete inflammatory proteins. Proteins such as TNF-α induce the phosphorylation of IRS-1 serine residues, thereby interfering with insulin signaling (e.g., causing depressed glucose transport and increased lipolysis). Inflammation worsens when macrophages and adipocytes activate JNK1/2, which are MAPK enzymes that increase production of other proinflammatory factors. When combined with depressed PKB activity, IL-6 induces adipocyte lipolysis, a process that results in the release of large amounts of fatty acids into the bloodstream. The fatty acids released by adipocyte lipolysis have an impact on the insulin sensitivity of other tissues, such as skeletal muscle and liver.

Two major processes promote insulin resistance in skeletal muscle: (1) fatty acid overload–mediated impairment of insulin/insulin receptor signaling and (2) TG accumulation in muscle cells. High levels of fatty acids drive TG synthesis in skeletal myocytes, causing levels of diacylglycerol (DAG), an intermediate in TG synthesis, to increase. Among the actions of DAG is the activation of an isoform of PKC referred to as PKCθ. By catalyzing the serine phosphorylation IRS1, PKCθ interferes with insulin signaling, resulting in reduced glycogen synthesis and GLUT4 transport to the myocyte plasma membrane. TG accumulation in myocytes is caused by sustained lipid overload and reduced efficiency of mitochondrial fatty acid oxidation, induced in part by ROS-induced damage.

Hepatic insulin resistance is caused by high blood concentrations of fatty acids and glycerol (both released by enlarged adipocytes) and glucose (diverted from insulin-resistant skeletal muscle cells). Excess fatty acids and glycerol drive the synthesis of TG, many of which are then exported in VLDL. If TG synthesis exceeds hepatocyte VLDL export capacity, excess TGs form potentially cytotoxic lipid droplets. Increased DAG levels in hepatocytes also result in PKCε activation, which interferes with insulin signaling, resulting in depressed glycogenesis. Increased fatty acid oxidation increases acetyl-CoA levels, which lead to pyruvate carboxylase (PC) activation. PC activation combined with high glycerol levels and depressed insulin signaling results in stimulating gluconeogenesis.

TREATMENT OF TYPE 2 DIABETES Type 2 diabetes may be treated with diet control and exercise. Obese patients often become more sensitive to insulin (i.e., there is an upregulation of insulin receptors) when they lose weight. Because sustained muscular activity increases the uptake of glucose without requiring insulin, exercise also decreases hyperglycemia.

Medications may also be effective in the treatment of type 2 diabetes. Oral hypoglycemic agents such as the sulfonylureas stimulate insulin release and, therefore, reduce gluconeogenesis and glycogenolysis in the liver while increasing transport of glucose into insulin-sensitive body cells. Metformin, a biguanide antidiabetic drug, inhibits hepatic gluconeogenesis and promotes peripheral glucose uptake and fatty acid oxidation by activating AMPK.

Failure to control hyperglycemia in type 2 diabetics with other serious medical conditions (e.g., renal insufficiency, myocardial infarction, or infections) can cause a serious metabolic state referred to as **hyperosmolar hyperglycemic nonketosis** (HHNK). (Ketoacidosis is rare in type 2 diabetes.) The additional metabolic stress exacerbates insulin resistance, and blood glucose levels rise. The patient may then become severely dehydrated. The resulting lower blood volume depresses renal function, which causes further increases in blood glucose concentrations.

Eventually, the patient becomes comatose. Because the onset is slow, it may not be recognized until the dehydration is severe. (This is especially true for elderly diabetics, who often have a depressed thirst mechanism.) For this reason, HHNK is often more life-threatening than ketoacidosis.

Type 3 Diabetes

In recent years, researchers have discovered that there is a connection between peripheral insulin resistance and the risk of Alzheimer's disease (AD) and that insulin has vital functions in the brain. Insulin readily enters the brain by a transport mechanism across the *blood-brain barrier* (BBB) involving insulin receptors. (The BBB is a specialized, protective cell network that restricts the movement of proteins and drug molecules into the central nervous system.) Insulin receptors are expressed by neurons and glial cells in specific areas of the brain, most notably in those structures involved in learning, reasoning, and memory (hippocampus, cerebellum, amygdala, and the cerebral cortex), and food intake suppression (hypothalamus). In contrast to peripheral cells such as skeletal myocytes and adipocytes, insulin's effect on glucose transport into brain cells is negligible. (Insulin-insensitive GLUT1 and GLUT3 are the major glucose transporter proteins. Insulin-sensitive cells have low levels of GLUT4 in addition to GLUT3.) Instead, insulin's effects on brain involve diverse functions that are controlled by IRS proteins (p. 611) and MAPK pathways (p. 612).

Alzheimer's disease (p. 436) is an irreversible, progressive brain disorder that gradually destroys memory and other cognitive skills and causes personality changes. Peripheral insulin resistance associated with poorly controlled type 2 diabetes or obesity is a major AD risk factor primarily because of *brain insulin resistance*, which results from the effects of inflammatory cytokines and hyperinsulinemia. (Other risk factors include aging, the *APOE4* gene [p. 437], poor diet, and sedentary lifestyle.) Inflammatory cytokines (TNF-α, IL-1β, and IL-6), synthesized by adipose tissue of obese patients and/or by neural cells in response to diabetic hyperglycemia, contributes to neuroinflammation, whose effects include activation of kinases such as JNK (*J*un *N*-terminal *k*inase) that inhibit insulin signaling, promote tau hyperphosphorylation (p. 436), and increase mitochondrial-damaging ROS formation (p. 385) and ER stress (p. 49). Hyperinsulinemia disrupts the clearance of amyloid β (Aβ) (pp. 436–37) one of the aggregation prone proteins associated with Alzheimers's disease. Insulin degrading enzyme (IDE) degrades both insulin and Aβ, and has an important role in clearing neurotoxic amyloid peptides from the brain. However, hyperinsulinemia contributes to Aβ accumulation because insulin has a higher affinity for the active site of IDE than Aβ, and insulin resistance diminishes insulin-dependent IDE synthesis.

Brain insulin resistance, now referred to as *type 3 diabetes*, begins decades before AD symptoms are observable. The progression of pathological changes in neural networks with high metabolic demands, especially those involved in learning and memory, is not only the result of Aβ and phosphorylated tau accumulation. AD's metabolic alterations are the consequence of several decades of increasingly severe deficits in glucose-derived ATP synthesis caused by insulin resistance, neuroinflammation, and cumulative mitochondrial damage.

Long-Term Complications of Diabetes

Despite the efforts of physicians and patients to control the symptoms of diabetes, few diabetics avoid the long-term consequences of their disease. In addition to AD risk, diabetics are especially prone to develop kidney failure, myocardial infarction, stroke, blindness, and neuropathy (nerve damage resulting in the loss of sensory and motor functions). In addition, circulatory problems

often cause gangrene, which leads to tens of thousands of amputations annually.

Many diabetic complications stem from damage to the vascular system. For example, damaged capillaries in the eye and kidney lead to blindness and kidney damage, respectively. Similarly, the accelerated form of atherosclerosis found in diabetics can lead to myocardial infarction and stroke. Much of this damage results from advanced glycation end products (AGEs). Whether they originate by endogenous glycation reactions (p. 256) or from exogenous sources (ingestion with food or smoking), AGEs initiate atherosclerosis. Another facet of this disease process is hyperglycemia-initiated ROS formation, which causes mitochondria in endothelial cells to produce large quantities of superoxide. Some superoxide species react with NO• to form peroxynitrite ($ONOO^-$; see p. 386). The subsequent nitrosylation of proteins, most notably antioxidant enzymes and NO synthase, leads to increased oxidative stress and a decrease in NO•, a critical vasodilator molecule.

SUMMARY Diabetes is an example of how a single defect (the inability to synthesize or respond to insulin) in a complex biological system can cause devastating damage.

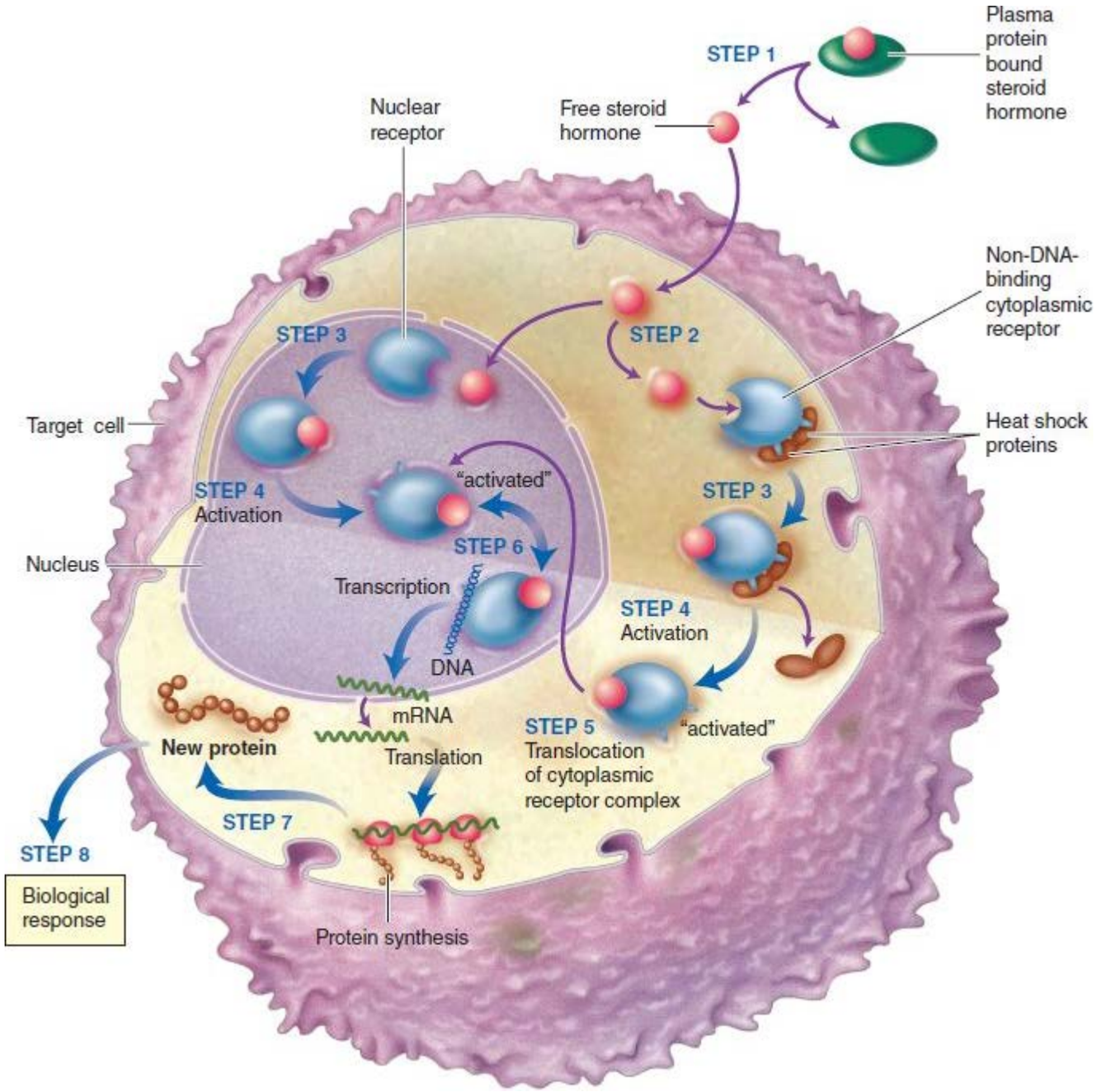


FIGURE 16.11

Model of Steroid Hormone Action within a Target Cell

Steroid hormones are transported in blood associated with plasma proteins. When they reach a cell and are released (1), the hormone molecules diffuse through the plasma membrane, where they bind to receptor molecules in cytoplasm (2) or nucleus (3). After activation (4), a cytoplasmic hormone-receptor complex, a dimer (not shown), migrates to the nucleus (5). The binding of an activated hormone-receptor complex to DNA HRE sequences (6) results in a change in the rate of transcription of specific genes and, therefore, in the pattern of proteins (7) that the cell produces. The net effect of the steroid hormone (8) is a change in the metabolic functioning of the cell.

16.3 METABOLISM IN THE MAMMALIAN BODY: THE FEEDING–FASTING CYCLE

Each of the major metabolic pathways that sustain life in multicellular organisms has now been

covered. Any true understanding of metabolism, however, requires a more integrated approach. A brief review of the feeding–fasting cycle, the self-regulating mechanism by which the mammalian body extracts energy and nutrients from food, provides an opportunity to observe biochemical reactions as they actually occur.

Each organ in the mammalian body contributes to the individual's function in several ways. For example, some organs are consumers of energy so that they may perform certain energy-driven tasks (e.g., brain and skeletal, cardiac, and smooth muscle). Other organs, such as those in the digestive tract, are responsible for efficiently supplying energy-rich nutrient molecules for use elsewhere. Information in the form of signal molecules (e.g., hormones and neurotransmitters) is used to regulate the balance between energy generation and energy expenditure. Prominent examples of protein hormones are ghrelin, peptide YY (PYY), cholecystokinin (CCK), and glucagon-like peptide 1 (GLP-1). **Ghrelin** (ghr), produced by cells in the stomach and small intestine, stimulates appetite (food intake), whereas insulin, PYY, CCK, and GLP-1 promote satiety (i.e., inhibit food intake). Nutrient transport across cell plasma membranes is also an important feature of organ function. Glucose transport is a well-researched example. Active transport of glucose by the Na^+/glucose transporter is linked to a Na^+ gradient established by the ATP-driven Na^+-K^+ pump (p. 429). The diffusion of glucose across cell membranes is facilitated by glucose carriers called GLUTs: GLUT1 (most cells), GLUT2 (liver, β-pancreatic cells, renal tubular cells, and intestinal enterocytes), GLUT3 (neurons), and GLUT4 (insulin-sensitive muscle and adipose tissue cells, and several types of brain neurons). GLUT5, principally found in intestinal enterocyte and liver cell plasma membranes, transports fructose.

Despite their consistent requirements for energy and biosynthetic precursor molecules, mammals consume food only intermittently. This is possible because of elaborate mechanisms for storing and mobilizing energy-rich molecules derived from food (**Figure 16.12**). The changes in the status of various biochemical pathways during transitions between feeding and fasting illustrate metabolic integration and the profound regulatory influence of hormones. Substrate concentrations are also an important factor in metabolism. In the **postprandial** state, which occurs directly after a meal has been digested or absorbed, blood nutrient levels are elevated above those in the fasting phase. During the **postabsorptive** state, for example, after an overnight fast, nutrient levels in blood are low.

The Feeding Phase

As the feeding phase begins, food is propelled along the GI tract by muscle contractions initiated and controlled by the enteric (intestinal) nervous system. As it moves through the digestive organs, food is broken into smaller particles and exposed to enzymes. Ultimately, the products of digestion (consisting largely of sugars, fatty acids, glycerol, and amino acids) are absorbed by the small intestine and transported into the blood and lymph. This phase is regulated by interactions between enzyme-producing cells of the digestive organs, the nervous system, and several hormones. The enteric nervous system, which is influenced by the parasympathetic and sympathetic nerves, is responsible for the waves of smooth muscle contraction that propel food along the tract, as well as for regulating the secretions of several digestive structures (e.g., salivary and gastric glands). Hormones such as gastrin, secretin, and CCK also contribute to the digestive process. They do so by stimulating the secretion of enzymes or digestive aids such as bicarbonate and bile.

The early postprandial state is illustrated in **Figure 16.13**. Sugars and amino acids are absorbed from the small intestine and transported by the portal blood to the liver. The portal blood also contains a high level of lactate, a product of enterocyte metabolism. Most lipid molecules are transported from the small intestine in lymph as chylomicrons. Chylomicrons pass into the

bloodstream, which carries them to tissues such as muscle and adipose tissue. After most triacylglycerol molecules have been removed from chylomicrons, the liver then takes up these structures, now referred to as chylomicron remnants. The phospholipid, protein, cholesterol, and a few remaining triacylglycerol molecules are then degraded or reused. For example, cholesterol is used to synthesize bile acids, and fatty acids are used in new phospholipid synthesis. Phospholipids, as well as other newly synthesized lipid and protein molecules, are then incorporated into lipoproteins by liver cells for export to other tissues.

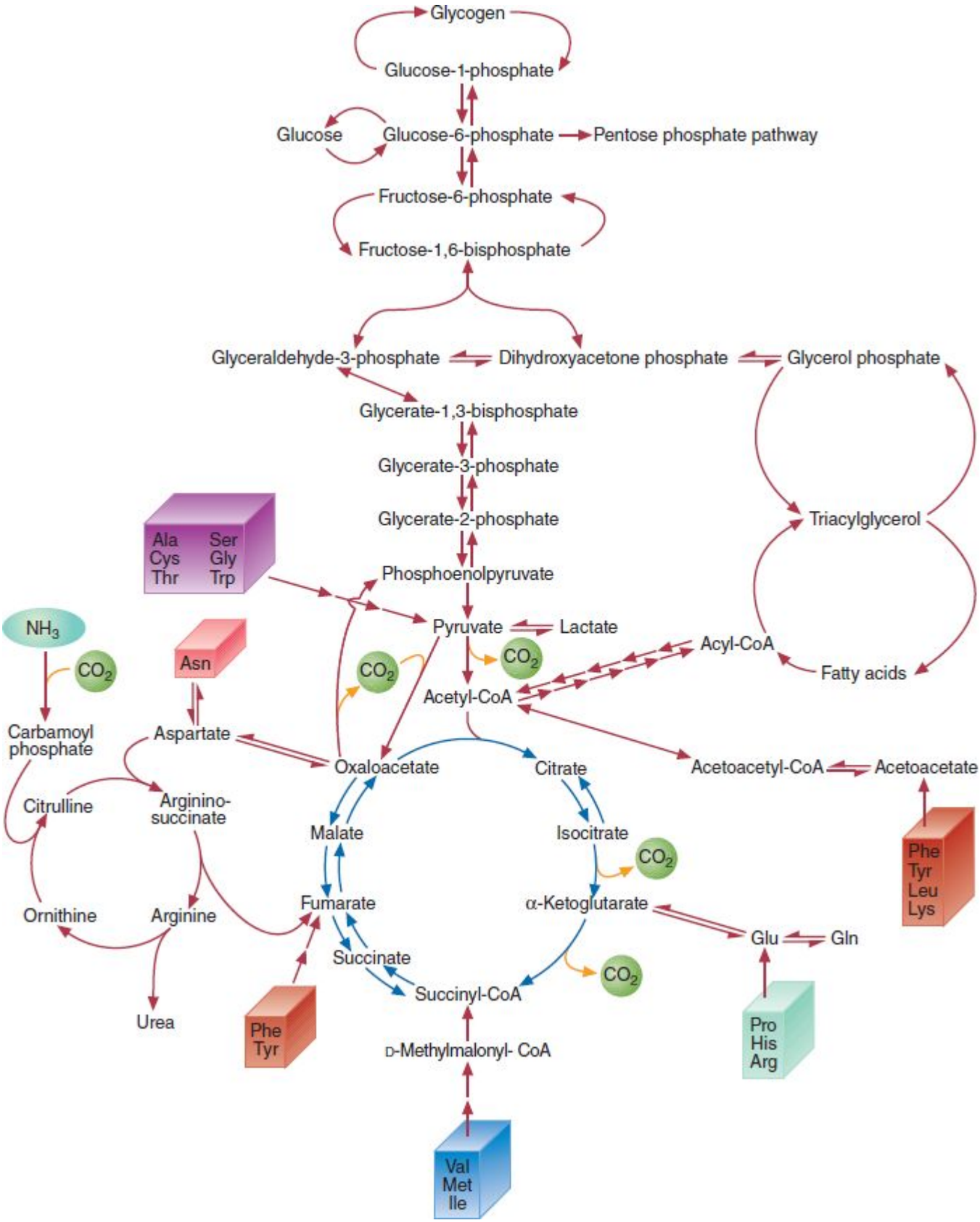


FIGURE 16.12

Nutrient Metabolism in Mammals

Despite the variability of the mammalian diet, these organisms usually provide their cells with adequate nutrients. Control mechanisms that regulate biochemical pathways are responsible for this phenomenon.

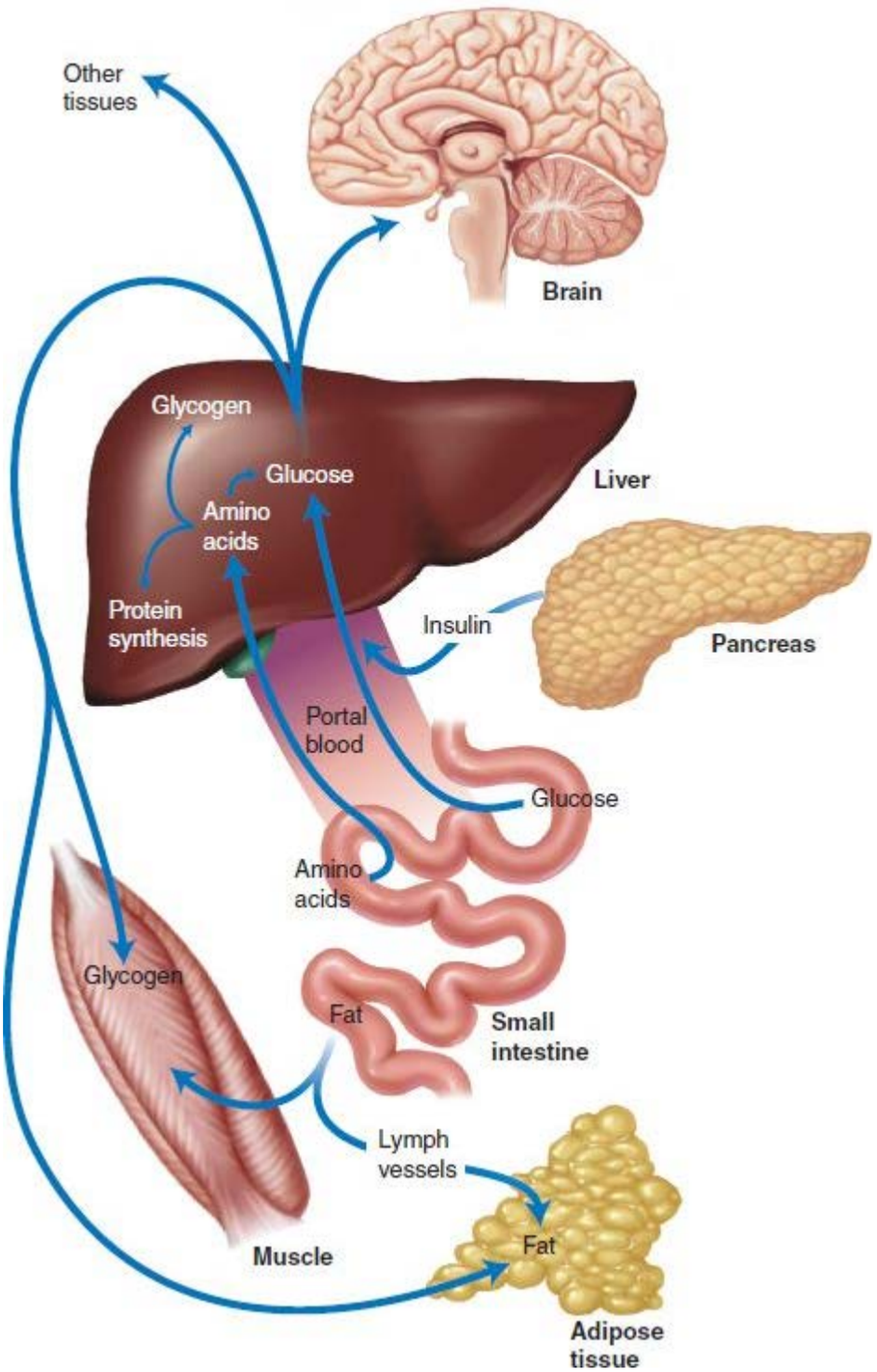


FIGURE 16.13

The Early Postprandial State

The primary substrates for glycogen synthesis in liver are amino acids and lactate (not shown) derived from portal blood. Note that the normal use for glucose in fat cells is as the precursor of glycerol. Fat cells do not carry out significant de novo fatty acid synthesis, but instead obtain most from the diet. Chylomicrons (not shown) carry lipids from the small intestine to the body's tissues, especially muscle and adipose tissue. Note that the brain usually uses glucose as its sole fuel.

As glucose moves through the blood from the small intestine to the liver, pancreatic β cells within the pancreas are stimulated to release insulin. (High blood glucose and insulin levels depress glucagon secretion by pancreatic α cells.) Insulin release triggers several processes that

ensure the storage of nutrients. These include glucose uptake by muscle and adipose tissue, glycogenesis in liver and muscle, fat synthesis in liver, and fat storage in adipocytes. Insulin also represses lipolysis in adipocytes, and gluconeogenesis and glycogenolysis in liver. In addition, insulin influences amino acid metabolism. For example, insulin promotes the transport of amino acids into cells (especially liver and muscle cells). In general, insulin stimulates protein synthesis in most tissues.

Glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic peptide (GIP), referred to as the *incretins*, are released into the bloodstream in response to glucose in the intestinal lumen. Both GLP-1 and GIP augment glucose-stimulated insulin secretion by pancreatic β-cells. In addition, GLP-1 promotes insulin synthesis and inhibits glucagon secretion, and GIP promotes lipoprotein lipase activity (p. 445) and lipogenesis (TG synthesis) in adipose tissue.

Although the effects of insulin on postprandial metabolism are profound, other factors (e.g., substrate supply and allosteric effectors) also affect the rate and degree to which these processes occur. For example, elevated levels of fatty acids in blood promote lipogenesis in adipose tissue. Regulation by several allosteric effectors further ensures that competing pathways do not occur simultaneously. In many cell types, for example, fatty acid synthesis is promoted by citrate (an activator of acetyl-CoA carboxylase), whereas fatty acid oxidation is depressed by malonyl-CoA (an inhibitor of carnitine acyltransferase I activity).

The Fasting Phase

The early postabsorptive state (**Figure 16.14**) of the feeding–fasting cycle begins as the nutrient flow from the intestine diminishes. As blood glucose and insulin levels fall, glucagon is released from pancreatic α-cells. Glucagon prevents hypoglycemia by promoting glycogenolysis and gluconeogenesis in liver. Decreased insulin reduces energy storage in several tissues and leads to increased lipolysis and the release of amino acids such as alanine and glutamine from muscle. Recall that several tissues use fatty acids in preference to glucose. Glycerol and alanine (i.e., the glucose–alanine cycle, p. 306) are substrates for gluconeogenesis, and glutamine is an energy source for enterocytes.

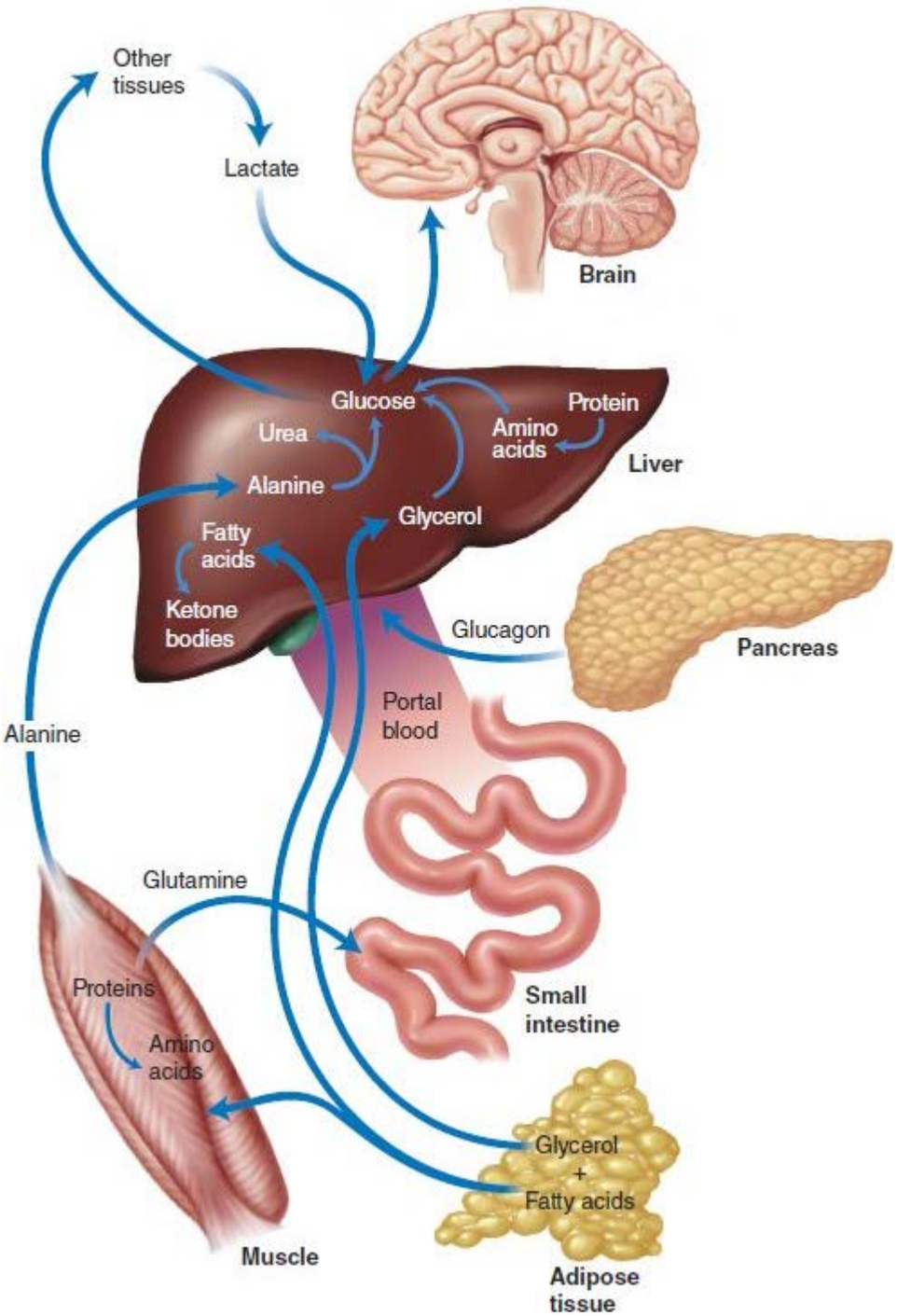


FIGURE 16.14

The Postabsorptive State

Between meals the body, under the influence of hormones, obtains nutrients from skeletal muscle (e.g., alanine for gluconeogenesis in liver and glutamine for energy generation in enterocytes) and adipose tissue (fatty acids). See the text for further details.

When a fast becomes prolonged (e.g., overnight), several metabolic strategies maintain blood glucose levels. Increased mobilization of fatty acids from adipose tissue during the postabsorptive state is stimulated by epinephrine. These fatty acids provide an alternative to glucose for muscle. (Reduced skeletal muscle consumption of glucose spares its use for brain.) In addition, the action of glucagon increases gluconeogenesis, using amino acids derived from muscle.

Under conditions of extraordinarily prolonged fasting (starvation), the body makes metabolic

changes to ensure that adequate amounts of blood glucose are available to sustain energy production in the brain and other glucose-requiring cells. Fatty acids from adipose tissue and ketone bodies from liver are mobilized to sustain the other tissues. Because glycogen is depleted after several hours of fasting, gluconeogenesis plays a critical role in providing sufficient glucose. During early starvation, large amounts of amino acids from muscle are used for this purpose. However, after several weeks, the breakdown of muscle protein declines because the brain is using ketone bodies as a fuel source.

KEY CONCEPTS

- During the feeding phase, food is consumed, digested, and absorbed.
- Absorbed nutrients are then transported to the organs, where they are either used or stored.
- During fasting, several metabolic strategies maintain blood glucose levels.

QUESTION 16.7

Explain the metabolic changes that occur during starvation. What appears to be the principal purpose for the preferential degradation of muscle tissue during starvation?

QUESTION 16.8

Explain the changes in liver metabolism that occur when blood glucose levels drop after a meal has been digested.

QUESTION 16.9

Researchers have observed that an oral dose of glucose elicits a higher release of insulin than does an intravenous injection of glucose. Explain, in general terms, this discrepancy.

Feeding Behavior

Feeding behavior is the complex mechanism by which animals, including humans, seek out and consume food. In mammals, regulation of feeding behavior involves hormonal and neural signals from peripheral organs (e.g., the GI tract and adipose tissue) and sensory input from the external environment (e.g., the sight, smell, and taste of palatable food) that together are integrated in the brain to regulate appetite and the body's metabolic processes (**Figure 16.15**).

For mammals, which have high-energy requirements, the consumption of sufficient food to ensure the energy needed to sustain life is of critical importance. To this end, mammals have evolved a robust food-seeking system involving several neuronal pathways and numerous signaling molecules. In addition to providing a mechanism for balancing energy consumption and use, the mammalian brain links appetite systems to taste, olfaction, and reward systems to create a powerful drive that ensures survival.

Although appetite regulation is still not completely understood, it is clear that the principal neural circuits that control appetite are in the hypothalamus, located in the ventral (underside)

portion of the vertebrate brain, and in the brain stem. Despite its small size, the hypothalamus, one of the most evolutionarily conserved regions of the brain, has a wide array of functions, including control of body temperature, electrolyte balance, monitoring of nutrient levels (e.g., blood glucose), and several aspects of emotional behavior.

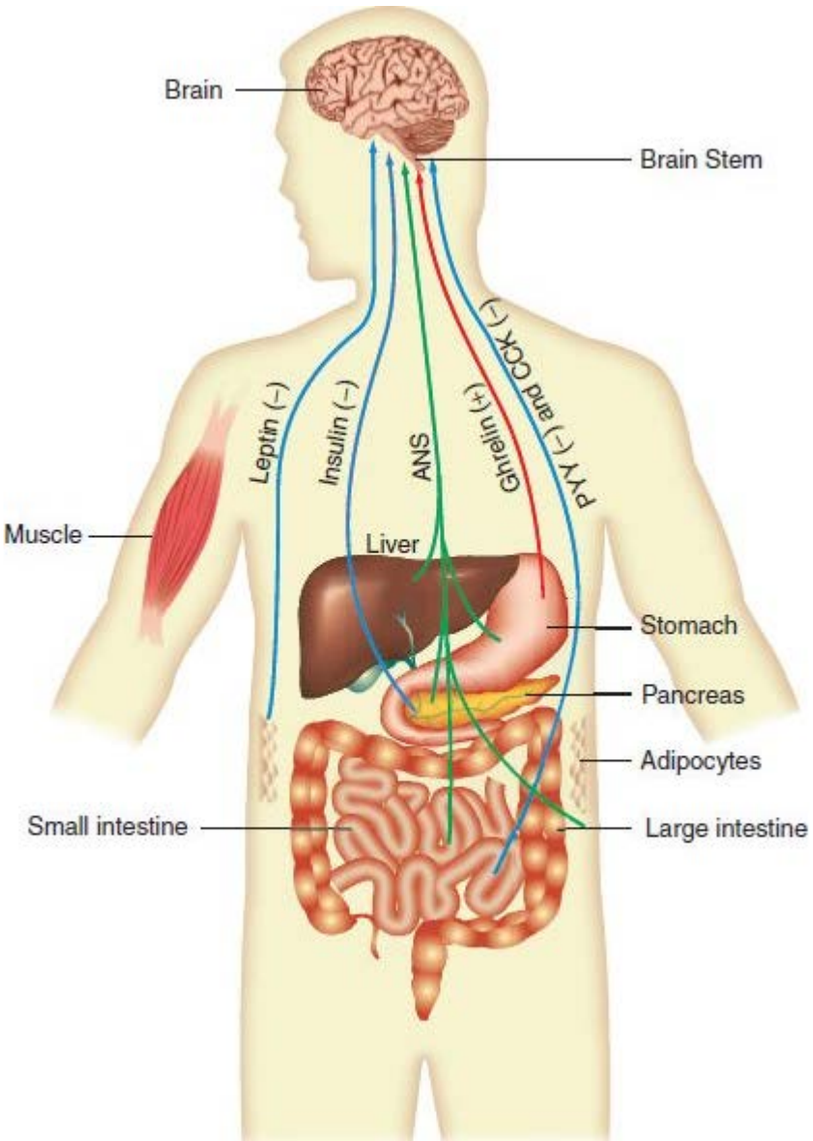


FIGURE 16.15

Feeding Behavior in Humans

Appetite and satiety in humans are regulated by hormonal and neural signals from peripheral organs. Peptide hormones such as PYY and CCK (produced by cells in the GI tract), insulin (produced by pancreatic β-cells), and leptin (produced by adipose tissue) inhibit appetite (−); that is, they promote satiety. Ghrelin (produced by cells in the stomach and small intestine) stimulates appetite (+). Neuronal pathways of the autonomic nervous system (ANS) such as the vagus nerve continuously supply the brain with information related to the status of the body's internal organs.

The primary neurons that control feeding behavior are in the *arcuate nucleus* (ARC) of the hypothalamus (**Figure 16.16**). Activation of ARC neurons that produce NPY (neuropeptide Y) and AgRP (agouti-related peptide) stimulates appetite, whereas stimulation of POMC (pro-opiomelanocortin) cells that produce α-MSH (α-melanocyte-stimulating hormone) suppresses

appetite. Hormones that affect ARC-regulated feeding behavior include **leptin** (a satiety-including protein secreted by adipose tissue in proportion to adipose tissue mass), insulin, ghr, and PYY. Under normal conditions, when leptin levels rise, indicating that the body's energy resources are sufficient, NPY/AgRP neurons are inhibited and POMC neurons are activated. In this circumstance, appetite for food is depressed. In contrast, increased food consumption, triggered by the activation of NPY/AgRP neurons and the inhibition of POMC neurons, results from falling leptin levels (caused by weight loss). All these neurons send signals to other neurons in the hypothalamus (referred to collectively as second-order neurons), which in turn relay them to other parts of the brain. Among these targets is the *nucleus tractus solitarius* (NTS) within the brain stem, which integrates this information with appetite-regulating signals to and from the GI tract via the vagus nerve.

The results of feeding behavior–related signaling, depending on circumstances, range from increased appetite to a sense of satiety. Insulin reduces food intake via the NPY/AgRP and POMC neurons, though to a lesser extent than leptin. Insulin regulates leptin synthesis. During fasting or calorie restriction (dieting), decreased leptin levels, caused by decreased adipose tissue mass, contribute to increased hunger and subsequent weight gain. Ghr, an orexigenic (appetite-stimulating) molecule released by cells within the stomach and the small intestine, activates NPY/AgRP neurons. PYY, an anorexigenic (appetite-inhibiting) molecule produced by cells in the small intestine and colon, inhibits NPY/AgRP neurons. CCK, which acts to stimulate the release of digestive aids from the pancreas and gall bladder, also acts as an appetite suppressant by delaying gastric emptying and as an indirect inhibitor of NPY/AgRP neurons. CCK inhibits appetite by binding to receptors within the NTS and through neurons linking the NTS to areas within the hypothalamus that modulate appetite suppression. In addition, leptin signaling enhances CCK's satiating effect. The appetite regulating ARC neurons are also sensitive to local levels of glucose, fatty acids, and the amino acid leucine.

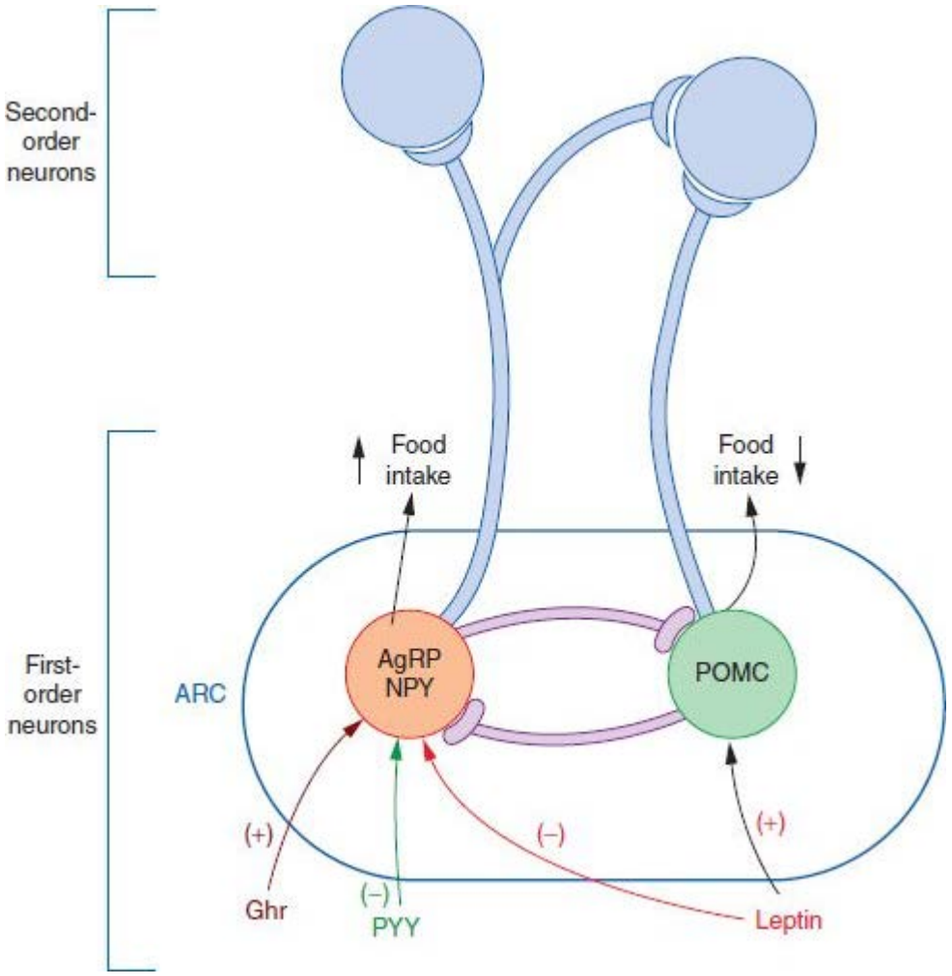


FIGURE 16.16

Appetite-Regulating Neurons in the Arcuate Nucleus

Within the arcuate nucleus (ARC) of the hypothalamus, there are two sets of appetite-regulating neurons with opposing effects: AgRP/NPY and POMC. Both neuron types are first-order neurons (i.e., they respond to peripheral signal molecules). Activation of AgRP/NPY neurons by Ghr increases appetite and energy-requiring metabolic processes. When POMC neurons are activated by leptin (and insulin to a lesser extent), appetite is depressed. Appetite is also negatively affected by the actions of leptin and PYY on AgRP/NPY neurons. When activated, AgRP/NPY suppresses POMC neurons. Appetite-regulating signals generated by AgRP/NPY and POMC neurons are transmitted via second-order neurons to other parts of the hypothalamus and then other brain centers. For example, AgRP and α-MSH compete for the same binding site on MC4R, a receptor on neurons in PVH (paraventricular hypothalamic nucleus). Signals from these centers are subsequently sent to the NTS in the brain stem, where they are integrated with neural signals from the GI tract and other organs.

The integration of disparate appetite-regulating hormonal and nutrient signals received within the hypothalamus appears to be mediated by AMPK (p. 470) and mTORC1 (p. 612). The signal transduction events that cause neurons to fire and release NPY and AgRP neurotransmitter molecules are triggered by activation of AMPK (p. 619) by appetite-promoting molecules such as ghr. The binding of these molecules to cell-surface receptors in neurons in ARC and other hypothalamic regions, in combination with low blood glucose levels, promotes appetite stimulation. When appetite-inhibiting hormones such as leptin and insulin bind to their cell-surface receptors, AMPK activity is inhibited, with the result that ArGP/NPY neurons are inhibited. ARC nutrient-sensing neurons also utilize mTORC1 to regulate feeding behavior. mTORC1 activity, which varies inversely with that of AMPK, is stimulated by leptin, insulin, and nutrients. As a result of the mTORC1-triggered signal transduction pathway, appetite is depressed. Low-energy

availability (i.e., a high AMP/ATP ratio) results in AMPK activation. Activated AMPK inhibits mTORC1, with the result that appetite increases.

Biochemistry IN PERSPECTIVE

Obesity and the Metabolic Syndrome

Why are so many humans predisposed to obesity in the modern world, especially in the past 30 years? Obesity, excess body weight that results from an imbalance between energy intake and energy expenditure, is a worldwide epidemic. If the human brain is so adept at balancing appetite, satiety, and physical activity, why has obesity become such a threat to health in the past 35 years? The answer lies in the long human struggle for survival and a significant recent change to the modern human diet.

For most of its existence, *Homo sapiens* has had physically challenging hunter–gatherer and agrarian lifestyles in which diets consisted of whole foods such as vegetables, fruit, nuts, seeds, eggs, fish, and lean meat. In the modern world, in contrast, an increasing proportion of human populations live under vastly different circumstances. The availability of inexpensive, calorie-dense, and nutrient-poor processed food has been combined with sedentary lifestyles. As a result, there has been significant weight gain among individuals genetically predisposed to what have now become problems with body weight regulation.

For many people, a calorie-restricted environment combined with vigorous physical work would mask such vulnerability. Several rare mutations, linked to monogenic disorders of body weight regulation (**Figure 16A**), provide insight into human feeding behavior. For example, obese patients with leptin or leptin receptor defects have insatiable appetites. In Prader–Willi syndrome (PWS), a rare obesity-related disorder, a deletion of a segment of paternal chromosome 15 results in disruption of hypothalamic satiety control, leading to exceptionally insatiable hunger. By one estimate, at least 40% of the factors causing obesity can be attributed to genetics. Obesity may result from a combination of a calorie-dense environment, sedentary lifestyles, and mutations in one or more genes coding for appetite and satiety signal transduction pathways components. Recent research has linked the obesity epidemic over the past several decades to the introduction in the 1970s of large amounts of fructose into processed foods. Fructose promotes weight gain because, unlike glucose, it does not suppress the release of ghr, the appetite-stimulating hormone, or stimulate the release of leptin and insulin, both of which promote satiety.

FIGURE 16A

Obesity and Genetics

Both mice have identical genomes except that the *leptin* gene was deleted from the mouse on the left.

In addition to the discomfort and severe social stigma associated with obesity, excessive body weight is a serious risk factor for a variety of illnesses such as hypertension, heart disease, osteoarthritis, diabetes mellitus, Alzheimer's disease, and several forms of cancer. Obesity is now recognized as a contributing factor in metabolic syndrome.

Metabolic Syndrome

Metabolic syndrome is the term used to describe a cluster of clinical disorders that include, in addition to obesity, hypertension, dyslipidemia (high blood levels of total cholesterol and triacylglycerol and low HDL levels), and insulin resistance. Insulin resistance, one of the earliest manifestations of the metabolic syndrome, originates in part from excess levels of free fatty acids (FFAs) in blood. Expanding adipocytes, especially those of visceral (abdominal) adipose tissue, release fatty acids into the bloodstream. As the FFA levels in blood rise, these molecules begin accumulating in other cells. In insulin-sensitive tissue (muscle, liver, and pancreatic β cells), FFAs disrupt signal transduction pathways. Among the consequences of this process, called *lipotoxicity*, are dyslipidemia and hyperinsulinemia caused by FFA-stimulated insulin secretion and insulin resistance. Other effects include excess glucose production in liver (uninhibited gluconeogenesis and glycogenolysis) and inhibited insulin-mediated glucose uptake by muscle. When adipocytes also become insulin-resistant, there is an increase in lipolysis that results in the release of more FFA into the bloodstream. In genetically predisposed individuals, the metabolic syndrome may develop into type 2 diabetes (see the earlier Biochemistry in Perspective essay in this chapter, pp. 615–17).

In addition to promoting obesity and hypertension (p. 600), excessive fructose consumption contributes to the metabolic syndrome by other means. First, fructose has a greater propensity than glucose toward glycation reactions (p. 256) and AGE formation (linked to atherosclerosis and other inflammation-based diseases) because it spends proportionally more time in its open-chain form. Second, hepatic fructose metabolism is highly lipogenic. Consumption of large amounts of fructose has been linked not only to dyslipidemia, but also to nonalcoholic fatty liver disease, a disorder that compromises liver function. (Fructose has such a negative effect on the liver because most dietary fructose molecules are metabolized by this organ owing to the presence of GLUT5 in the hepatocyte plasma membrane.) Third, fructose consumption preferentially results in increased accumulation of abdominal fat, a major cause of inflammation. Finally, fructose-induced lipogenesis and uric acid production contribute to insulin resistance.

SUMMARY Natural selection in response to the rigors of chronic food scarcity has left many humans with the propensity to gain weight when calorie-dense food is plentiful. The body's inability to cope with lipotoxicity, caused by excessive body weight and fructose-caused metabolic stress, can result in metabolic syndrome.

WORKED PROBLEM 16.1

Body mass index (BMI) is a measure of a person's body composition that is based on both weight and height. It is defined as weight in kg/(height in m)^2. A normal healthy person's BMI falls within the range of 18.5 to 24.9. Individuals who have BMI values between 25 and 29.9 or above 30 are designated as overweight and obese, respectively. Calculate the BMI values for the following individuals and determine how they would be classified: three 6-foot (1.829 m) men with weights of 150 pounds (68 kg), 200 pounds (91 kg), and 250 pounds (115 kg).

SOLUTION

Substitute the given values into the equation BMI = weight/(height in m)^2. For the first man, his BMI is calculated as

$$\text{BMI} = 68 \text{ kg}/(1.829 \text{ m})^2 = 68 \text{ kg}/3.349 \text{ m}^2 = 20.3 \text{ (normal)}$$

For the second man, his BMI is calculated as

$$\text{BMI} = 91 \text{ kg}/3.349 \text{ m}^2 = 27 \text{ (overweight)}$$

For the third man, his BMI is calculated as

$$\text{BMI} = 115 \text{ kg}/3.349 \text{ m}^2 = 34 \text{ (obese)}$$ ■

Chapter Summary

1. Multicellular organisms require sophisticated regulatory mechanisms to ensure that all their cells, tissues, and organs cooperate.
2. Hormones are molecules that multicellular organisms use to convey information between cells. When target cells are distant from the hormone-producing cell, such molecules, released by glandular cells, are called endocrine hormones. To ensure proper control of metabolism, the synthesis and secretion of many mammalian hormones are ultimately controlled by the central nervous system. In addition, a negative feedback mechanism precisely controls various hormone syntheses. A variety of diseases are caused by either overproduction or underproduction of a specific hormone or by the insensitivity of target cells.
3. Signal molecules initiate their effects in the target cell by binding to a specific receptor. Polar, water-soluble signal molecules, such as amines and peptides, bind to cell-surface receptors. They alter the activities or synthesis of several enzymes and/or transport mechanisms in the cell. G protein-coupled receptors use one or more of the second messengers (cAMP, cGMP, IP_3, DAG, and Ca^{2+}) to mediate the primary signal's effect on the target cell in a process that has a significant amplifying effect on signal transduction. Enzyme-linked receptors such as tyrosine kinase receptors initiate signal transduction via intrinsic enzymatic activity or recruitment of intracellular enzymes. Insulin/RTK signaling utilizes two major mechanisms: the PI3K/PKB pathway and the Ras/MAP kinase pathway. Tyrosine kinase receptors do not directly involve the generation of a second messenger. The nonpolar steroid and thyroid hormones diffuse through the lipid bilayer and bind to intracellular receptors. The hormone-receptor complex subsequently binds to a DNA sequence referred to as a hormone response element (HRE). The binding of a hormone-receptor complex to an HRE enhances or diminishes the expression of specific genes.
4. The feeding–fasting cycle illustrates how a variety of organs contribute via hormones and neurotransmitters to the acquisition of food molecules and their use. Feeding behavior is a mechanism by which animals seek out and consume food. The goal is to maintain balance between energy acquisition and energy expenditure. The hypothalamus contains the critical neural circuits that control appetite and

satiety.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on integration of metabolism to help you prepare for exams.

Chapter 16 Review Quiz

Suggested Readings

Bedse G. 2015. Aberrant insulin signaling in Alzheimer's disease: current knowledge. Frontiers Neuroscience 9:204.

Bidwell AJ. 2017. Chronic fructose ingestion as a major health concern: is a sedentary lifestyle making it worse? Nutrients doi:10.3390/nu906059.

Haeusler RA et al. 2018. Biochemical and cellular properties of insulin receptor signaling. Nat Rev Mol Cell Biol 19:31–44.

Hilgendorf KI et al. 2016. The primary cilium as a cellular receiver: organizing ciliary GPCR signaling. Curr Opin Cell Biol 39:84–92.

Kuhn M. 2016. Molecular physiology of membrane guanylyl cyclase receptors. Physiol Rev 96:751–804.

Kullmann S, et al. 2016. Brain insulin resistance at the crossroads of metabolic and cognitive disorders in humans. Physiol Rev 96:1169–209.

Lee S-H et al. 2016. Insulin in the nervous system and the mind: functions in metabolism, memory, and mood. Mol Medicine 5:589–601.

Pal M et al. 2016. The roles of c-Jun NH2-terminal kinases (JNKs) in obesity and insulin resistance. J Physiol 594(2):267–79.

Peterson YK, Luttrell LM. 2017. The diverse roles of arrestin scaffolds in G protein–coupled receptor signaling. Pharmacol Rev 69:256–97.

Samuel VT, Shulman GI. 2016. The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux. J Clin Invest 126(1):12–22.

Solinas G, Becattini B. 2017. JNK at the crossroads of obesity, insulin resistance, and cell stress response. Mol Metab 6:174–84.

Key Words

autocrine, *603*

cytokine, *603*

DAG, *609*

desensitization, *605*

docking protein, *611*

downregulation, *605*

dyslipidemia, *615*

endocrine system, *601*
G protein, *605*
G protein-coupled receptor, *605*
G protein-coupled receptor kinase, *610*
ghrelin, *619*
glucosuria, *615*
growth factor, *603*
GTPase-activating protein, *607*
guanine nucleotide exchange factor, *606*
hormone response element, *614*
hyperglycemia, *615*
hyperinsulinemia, *616*
hyperosmolar hyperglycemic nonketosis, *616*
incretin, *621*
insulin-like growth factor, *603*
insulin resistance, *605*
IP_3, *609*
ketoacidosis, *615*
ketosis, *615*
leptin, *624*
metabolic syndrome, *626*
osmotic diuresis, *615*
paracrine, *603*
postabsorptive, *619*
postprandial, *619*
receptor tyrosine kinase, *611*
second messenger, *604*
steady state, *601*
target cell, *601*
tropic hormone, *605*
type 1 diabetes, *615*
type 2 diabetes, *615*
type 3 diabetes, *617*

Review Questions

SECTION 16.1

Comprehension Questions

1. Define the following terms:
    a. steady state

b. target cells
c. endocrine
d. paracrine
e. autocrine

2. Define the following terms:

a. TSH
b. T_3
c. T_4
d. endocrine system
e. SUA

Fill in the Blanks

3. High serum uric acid levels are linked to _____,_____, _____, _____, and _____.
4. In _____, the rate of anabolic processes in the body is approximately equal to that of catabolic processes.
5. Each hormone molecule in living organisms is recognized by specific cells called _____.
6. _____ hormones are secreted directly into the bloodstream.
7. _____ signal molecules bind to receptors on the surface of neighboring cells.

Short-Answer Questions

8. NADH is an important reducing agent in cellular catabolism, whereas NADPH is an important reducing agent in anabolism. Review previous chapters and identify how the synthesis and degradation of these two molecules are interconnected.
9. Explain, in general terms, the roles of the nervous and endocrine systems in metabolism.
10. Explain the relationship between blood uric acid levels and hypertension.

Critical-Thinking Questions

11. Explain how consumption of fructose-containing soft drinks contributes to gout symptoms.
12. Why did researchers investigate the relationship of uric acid to hypertension in young patients rather than adults?

SECTION 16.2

Comprehension Questions

13. Define the following terms:

a. growth factors
b. cytokine
c. second messenger
d. enzyme cascade
e. desensitization

14. Define the following terms:

a. downregulation

b. tropic hormone
c. G protein-coupled receptor
d. G protein
e. cAMP

15. Define the following terms:

a. adenylate cyclase
b. $G\alpha_s$
c. $G\beta\gamma$
d. PKA
e. CRE

16. Define the following terms:

a. atrial natriuretic factor
b. heat stable enterotoxin
c. PKG
d. PKB
e. mTORC1

17. Define the following terms:

a. G protein-coupled receptor kinase
b. β-arrestin
c. enzyme-linked receptor
d. guanylin
e. RTK

18. Define the following terms:

a. phosphatidylinositol cycle
b. PIP_2
c. IP_3
d. DAG
e. PKC

19. Define the following terms:

a. SREBP-1c
b. PPARγ
c. ChREBP
d. Grb2
e. SOS

20. Define the following terms:

a. type 1 diabetes
b. type 2 diabetes
c. type 3 diabetes
d. dyslipidemia
e. glucosuria

21. Define the following terms:

a. glucosuria

b. osmotic diuresis
c. ketoacidosis
d. ketosis
e. hyperinsulinemia

22. Define the following terms:

a. hyperosmolar hyperglycemic nonketosis
b. insulin resistance
c. JNK
d. PKCθ
e. PKCε

23. Define the following terms:

a. IDE
b. MAPK
c. blood-brain barrier
d. TNF-α
e. HRE

Fill in the Blanks

24. The reduction in the numbers of cell-surface receptors in response to stimulation by their specific hormone molecule is called _____.
25. _____ is a peptide hormone released by heart atrial cells in response to increased blood volume.
26. Type 1 diabetes is caused by the destruction of _____.
27. Type 2 diabetes is caused by an insensitivity to _____.
28. Glucose in the urine is referred to as _____.
29. Excessive loss of water and electrolytes in diabetes is referred to as _____.
30. The most serious symptom of type 1 diabetes is _____.
31. Elevated concentration of blood ketones is referred to as _____.
32. The two major types of cell-surface receptors that bind to hormone molecules are G protein-linked receptors and _____ receptors.
33. A group of molecules called _____ mediate the actions of many hormones.
34. Phosphorylation of CFTR by _____ results in Cl^- and HCO_3^- secretion into the intestinal lumen.

Short-Answer Questions

35. The binding of insulin to its receptor on the surface of a target cell activates the receptor's tyrosine kinase activity, which in turn causes several phosphorylation cascades that alter the activity of numerous enzymes as well as the expression of specific genes. Describe how one of these phosphorylation cascades stimulates glycogen synthesis.
36. Bodybuilders often consume anabolic steroids to increase their skeletal muscle mass. How do these steroid molecules achieve this effect? (Note that the common side effects of anabolic steroid abuse include heart failure, violent behavior, and liver cancer.)
37. Extreme thirst is a characteristic of diabetes. Explain.
38. Briefly discuss the major classes of second messenger that are currently recognized.
39. How do phorbol esters promote tumor growth?
40. Hemoglobin molecules exposed to high levels of glucose are converted to glycated products.

The glycated product used to measure blood glucose control in diabetics is hemoglobin A_{1C} (HbA_{1C}), contained in red blood cells that last around 3 months. In general terms, explain how HbA_{1C} forms and why it is a useful marker for blood glucose control.

41. Ketoacidosis is a common feature of insulin-dependent diabetes mellitus, but not insulin-independent diabetes mellitus. Explain.
42. What are the most common sites on proteins that are phosphorylated during signal transduction cascades?
43. Type 2 diabetics are often obese. Explain how obesity contributes to the onset of diabetes.
44. The principal early physiological symptom of diabetes is high blood levels of glucose. As a result, diabetes is often associated with abnormal carbohydrate metabolism. List several other insulin-dependent processes that are affected.

Critical-Thinking Questions

45. In severely diabetic patients, the blood glucose level is so high that it appears in the urine. Before the development of blood tests by modern medical research, diabetics could often be recognized by the appearance of flies around their feet. Suggest a reason for this observation.
46. Explain why obese individuals are often insulin-resistant.
47. Explain how second messenger molecules work. Why use a second messenger rather than simply relying on a more direct hormone-stimulated transduction mechanism?
48. Steroid hormones are often present in cells in low concentrations that make them difficult to isolate and identify. Researchers find it easier to isolate the proteins that bind steroids using affinity chromatography. [*Hint*: Refer to the Biochemistry in the Lab essay on protein technology on pp. *184–89*]. Explain how you would use this technique to isolate a protein suspected of steroid hormone binding.
49. Describe the relationship between AMPK and mTORC1.
50. Several hormones may activate the same G protein. Therefore, different hormones may have the same or a similar effect. For example, glycogen degradation is initiated by both epinephrine and glucagon. Why is overlap of function an advantage?
51. Hyperinsulinism, the result of an excessive insulin dose or an insulin-secreting tumor, can result in brain damage. Explain.
52. Describe the mechanisms in adipose tissue, skeletal muscle, and liver that result in insulin resistance in these tissues.

SECTION 16.3

Comprehension Questions

53. Define the following terms:
    a. ghrelin
    b. PYY
    c. CCK
    d. GLP-1
    e. incretin
54. Define the following terms:
    a. postprandial
    b. postabsorption

c. chylomicron remnants
d. feeding behavior
e. satiety

55. Define the following terms:

a. NPY/AgRP
b. POMC
c. NTS
d. leptin
e. α-MSH

Fill in the Blanks

56. _____ is a cluster of disorders that include obesity, hypertension, dyslipidemia, and insulin resistance.
57. The _____ within the brain stem integrates appetite-regulating signals.
58. The term _____ is used to refer to molecules that stimulate appetite.

Short-Answer Questions

59. During periods of fasting, some muscle protein is depleted. How is this process initiated, and what happens to the amino acids in these proteins?
60. Explain why a diet high in fructose is a major contributing factor in atherosclerosis.
61. What do leptin deficiency and Prader–Willi syndrome have in common?
62. Peptide YY (PYY) is a 36-amino acid peptide that is so named because it contains two biochemically relevant tyrosine residues. (The one-letter symbol for tyrosine is Y; refer to Table 5.1) Describe the function of PYY and state where it is synthesized.
63. Describe the effects of a diet that includes large amounts of fructose.
64. After about 6 weeks of fasting, the production of urea is decreased. Explain.

Critical-Thinking Questions

65. When humans fast, virtually all their glucose reserves are consumed on the first day. The brain requires glucose to function and adjusts slowly to other energy sources such as ketone bodies. Explain how the body supplies glucose to the brain during this process.
66. Long-term weight loss, as the result of dieting (calorie restriction), has a failure rate of about 95%. Review the basic principles of systems biology and explain how the body resists conscious efforts to lose weight.
67. In controlled clinical trials, the consumption of soft drinks containing artificial sweeteners was observed to result in an increased appetite for carbohydrate-containing foods. Based on your knowledge of appetite control by the brain, speculate as to why this phenomenon occurs.
68. As a result of social and economic changes after World War II, the Pima Indians of Arizona began to adopt a “Westernized” lifestyle that included high-calorie diets and sedentary occupations. Soon afterward, obesity and type 2 diabetes became common. A genetically related group, Pima Indians in Mexico, who lived in remote mountain villages, had a low rate of obesity and diabetes. Researchers have discovered that IRS1 levels are reduced in obese individuals from the Arizona Pima population, as compared with lean individuals. It was also shown that specific mutations in the IRS1 gene in combination with obesity led to a 50% reduction in insulin sensitivity. Explain this phenomenon.
69. Pima Indians are encouraged to exercise regularly to delay the onset of diabetes and/or improve diabetic symptoms. What impact does vigorous physical exercise have on their health?

70. The fat store of a normal 150-pound (about 68 kg) man is about 1.5 kg or 141,000 cal. Assuming that during a prolonged fast such a person “burns” 2000 cal/day, determine how long this reserve will last.

MCAT Study Question

71. Which of the following molecules is not involved in the early postprandial state?

    a. insulin
    b. incretin
    c. glucagon
    d. lactate

72. Which of the following molecules inhibits appetite for food?

    a. cholecystokinin
    b. ghrelin
    c. AgRP
    d. AMPK

73. PKA activity initiates a signal transduction mechanism that involves all of the following except

    a. CREB
    b. glycogen phosphorylase
    c. nitric oxide
    d. CRE

74. $G\alpha_q$ activates which of the following enzymes?

    a. phospholipase C
    b. PDK1
    c. adenylate cyclase
    d. PKG

75. All of the following molecules are involved in insulin signal transduction except

    a. IRS
    b. RAS
    c. ANF
    d. PIP_3

CHAPTER 17 Nucleic Acids

Genetic Inheritance The traits of human parents, such as eye and hair color, are inherited by their children. It is the nucleotide base sequence of DNA that transmits genetic information from one generation to the next.

OUTLINE

WHAT MAKES US HUMAN?

17.1 DNA
DNA Structure: The Nature of Mutation
DNA Structure: The Genetic Material
DNA Structure: Variations on a Theme
DNA Supercoiling
Chromosomes
Genome Structure

17.2 RNA
Transfer RNA
Ribosomal RNA
Messenger RNA
Noncoding RNA

17.3 VIRUSES
Bacteriophage T4: A Viral Lifestyle

Biochemistry in Perspective
Epigenetics and the Epigenome: Genetic Inheritance beyond DNA Base Sequences

Biochemistry in the Lab
Nucleic Acid Methods

Biochemistry in Perspective
Forensic Investigations

Biochemistry in Perspective
Ebola Virus

AVAILABLE ONLINE
Biochemistry in Perspective
A Short History of DNA Research: The Early Years
Biochemistry in Perspective
HIV Infection

What Makes Us Human?

Chimpanzees (*Pan troglodytes*) are our closest living relatives. In addition to obvious similarities between humans and chimpanzees in anatomy, physiology, and social behavior, we share an astonishing 99.5% of our genome with them (after taking inserted and deleted DNA sequences into account). And yet humans differ in significant ways from chimpanzees. In the 6 million years since the last common ancestor of chimpanzees and humans lived, we have evolved such uniquely human traits as bipedalism (upright walking), opposable thumbs, and the cognitive power and flexibility that make tool

making, abstract thought, complex language, and art possible.

In 2005, biostatisticians began the process of determining exactly how we differ from chimpanzees. Equipped with complete chimpanzee and human genomes and powerful clustered computers, they compared the DNA sequences of both species and ascertained which sequences are exclusively human (altogether about 15 million bases). Although there are some sequence changes in protein-coding genes, most are found in noncoding DNA, formerly referred to as "junk DNA." Acting as genetic switches, these fast-evolving human regulatory sequences are involved in the activation or repression of nearby genes. Instead of having a lot of new protein-coding genes, humans regulate their genes somewhat differently than other primates. Not surprisingly, many human genetic switch sequences are active in the brain. Specific human changes have also occurred as copy number variations (CNVs), or changes in the number of copies of a specific DNA sequence. CNVs are often the result of deletions or duplications. Among the most distinctive human sequences are the following:

HARs

HARs (human accelerated regions) are DNA sequences that are highly conserved in vertebrates but have changed significantly during human evolution. Of the 202 HARs that have been identified, HAR1 and HAR2 are the most notable. HAR1 codes for a 106 RNA-base sequence (a noncoding RNA; see pp. 13 and 673–75), which differs from the chimpanzee sequence by 18 bases. It is located on chromosome 20, with two overlapping ncRNA genes, HAR1A and HAR1B, which are also known as HACNS1. HAR1A is active in the fetal and adult brain where it is involved in cerebral cortex development. With 12 human-specific base substitutions in its 119-base sequence, HAR2 codes for a human-specific developmental enhancer sequence that contributes to human wrist and thumb formation and ankle and foot modifications that facilitate bipedalism. (An *enhancer* is a short DNA sequence that can increase the transcription of one or more specific genes.)

ASPM

ASPM (*a*bnormal *s*pindle-like *p*rotein *m*icrocephaly-associated) is a gene located on chromosome 1 that codes for a protein involved in mitotic spindle formation and that contributes to enlarged cerebral cortex size. When human ASPM is mutated, babies are born with very small heads, a disorder called microcephaly.

FOXP2

The *FOXP2* (*fo*rkhead bo*x* protein P2) gene, located on chromosome 7, codes for a highly conserved transcription factor that is most active in brain, lung, heart, and GI tract organs. In the developing human brain, FOXP2 regulates the expression of genes involved in several aspects of language and speech production. These include motor control of the mouth and tongue that makes human vocalization possible and cognitive skills, including those required for language comprehension (i.e., processing of words according to grammatical and syntax rules). When compared with the chimpanzee version, human FOXP2 has two specific amino acid changes, which allow the transcription factor to play a key role in synaptic plasticity. (*Synaptic plasticity*, required in such functions as learning and memory, is the capacity to change the strength of a synaptic connection, often by altering the amount of neurotransmitter molecules that are released.) The brain regions most affected are the *basal ganglia*, which is associated with motor control and learning, and the *inferior frontal cortex*, which contains *Broca's area*, linked to speech production. Mutations in human FOXP2 are exceedingly rare. People with this anomaly have severe speech and language defects. They are not only incapable of intelligible speech; they also have cognitive deficits, with lower verbal and nonverbal IQ values.

AMY1

AMY1 on chromosome 1 codes for *salivary amylase*, the enzyme that initiates starch degradation in the mouth. Although many mammals have multiple copies, humans have an especially large number (an

average of seven copies). Consequently, starch is a large component in the human diet. With only two copies, chimpanzees consume a low-starch diet.

LCT

The *LCT* gene located on chromosome 2 codes for the enzyme lactase that hydrolyzes lactose to yield glucose and galactose. Mammals synthesize lactase so that newborn animals can digest milk sugar until weaning when the enzyme is no longer produced. Some humans whose ancestors lived in select areas in Europe and Africa in which cattle were domesticated possess a trait, called *lactose persistence*, which allows them to produce lactase throughout life. The single-base change mutations in a regulatory sequence within a nearby gene allow continuous lactase synthesis in both human populations.

As researchers continue to identify and characterize other DNA sequence differences between our closest relatives and ourselves, this work will reveal the genetic events that have shaped our species.

Overview

THE NUCLEIC ACIDS DNA AND RNA ARE POLYNUCLEOTIDES THAT ENCODE THE GENETIC INFORMATION USED TO CONSTRUCT AND MAINTAIN LIVING organisms. Double-stranded DNA is, in effect, the blueprint used to direct cell processes. Cells then convert DNA's operating instructions into the nucleotide sequence of single-stranded RNA molecules. RNAs have numerous functions, which include polypeptide synthesis, regulation of gene expression (control of when or if a specific gene product is synthesized), and protection from foreign nucleic acids introduced by viral infections. Investigations of nucleic acid structure and function, now almost 70 years old, have given humans a previously unimagined understanding of biological processes and a powerful tool used in such diverse fields as disease diagnosis and treatment and forensic investigations.

For countless centuries, humans have observed inheritance patterns without understanding the mechanisms that transmit physical traits and developmental processes from parent to offspring. Many human cultures have used such observations to improve their economic status, as in the breeding of domesticated animals or seed crops. It was not until the nineteenth century that the scientific investigation of inheritance, now referred to as **genetics**, began. By the beginning of the twentieth century, scientists generally recognized that physical traits are inherited as discrete units (later called "genes") and that chromosomes within the nucleus are the repositories of genetic information. Eventually, the chemical composition of chromosomes was elucidated, and, after many decades of investigation, deoxyribonucleic acid (DNA) was identified as the genetic information. In the decades that followed the 1953 discovery of DNA structure by James Watson and Francis Crick (**Figure 17.1**), a new science emerged. (Refer to the online Biochemistry in Perspective essay A Short History of DNA Research: The Early Years for a brief overview of the work of other scientists that led to the discovery of DNA structure by Watson and Crick.) **Molecular biology** is devoted to the investigation of gene structure and genetic information processing. Using the technologies developed by molecular biologists and biochemists, life scientists have studied the processes by which living organisms organize and process genetic

information. This work has revealed the following principles.

FIGURE 17.1

The First Complete Structural Model of DNA

When James Watson (left) and Francis Crick discovered the structure of DNA in 1953 using Rosalind Franklin's X-ray diffraction images, they were research students at the Henry Cavendish Laboratory of Cambridge University.

1. DNA directs the functioning of living cells and is transmitted to offspring. DNA is composed of two polydeoxynucleotide strands that form a double helix (**Figure 17.2**). The information in DNA is encoded in the form of the sequence of purine and pyrimidine bases (refer to **Figure 14.22**). A **gene** is a DNA sequence that contains the base sequence information necessary to code for a gene product (a polypeptide or several types of RNA molecules) and regulatory sequences that control synthesis of the gene product. The complete DNA base sequence in an organism is referred to as the **genome**. DNA synthesis, referred to as **replication**, involves the complementary pairing of purine and pyrimidine bases between the old parental strand and the newly synthesized strand. The physiological and genetic function of DNA requires the synthesis of error-free copies. Consequently, most organisms employ several DNA repair mechanisms.
2. The mechanism by which genetic information is decoded and used to direct cellular processes begins with the synthesis of ribonucleic acid (RNA). Its synthesis, referred to as **transcription** (**Figure 17.3a**), involves the complementary pairing of ribonucleotide bases with the bases in a DNA molecule. Each newly synthesized RNA molecule is called a **transcript**. The term **transcriptome** designates the complete set of RNA molecules that are transcribed from a cell's genome.

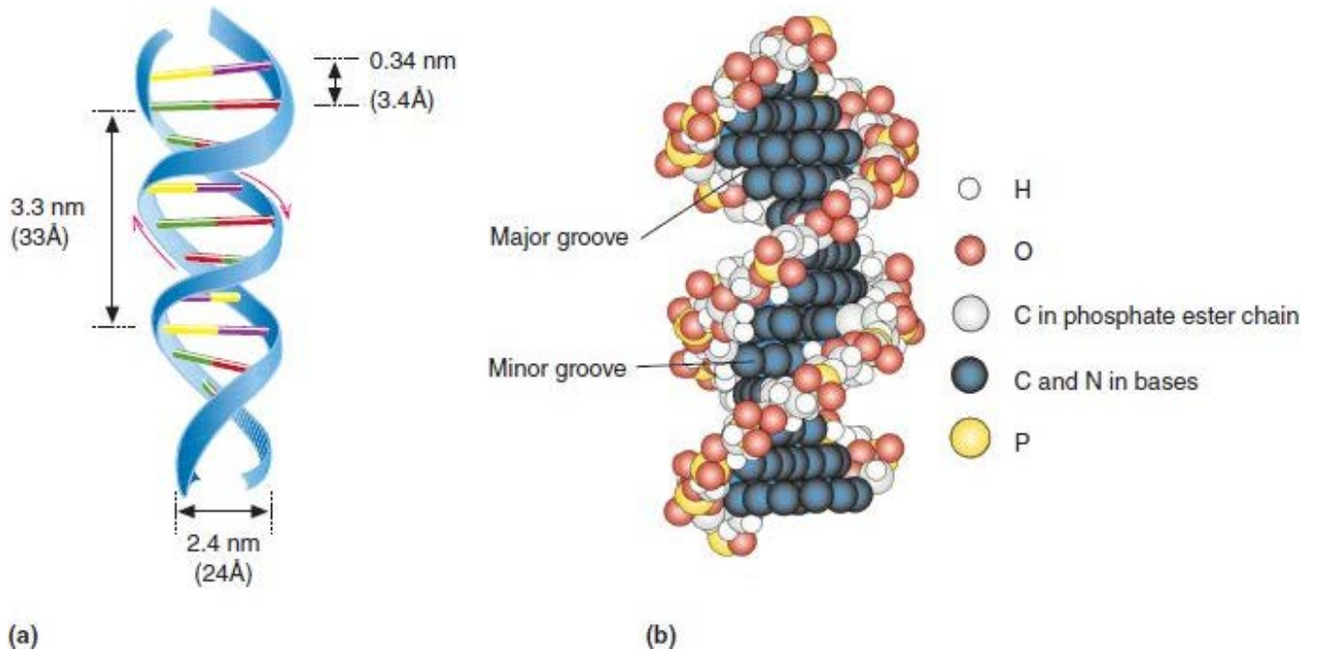


FIGURE 17.2

Two Models of DNA Structure

(a) The DNA double helix is represented as a spiral ladder; this is the conformation originally proposed by Watson and Crick and now known as B-DNA. (For the structural properties of three forms of DNA, see **Table 17.1** later in this chapter.) The sides of the spiral ladder represent the sugar-phosphate backbones. The rungs represent the base pairs. (b) In a space-filling model, colored spheres represent the sugar-phosphate backbones. The base pairs consist of horizontal arrangements of dark blue spheres. Wide and narrow grooves are created by twisting the two strands around each other in a right-handed sense.

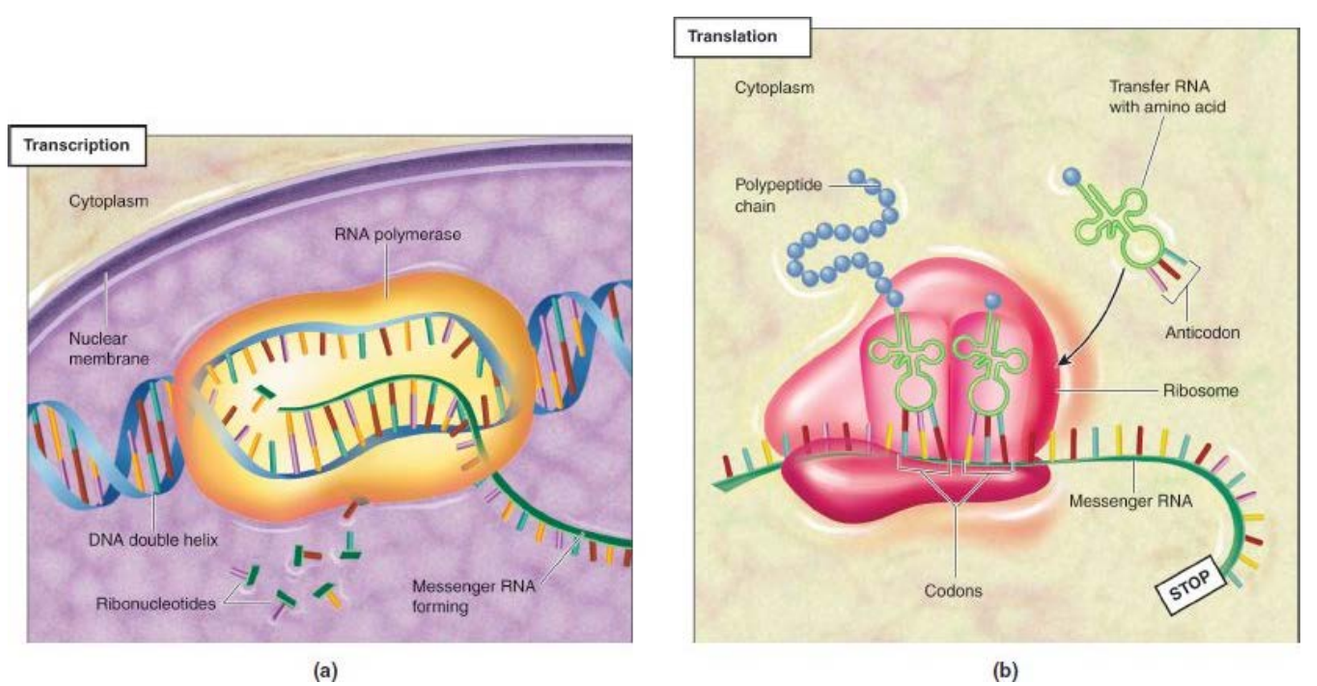


FIGURE 17.3

An Overview of Genetic Information Flow

The genetic information in DNA is converted into the linear sequence of amino acids in polypeptides in a two-phase process. During *transcription* (a), RNA molecules are synthesized from a DNA strand through complementary base pairing between the bases in DNA and the bases in free ribonucleoside triphosphate molecules. During the second phase, called *translation* (b), mRNA molecules bind to ribosomes that are composed of rRNA and ribosomal proteins. Transfer RNA–aminoacyl complexes position their amino acid cargo in the catalytic site within the ribosome in a process that involves complementary base pairing between the mRNA base triplets called codons and tRNA base triplets called anticodons. When the amino acids are correctly positioned within the catalytic site, a peptide bond is formed. After the mRNA molecule moves relative to the ribosome, a new codon enters the ribosome's catalytic site and base pairs with the appropriate anticodon on another aminoacyl-tRNA complex. As a stop codon in the mRNA enters the catalytic site, the newly formed polypeptide is released from the ribosome.

3. Several types of RNA participate directly in the synthesis of the enzymes and other proteins required for the regulated manufacture of all other biomolecules needed in organismal function. The base sequence of each messenger RNA (mRNA) specifies the primary sequence of a specific polypeptide. Ribosomal RNA (rRNA) molecules are components of the ribosomes. Each transfer RNA (tRNA) molecule is covalently bound to a specific amino acid and delivers it to the ribosome for incorporation into a polypeptide chain. Protein synthesis, called **translation** (**Figure 17.3b**), occurs within ribosomes, the ribonucleoprotein molecular machines that translate the base sequences of mRNAs into the amino acid sequences of polypeptides. The entire set of proteins synthesized by a cell is referred to as the **proteome**.
4. **Gene expression** is the set of mechanisms whereby cells control the timing of gene product synthesis in response to environmental or developmental cues. A vast array of proteins, called *transcription factors*, and RNA molecules, called *noncoding RNA molecules* (ncRNAs), regulate gene expression when they bind to specific DNA sequences and/or transcription factors. The term **metabolome** refers to the sum total of all the low-molecular-weight metabolite molecules produced by a cell as the result of its gene expression pattern.

The flow of genetic information can be summarized by a sequence called the *central dogma*:

This diagram illustrates that the genetic information encoded in DNA base sequences flows from DNA, a molecule that is replicated during cell division (indicated by the arrow encircling DNA), to RNA, which specifies the primary structure of proteins. As originally conceived, the central dogma asserted that genetic information flows in one direction only: from DNA to RNA to protein. Several years ago, however, an important exception to the central dogma was revealed. Some of the viruses that have RNA genomes also possess an enzyme activity referred to as *reverse transcriptase*. Once such a virus has infected a host cell, the reverse transcriptase uses a viral RNA base sequence template to form a DNA copy. The viral DNA is then inserted into a host chromosome. One such virus is HIV, discussed in the online Biochemistry in Perspective essay HIV Infection.

Chapter 17 focuses on the structure of the nucleic acids. It begins with a description of DNA structure and how that structure can be altered by mutations. This is followed by a discussion of current knowledge of genome and chromosome structure, as well as the structure and roles of RNA's several forms. A Biochemistry in Perspective essay provides a brief overview of epigenetics, the covalent modification of DNA that adds another layer of gene regulation and

inheritance but is not bound by the Mendelian laws of inheritance. The chapter ends with a description of viruses, macromolecular complexes composed of nucleic acid and proteins that are cellular parasites. Chapter 18 discusses several aspects of nucleic acid synthesis and function (i.e., DNA replication and transcription). Protein synthesis (translation) is described in Chapter 19. The strategies and techniques that are routinely used to isolate, purify, characterize, and manipulate nucleic acids are described in Biochemistry in the Lab boxes in Chapters 17 (Nucleic Acid Methods, pp. 666–70) and 18 (Genomics, pp. 712–18).

17.1 DNA

DNA consists of two polydeoxyribonucleotide strands that wind around each other to form a right-handed double helix (**Figure 17.2**). The structure of DNA is so distinctive that this molecule is often referred to as *the double helix*. As described earlier (Sections 1.3 and 14.3), each nucleotide monomer in DNA is composed of a nitrogenous base (either a purine or a pyrimidine), a deoxyribose sugar, and phosphate. The mononucleotides are linked to each other by 3′,5′-phosphodiester bonds. These bonds join the 5′-hydroxyl group of the deoxyribose of one nucleotide to the 3′-hydroxyl group of the sugar unit of another nucleotide through a phosphate group (**Figure 17.4**). The antiparallel orientation of the two polynucleotide strands allows hydrogen bonds to form between the nitrogenous bases that are oriented toward the helix interior (**Figure 17.5**). There are two types of base pair (bp) in DNA: (1) adenine (a purine) pairs with thymine (a pyrimidine) (AT pair), and (2) the purine guanine pairs with the pyrimidine cytosine (GC pair). The overall structure of DNA resembles a twisted staircase because each base pair is oriented at an angle to the long axis of the helix. The average dimensions of crystalline B-DNA have been measured.

1. One turn of the double helix spans 3.32 nm and consists of approximately 10.3 base pairs. (Changes in pH and salt concentrations affect these values slightly.)
2. The diameter of the double helix is 2.37 nm. Note that the interior space of the double helix is suitable only for base pairing a purine and a pyrimidine. Pairing two pyrimidines would create a gap, and two purines would not fit in the interior space of the double helix. The relative dimensions of both types of base pair are illustrated in **Figure 17.6**.

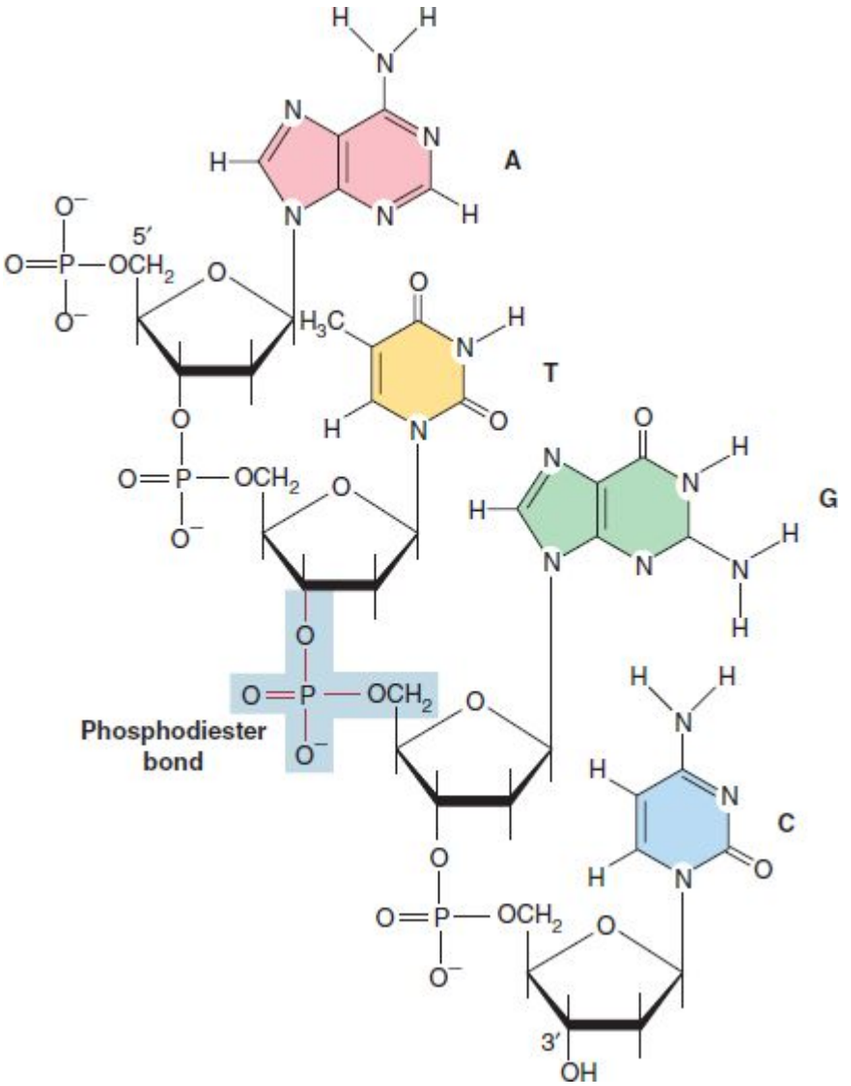


FIGURE 17.4

DNA Strand Structure

In each DNA strand, the deoxyribonucleotide residues are connected to each other by 3′,5′-phosphodiester linkages. The sequence of the strand section illustrated in this figure is 5′-ATGC-3′. Refer to p. 554 for the numbering system of the atoms in nucleotides. (Note that in sugar molecules, the hydrogens bonded to carbon can be represented by single lines.)

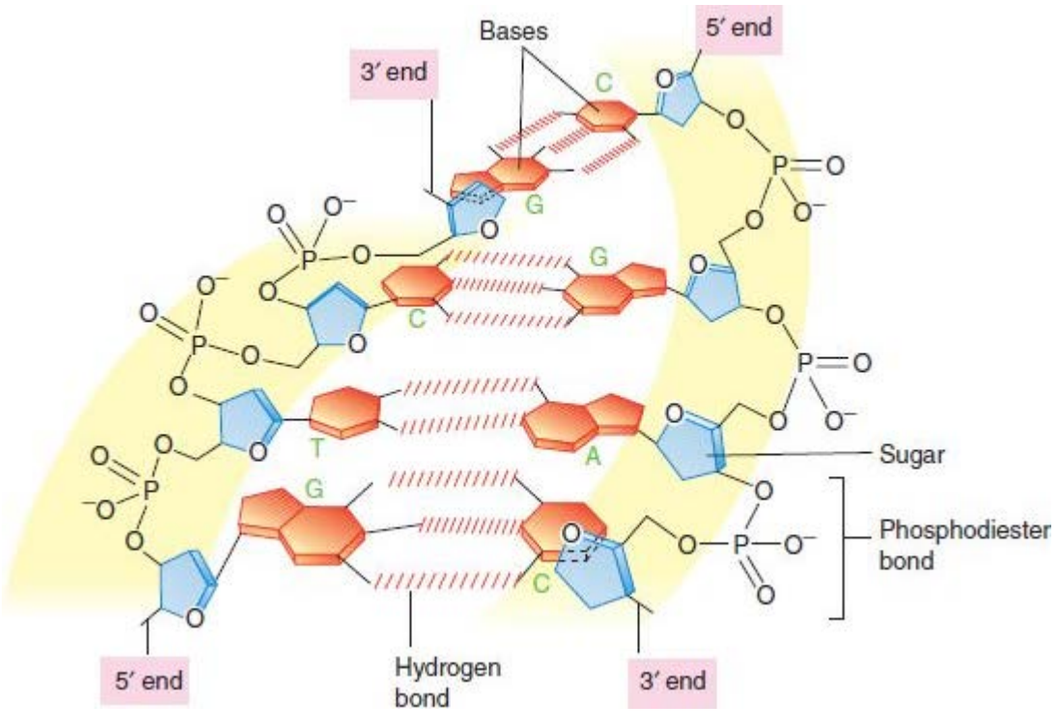


FIGURE 17.5

DNA Structure

In this short segment of DNA, the bases are shown in orange and the sugars are blue. Each base pair is held together by either two or three hydrogen bonds. The two polynucleotide strands are antiparallel. Because of base pairing, the order of bases in one strand determines the order of bases along the other.

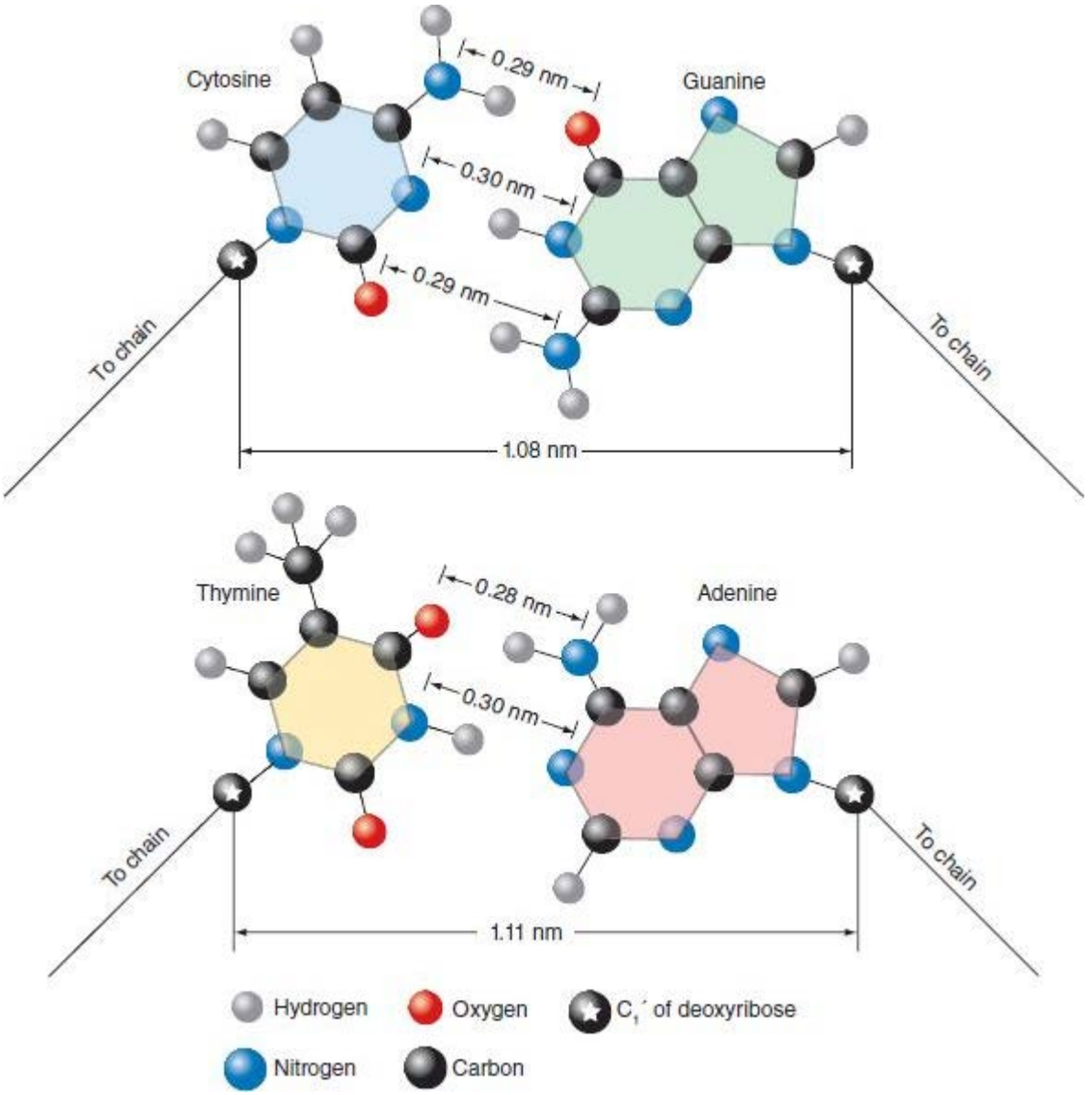


FIGURE 17.6

DNA Structure: AT and GC Base Pair Dimensions

Two hydrogen bonds are formed in each AT base pair and three in each GC base pair. The near-equal dimensions of base pairs of both types allow the formation of uniform helical conformations of the two polynucleotide strands.

As befits its genetic information storage role in living processes, DNA is a relatively stable molecule. Several types of noncovalent interactions contribute to the stability of its helical structure.

1. **Hydrophobic interactions**. The base ring π cloud of electrons between stacked purine and pyrimidine bases is relatively nonpolar. The clustering of the base components of nucleotides within the double helix is a stabilizing factor in the three-dimensional macromolecule because it minimizes their interactions with water, thereby increasing overall entropy.
2. **Hydrogen bonds**. The base pairs, on close approach, form a preferred set of hydrogen bonds, three between GC pairs and two between AT pairs. The cumulative "zippering" effect of these hydrogen bonds keeps the strands in correct complementary orientation.
3. **Base stacking**. Once the antiparallel polynucleotide strands have been brought together by

base pairing, the parallel stacking of the nearly planar heterocyclic bases stabilizes the molecule because of the cumulative effect of weak van der Waals forces generated by π cloud shifts as the bases stack.

4. **Hydration**. As with proteins, water stabilizes the three-dimensional structure of nucleic acids. DNA molecules bind a significant number of water molecules. The water content of B-DNA, the conformation illustrated in **Figure 17.2**, is about 30% by weight. Water molecules bind to phosphate groups, ribose 3′- and 5′-oxygen atoms, and electronegative atoms in the nucleotide bases. When measured under laboratory conditions, each nucleotide in B-DNA binds about 18 to 19 water molecules. Each phosphate can bind a maximum of 6 water molecules.
5. **Electrostatic interactions**. DNA's external surface, referred to as the *sugar–phosphate backbone*, possesses negatively charged phosphate groups. Repulsion between nearby phosphate groups, a potentially destabilizing force, is minimized by the shielding effects of water and divalent cations such as Mg^{2+} and polycationic molecules such as the polyamines (p. 651) and histones (see pp. 651–53).

KEY CONCEPT

DNA is a relatively stable molecule composed of two antiparallel polynucleotide strands wound around each other to form a right-handed double helix.

WORKED PROBLEM 17.1

Deoxynucleotides (deoxynucleoside triphosphates) are the substrates for DNA synthesis. Calculate the number of glucose molecules that must be degraded to generate the energy required for the synthesis of 1000 deoxynucleotides used in DNA synthesis. Assume that glucose is oxidized completely to CO_2 and H_2O, the malate–aspartate shuttle is in operation, and NADPH is the equivalent of 4 ATP.

SOLUTION

The biosynthesis of each deoxynucleotide from a ribonucleoside diphosphate requires two reactions. Ribonucleoside diphosphate molecules containing the bases guanine, thymine, and cytosine are converted to their corresponding nucleoside triphosphates by ATP-requiring nucleoside kinase (p. 557). Each ribonucleoside triphosphate is then converted to a deoxyribonucleoside triphosphate in an NADPH-requiring reaction catalyzed by ribonucleotide reductase (p. 560). Because each NADPH is the equivalent of 4 ATP, the total number of ATPs required to synthesize a deoxyribonucleoside triphosphate is 5 ATP. As a result, the cost of synthesizing 1000 deoxyribonucleotides (1000 × 5) is 5000 ATPs. Because each completely oxidized glucose molecule yields 31 ATPs, the total number of glucose molecules required to synthesize 5000 nucleotides is approximately 161 (i.e., 5000/31). ■

QUESTION 17.1

When DNA is heated, it denatures; that is, the strands separate because hydrogen bonds are

broken and some base-stacking and hydrophobic interactions are disrupted. The higher the temperature, the larger the number of hydrogen bonds that are broken. After reviewing DNA base pair structure, determine which of the following molecules will denature first as the temperature is raised. Explain your reasoning.

a. 5′-GCATTTCGGCGCGTTA-3′
   3′-CGTAAAGCCGCGCAAT-5′
b. 5′-ATTGCGCTTATATGCT-3′
   3′-TAACGCGAATATACGA-5′

DNA Structure: The Nature of Mutation

Although its structure makes DNA eminently suited for information storage, it is not a static molecule. DNA is vulnerable to several types of disruptive forces that can cause mutations, permanent base sequence changes. Mutations range from small-scale alterations (e.g., single-base changes) to large-scale chromosomal abnormalities. Although most mutations are either deleterious or neutral (having no discernible effect on an organism's fitness), on rare occasions a mutation can enhance the adaptation of an organism to its environment. In other words, mutational change is the raw material of evolution. It should also be noted that the mutation rates of most organisms (mutation frequency per cell division or generation) are usually low because of two factors: the accuracy of the DNA replication process (p. 689) and the efficiency of DNA repair (p. 699).

MUTATION TYPES The most commonly observed mutations include single-base changes, insertions, deletions, duplications, and genome rearrangements. As their name implies, a single-base change involves a change in the identity of a single base in one strand of a DNA sequence. Also referred to as **point mutations**, single-base changes can result, if the damage is unrepaired, in transition or transversion mutations. In **transition mutations** (**Figure 17.7**), caused by deamination reactions or tautomerization (p. 554), a pyrimidine is substituted for another pyrimidine or a purine is substituted for another purine. **Transversion mutations**, caused by alkylating agents or ionizing radiation, occur when a pyrimidine is substituted for a purine or vice versa. Point mutations that occur in a population to any extent are referred to as **single nucleotide polymorphisms**. Point mutations can also result from insertions or deletions of an individual base pair. Point mutations within DNA sequences that code for gene products (e.g., polypeptides) are classified as silent, missense, or nonsense, according to their impact on the gene product's structure and/or function. In **silent mutations**, a base change has no discernible effect (e.g., coding for the same or a different amino acid in a polypeptide results in no functional difference), whereas in a **missense mutation**, there is an observable effect (e.g., coding for a different amino acid that causes a change in a polypeptide's structure and function). In a **nonsense mutation**, a point mutation converts the code for an amino acid into a premature stop signal. The polypeptide product of the transcribed sequence containing the nonsense mutation will be incomplete and probably nonfunctional.

Insertions and deletions, referred to as **indels**, are mutations that occur when one to thousands of bases are either inserted or removed from a DNA sequence. If the bases that are inserted or deleted into a polypeptide coding sequence are not divisible by 3 (the length of an individual amino acid coding sequence), a *frameshift mutation* (p. 755) can occur that results in either an altered or a truncated polypeptide. Indels are the result of either unequal crossing over of misaligned homologous chromosomes in meiosis or *slipped strand mispairing*, a mistake made during DNA

replication in which there is displacement and then mispairing of the DNA strands. *Repeat expansion mutations*, a type of insertion mutation also caused by DNA replication mistakes, result in an increased number of small nucleotide sequence repeats (e.g., a trinucleotide). If the repeat expansion becomes large enough, disease can occur. For example, Huntington's disease, a devastating inherited and progressive brain disorder, results when the number of CAG repeats (coding for the amino acid glutamine) in the *huntingtin* gene reaches 36 or more, thereby causing the synthesis of a toxic form of the huntingtin protein.

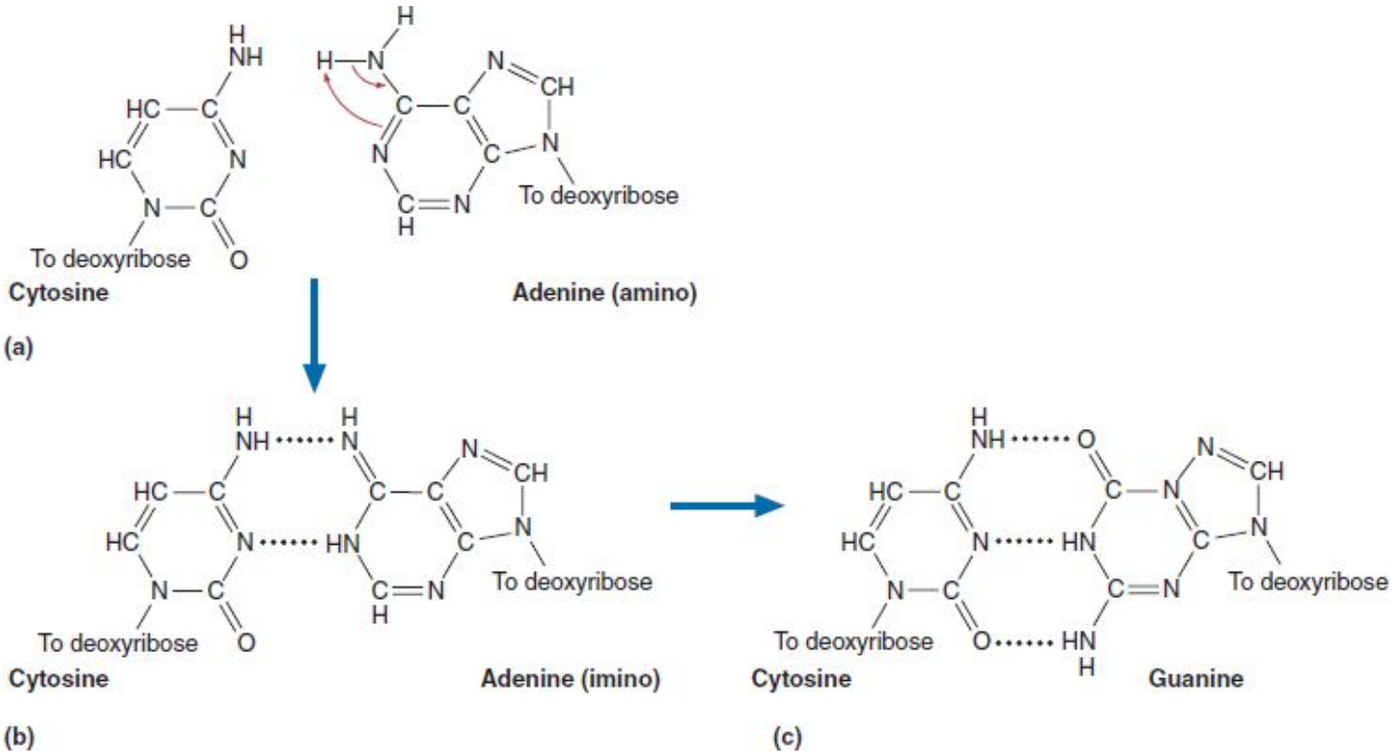


FIGURE 17.7

A Tautomeric Shift Causes a Transition Mutation

As adenine (a) undergoes a tautomeric shift, its imino form (b) can base-pair with cytosine. The transition shows up in the second generation of DNA replication when cytosine base-pairs with guanine. In this manner, an A-T base pair is replaced by a C-G base pair (c). Refer to pp. 554–56 for a description of tautomeric shifts of purine and pyrimidine bases.

3D animation of cytosine

Genome rearrangements such as inversions, translocations, and duplications can cause disruptions in gene structure or regulation. They can result from double-stranded DNA breaks caused by a variety of circumstances, including errors in meiosis and exposure to mutagens or radiation. An **inversion** results when a deleted DNA fragment is reinserted into its original position but in the opposite orientation. **Translocation** is a chromosomal abnormality observed in eukaryotes in which a DNA fragment from one chromosome inserts into a different position on the same chromosome or into a different (nonhomologous) chromosome. **Gene duplication**, the creation of duplicate genes or parts of genes, can result from unequal crossing over in meiosis or from retrotransposition, a process in which genetic elements called retrotransposons (p. 660) insert

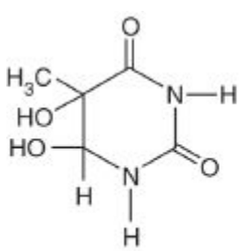
themselves into a DNA sequence. In rare instances, gene duplication is an important feature in evolution because duplicates can, as the result of mutation over long periods of time, take on different functions.

CAUSES OF DNA DAMAGE DNA damage can result from endogenous and exogenous disruptive forces. Endogenous causes of mutations include spontaneous events such as tautomeric shifts, depurination, deamination, and ROS-induced oxidative damage. Exogenous factors such as radiation and xenobiotic exposure (p. 644) can also be mutagenic.

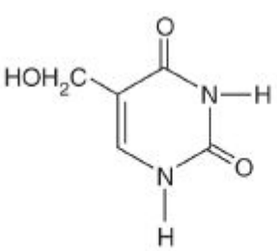
Tautomeric Shifts and Spontaneous Hydrolytic Reactions. Tautomeric shifts (**Figure 14.23**) are spontaneous changes in nucleotide base structure that result in amino to imino and keto to enol changes in configuration. Usually, tautomeric shifts have little effect on overall, three-dimensional DNA structure. However, if tautomers form during DNA replication, base mispairings may result. For example, the imino form of adenine will not form a base pair with thymine. Instead, it forms a base pair with cytosine (**Figure 17.7**). If this pairing is not corrected immediately, a transition mutation occurs because cytosine has been incorporated during the replication process in a position that should carry thymine.

Several spontaneous hydrolytic reactions also cause DNA damage. For example, it has been estimated that several thousand purine bases are lost daily from the DNA in each human cell. In depurination reactions, the N-glycosyl linkage between a purine base and deoxyribose is cleaved. The protonation of N-3 and N-7 of guanine promotes hydrolysis. A point mutation results if repair mechanisms do not replace the purine nucleotide. Similarly, bases can be spontaneously deaminated. For example, the deaminated product of cytosine converts to uracil via a tautomeric shift. Eventually, what should be a CG base pair is converted to an AT base pair. (Uracil is similar in structure to thymine.)

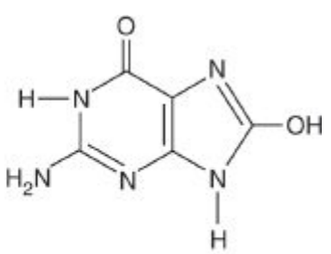
Ionizing Radiation. Some types of ionizing radiation (e.g., UV, X-rays, and γ-rays) can alter DNA structure. Low-radiation levels may cause mutation; high levels can be lethal. UV light-induced damage caused by a free radical mechanism (either abstraction of hydrogen atoms or the creation of •OH and other ROS) includes strand breaks, DNA-protein cross-linking (e.g., via thymine-tyrosine linkages), ring openings, and base modifications. The hydroxyl radical, formed by the radiolysis of water, as well as oxidative stress is known to cause strand breakage and numerous base modifications (e.g., thymine glycol, 5-hydroxymethyluracil, and 8-hydroxyguanine). The enzyme superoxide dismutase (p. 389) plays a key role in preventing ROS-induced damage to both nuclear and mitochondrial DNA.



Thymine glycol



5-Hydroxymethyl uracil



8-Hydroxyguanine

3D animation for 5-hydroxymethyl uracil

3D animation for 8-hydroxyguanine

8-Hydroxydeoxyguanosine (8-OHdG) levels in urine are used to measure the body's production of unquenched ROS. Smoking, for example, results in increases of 8-OHdG excretion by as much as 50%.

The most common UV-induced products, created by UV-B energy absorption by double bonds, are thymine dimers (**Figure 17.8**). The helix distortion that results from dimer formation stalls DNA replication machinery.

Xenobiotics. A large number of xenobiotics can damage DNA. The most important of these molecules are base analogues, alkylating agents, nonalkylating agents, and intercalating agents. The structures of **base analogues** are so similar to normal nucleotide bases that they can be incorporated into DNA. For example, 5-bromouracil (5-BU) is a base analogue of thymine. Because the enol tautomer (p. 556) of 5-BU forms a base pair with guanine, an AT pair can be converted into a GC pair in the next round of replication. **Alkylating agents** are electrophiles that attract molecules that possess an unshared pair of electrons, adding alkyl groups primarily to adenine and guanine, although thymine and cytosine can also be affected. Alkylated bases often pair incorrectly (e.g., methylguanine with thymine instead of with cytosine), leading to possible transition mutations on subsequent rounds of replication. In the case of methylguanine, a GC pair becomes an AT pair. Transversion mutations may also occur when the alkylating group is bulky. The polycyclic aromatic hydrocarbon benzo[*a*]pyrene, found in cigarette smoke, causes transversion mutations because it is converted to a highly reactive epoxide derivative by several biotransformation reactions, including those catalyzed by cytochrome P_{450} (p. 485). (An *epoxide* is cyclic ether with three ring atoms.) The benzo[*a*]pyrene epoxide then forms an adduct of guanine, causing distortion of DNA structure that results in G-to-T transversions as well as disrupted DNA replication. Alkylations can also promote tautomer formation, which may result in transition mutations. Examples of alkylating agents include dimethylsulfate and dimethylnitrosamine.

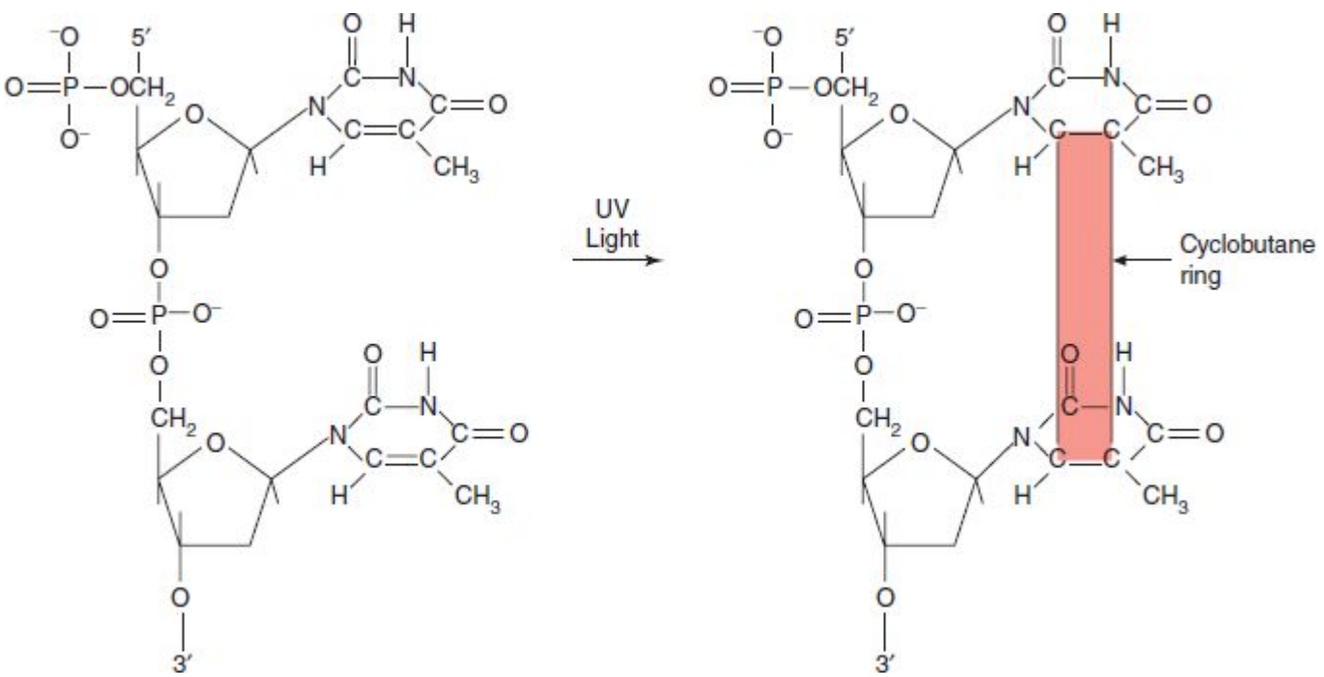


FIGURE 17.8

Thymine Dimer Formation

UV light induces the formation of a covalent linkage in the form of a cyclobutane ring (red) between adjacent thymine bases in a strand of DNA.

A variety of **nonalkylating agents** can modify DNA structure. Nitrous acid (HNO_2), derived from the nitrosamines and from sodium nitrite ($NaNO_2$), deaminates bases. (Both nitrosamines and $NaNO_2$ are found in processed meats and in any foodstuffs preserved with nitrite pickling salt.) HNO_2 oxidatively deaminates adenine, guanine, and cytosine to hypoxanthine, xanthine, and uracil, respectively. Certain planar polycyclic aromatic molecules are referred to as **intercalating agents** because they can distort DNA by inserting themselves (intercalating) between the stacked base pairs of the double helix. The resulting distortion of local DNA structure can disrupt DNA replication. Depending on the level of exposure, intercalating agents can cause damage that ranges from deletions or insertions to cell death. Doxorubicin and ethidium bromide are examples of intercalating agents. The chemotherapeutic agent doxorubicin inhibits DNA replication in fast-growing cancer cells. Ethidium bromide is a fluorescent tag molecule used as a nucleic acid stain in a variety of molecular biology lab techniques.

Mutations caused by DNA replication machinery errors and transposition (the movement of DNA sequences within a genome) are described in Chapter 18.

KEY CONCEPT

DNA is vulnerable to certain types of disruptive force that can result in mutations, permanent changes in its base sequence.

QUESTION 17.2

How will each of the following substances or conditions affect DNA structure?
a. ethanol; b. heat; c. dimethylsulfate; d. nitrous acid; e. 5-BU

QUESTION 17.3

The accumulation of oxidative DNA damage now appears to be a major cause of aging in mammals. Animals that have high metabolic rates (i.e., use large amounts of oxygen) or excrete large amounts of modified bases in the urine typically have shorter life spans. The excretion of relatively large amounts of oxidized bases indicates a reduced capacity to prevent oxidative damage. Despite substantial evidence that oxygen radicals damage DNA, the actual radicals that cause the damage are still not clear. Suggest possible culprits in addition to the hydroxyl radical. Some tissues sustain more oxidative damage than others. For example, the human brain is believed to sustain more oxidative damage than most other tissues during an average life span. Suggest two reasons for this phenomenon.

Radicals and Oxidative Damage

DNA Structure: The Genetic Material

Visit the companion website to read the Biochemistry in Perspective essay A Short History of DNA Research: The Early Years, describing the critical experiments that led to the discovery of DNA as the genetic material.

The publication of James Watson and Francis Crick's model of DNA structure in the April 25, 1953, issue of *Nature* was both an end and a beginning. Their work was the culmination of research over the course of nearly a century. Among the most important discoveries were (1) the deduction in 1928 by the British bacteriologist Frederick Griffith that bacteria can transfer genetic information between cells; (2) the identification of Griffith's "transforming factor" as DNA by Oswald Avery, Colin MacLeod, and Maclyn McCarty in 1944; and (3) the experiment by Alfred Hershey and Martha Chase in 1952 that confirmed that DNA is the genetic material of living organisms. Publication of the Watson–Crick model marked the beginning of a new field called **molecular biology**, which deals with the biosynthesis of DNA, RNA, and proteins and the mechanisms that regulate these processes.

The information used by Watson and Crick to construct their DNA model included the following:

1. The chemical structures and molecular dimensions of deoxyribose, the nitrogenous bases, and phosphate.
2. The 1:1 ratios of adenine to thymine and guanine to cytosine in the DNA isolated from a wide variety of species investigated by Erwin Chargaff between 1948 and 1952. (These 1:1 relationships are sometimes referred to as **Chargaff's rules**.)
3. Superb X-ray diffraction studies performed by Rosalind Franklin (**Figure 17.9**) indicating that DNA is a symmetrical molecule and probably a helix.
4. The diameter and pitch of the helix estimated by Maurice Wilkins and his colleague Alex Stokes from other X-ray diffraction studies.
5. The demonstration by Linus Pauling that protein, another class of complex molecule, could exist in a helical conformation.

KEY CONCEPT

The model of DNA structure proposed by James Watson and Francis Crick in 1953 was based on information derived from the efforts of many individuals.

The 1962 Nobel Prize in Medicine or Physiology was awarded to James Watson, Francis Crick, and Maurice Wilkins.

DNA Structure: Variations on a Theme

The structure discovered by Watson and Crick, referred to as **B-DNA**, represents the sodium salt of DNA under highly humid conditions. DNA can assume different conformations because deoxyribose is flexible and the C^1-N-glycosidic linkage rotates. (Recall that furanose rings have a

puckered conformation.)

A-DNA When DNA becomes partially dehydrated—that is, when the number of water molecules bound to each of the nucleotides drops to about 13 to 14—the molecule assumes the A form (**Figure 17.10** and **Table 17.1**). In **A-DNA**, the base pairs are no longer at right angles to the helical axis. Instead, they tilt 20° away from the horizontal. In addition, the distance between adjacent base pairs is slightly reduced, with 11 bp per helical turn instead of the 10.5 bp found in the B form. Each turn of the double helix occurs in 2.46 nm, instead of 3.32 nm, and the molecule's diameter swells to approximately 2.55 nm from the 2.37 nm observed in B-DNA. The A form of DNA is observed when it is extracted with solvents such as ethanol. The significance of A-DNA under cellular conditions is that the structure of RNA duplexes and RNA/DNA duplexes formed during transcription resembles the A-DNA structure.

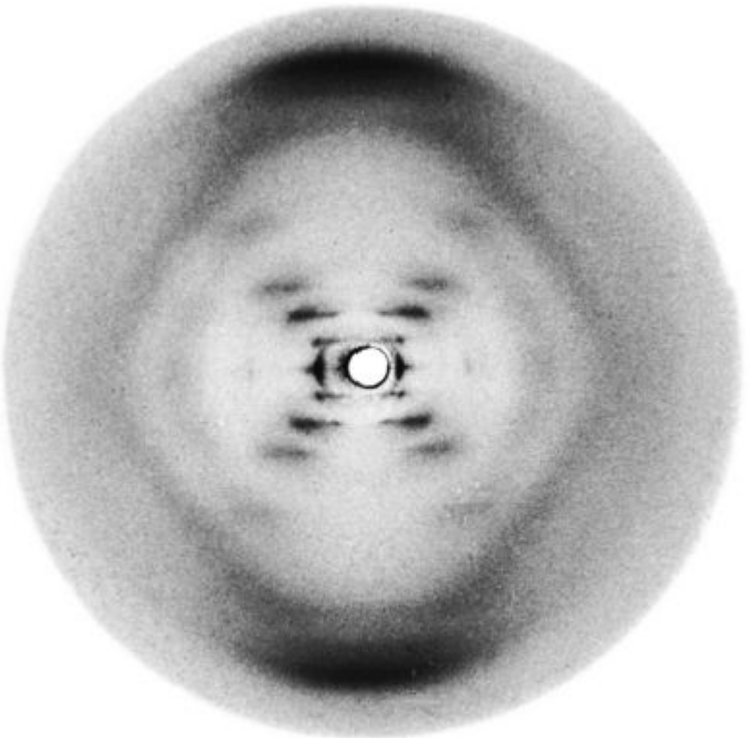


FIGURE 17.9

X-Ray Diffraction Study of DNA by Rosalind Franklin and R. Gosling

The symmetry of the X-ray diffraction pattern indicates a helical structure.

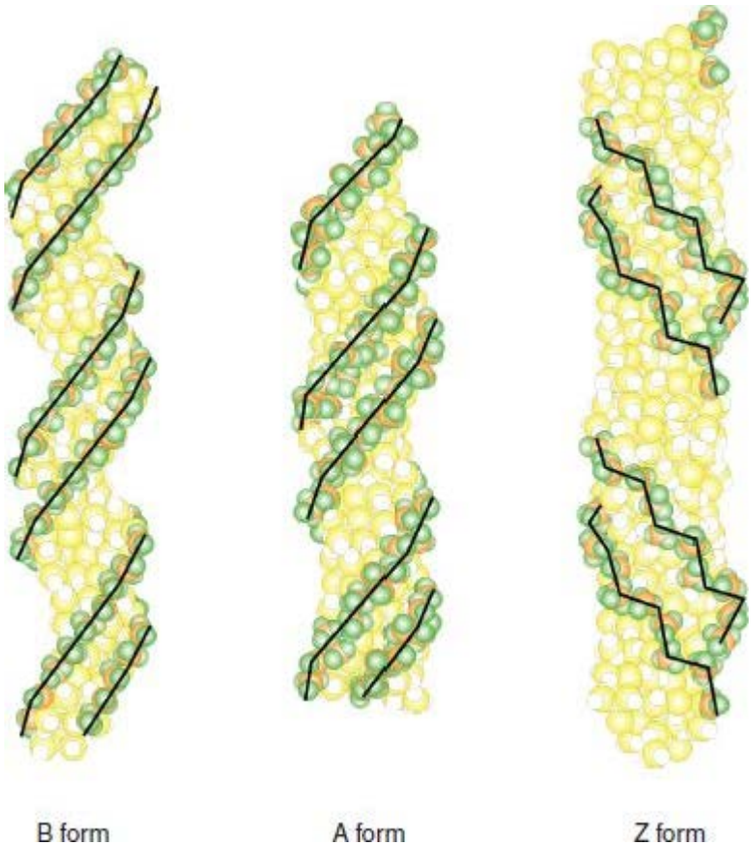


FIGURE 17.10

B-DNA, A-DNA, and Z-DNA

Because DNA is a flexible molecule, it can assume different conformational forms depending on its base sequence and/or isolation conditions. Each molecular form in the figure possesses the same number of base pairs. Refer to **Table 17.1** for the dimensions of these three DNA structures.

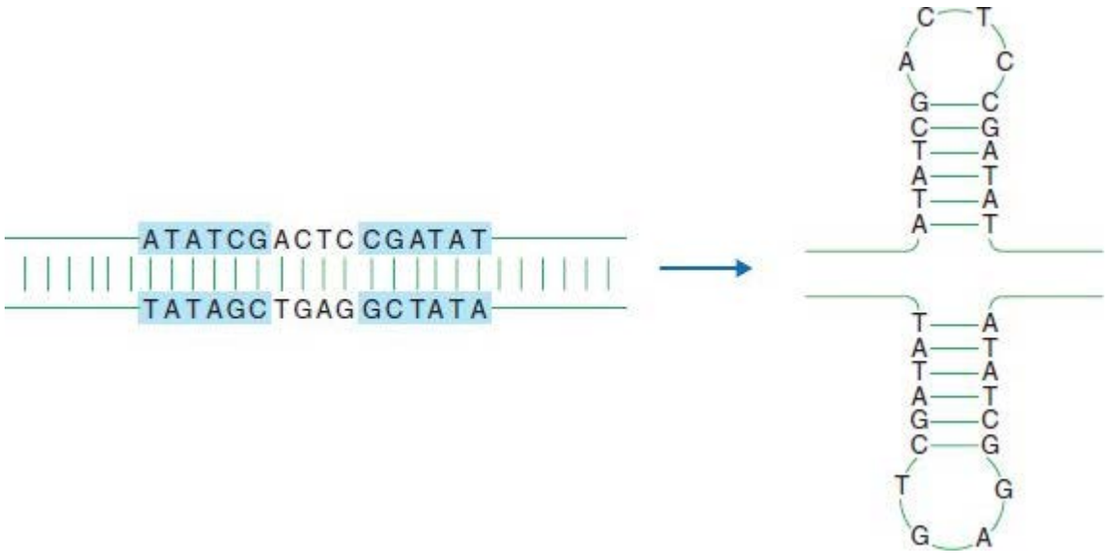
TABLE 17.1 Selected Structural Properties of B-, A-, and Z-Dna

	B-DNA (Watson–Crick Structure)	A-DNA	Z-DNA
Helix diameter	2.37 nm	2.55 nm	1.84 nm
Base pairs per helical turn	10.5	11	12
Helix rise per helical turn	3.32 nm	2.46 nm	4.56 nm
Helix rise per base pair	0.34 nm	0.24 nm	0.37 nm
Helix rotation	Right-handed	Right-handed	Left-handed

Z-DNA The Z form of DNA (named for its “zigzag” conformation) radically departs from the B form. **Z-DNA** (D = 1.84 nm), which is considerably slimmer than B-DNA (D = 2.37 nm), is twisted into a left-handed spiral with 12 bp per turn. Each turn of Z-DNA occurs in 4.56 nm, compared with 3.32 nm for B-DNA. DNA segments with alternating purine and pyrimidine bases (especially CGCGCG) are most likely to adopt a Z configuration. In Z-DNA, the bases stack in a left-handed staggered dimeric pattern, which gives the DNA a zigzag appearance and its flattened, nongrooved surface. Regions of DNA rich in GC repeats are often regulatory, binding specific proteins that initiate or block transcription. Although the physiological significance of Z-DNA is

unclear, it is known that certain physiologically relevant processes such as methylation and negative supercoiling (discussed on p. 649) stabilize the Z form. In addition, short segments have been observed to form as the result of torsional strain during transcription.

Certain segments of DNA have been observed to have higher-order structures. Cruciforms are an important example. As their name implies, *cruciforms* are cross-like structures. They are likely to form when a DNA sequence contains a palindrome, a sequence that provides the same information whether it is read forward or backward (e.g., "MADAM, I'M ADAM."). In contrast to language palindromes, the nucleotide base "letters" are read in one direction on one of the complementary strands of DNA and in the opposite direction on the other strand. One-half of the palindrome on each strand is complementary to the other half. The DNA sequences that form palindromes, which may consist of several bases or thousands of bases, are called *inverted repeats*. Cruciform formation, which occurs during DNA recombination (pp. 704–11) and DNA repair (pp. 699–704), begins with a small bubble, or *protocruciform*, and progresses as intrastrand base pairing occurs.



WORKED PROBLEM 17.2

The haploid human genome consists of approximately 3.2×10^9 bp (or 6.4×10^9 bp per diploid genome). Assuming that DNA is B-DNA, calculate the total length of a single cell's DNA.

SOLUTION

Given that the helix rise per base pair of B-DNA is 0.34 nm (**Table 17.1**), the total length of the DNA in a diploid cell is

$$6.4 \times 10^9 \text{ bp} \times 0.34 \times 10^{-9} \text{ m/bp} = 2.2 \text{ m}$$ ■

QUESTION 17.4

Compare the structural features that distinguish B-DNA from A-DNA and Z-DNA. What is known about the functional properties of these variants of B-DNA, the Watson–Crick structure?

.

DNA Supercoiling

DNA supercoiling, the over- or underwinding of DNA, facilitates several biological processes. Examples include packaging DNA into a compact form that can fit within cells, as well as replicating and transcribing DNA (Chapter 18). To undergo supercoiling, a single strand of double-helical DNA must be broken, or "nicked," and then either overwound or underwound before resealing. (A "nick" is a break in a single strand of double-helical DNA.) Because DNA supercoiling is a dynamic three-dimensional process, the information that two-dimensional illustrations can convey is limited. To understand supercoiling, therefore, consider the following thought experiment. A long, linear DNA molecule is laid on a flat surface. Then the ends are brought together and sealed to form an unpuckered circle (**Figure 17.11**). Because this molecule is sealed without under- or overwinding, the helix is said to be relaxed, and it remains flat on a surface. If the relaxed circular DNA molecule is held and twisted a few times, it takes the twisted shape shown in the figure. When this twisted molecule is returned to the flat surface and made to lie on the plane, it spontaneously rotates to eliminate the twist. Note that supercoiling is a mathematical concept borrowed from knot theory. The supercoiling of a DNA molecule is the sum of *twist* (the number of helical turns) and *writhe* (the number of times a DNA molecule crosses over itself).

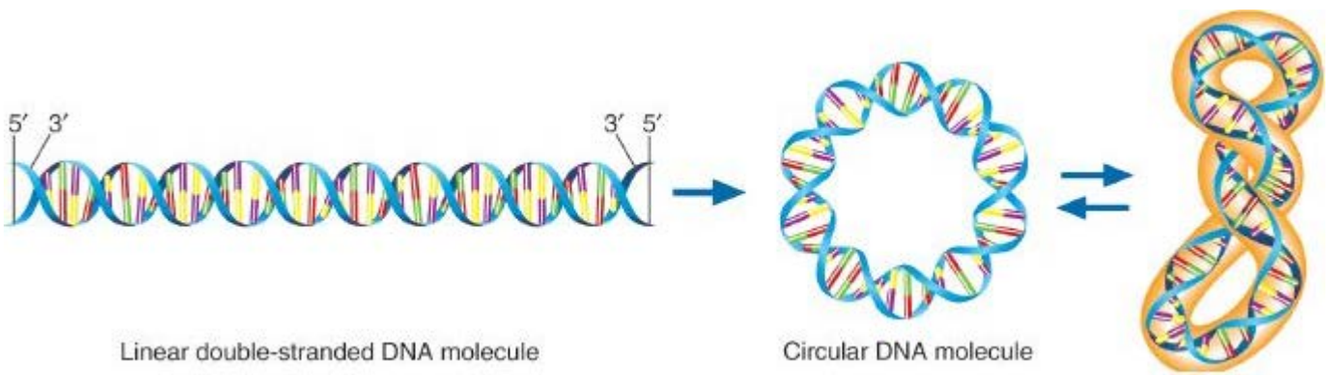


FIGURE 17.11

Linear and Circular DNA and DNA Winding

A linear DNA molecule circularizes to form a relaxed circular DNA molecule. When a relaxed circular DNA molecule is twisted, it reverts to its flat structure upon release.

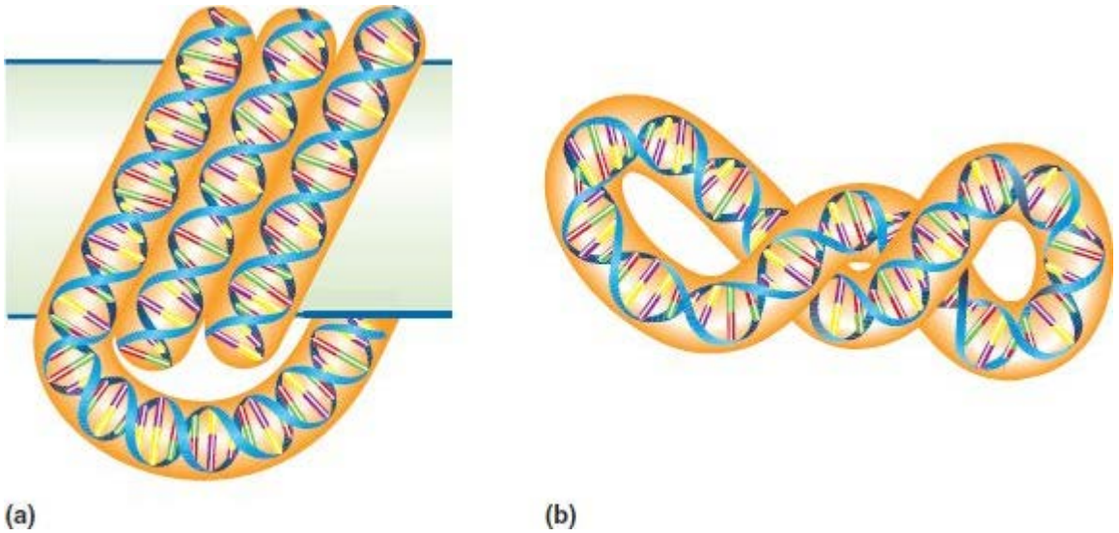


FIGURE 17.12

Supercoils

Supercoils occur in two major forms: (a) toroidal (spiral) and (b) plectonemic. In plectonemic DNA, the DNA coils are interwound, i.e. wrapped around each other.

When a linear DNA molecule is *underwound* (i.e., the right-handed DNA helix is twisted in a left-handed direction) and then sealed, the circular molecule twists to the right to relieve strain, with negative supercoiling as the result. Negatively supercoiled DNA (most naturally occurring DNA molecules) can form either of two interconvertible shapes: a toroidal supercoil or a plectonemic (interwound) supercoil (**Figure 17.12**). A negatively supercoiled DNA molecule stores potential energy in the form of torque (force that causes rotation). The stored energy, in turn, facilitates strand separation during processes such as DNA replication and transcription (**Figure 17.13**). An *overwound* DNA molecule (i.e., twisted in the right-handed direction before it is sealed to form a circle) twists to the left to relieve stress and is positively supercoiled. Positive supercoils that form during strand separation in DNA replication interfere with the replication machinery. Enzymes called topoisomerases (e.g., DNA gyrase in *E. coli*) make reversible cuts that allow the supercoiled DNA segments to relax.

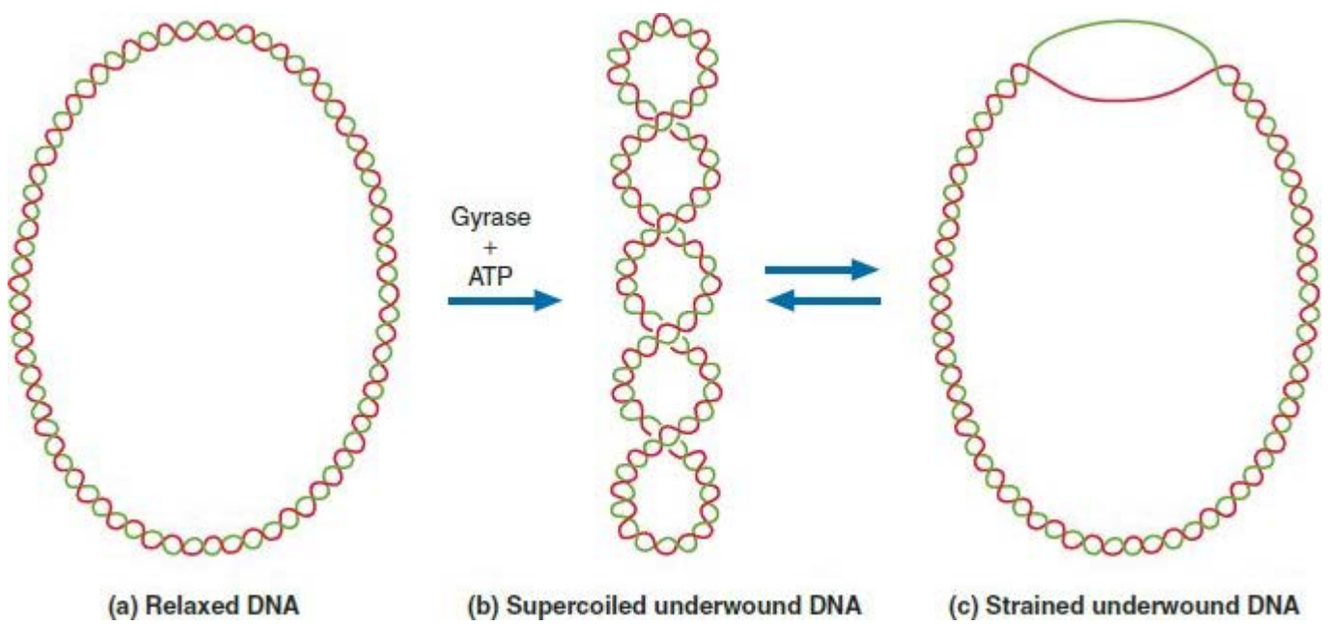


FIGURE 17.13

Effect of Strain on a Circular DNA Molecule

When a negatively supercoiled DNA molecule is forced to lie in a plane, the strain relieved by the formation of the negative supercoiling is reintroduced. Breakage and re-formation of a phosphodiester linkage allows the conversion of a relaxed circular form (a) to the negatively supercoiled (underwound) form (b). The strain relieved by the supercoiling process will be reintroduced when the underwound molecule is forced to lie in a plane (c).

DNA coiling can be compared with a coiled telephone cord (**Figure 17.14**). The coiled cord can only lie flat if it is rotated so as to undo the supercoils. Underwinding and overwinding can be observed if the hands of two individuals grasp each end of an unconnected telephone cord. One end is held stationary, while the second end is twisted. If twisting occurs in the same direction of the cord's coil (e.g., a right-handed coil twisted to the right), then the coil becomes overwound (positively supercoiled). Twisting in the opposite direction (e.g., a right-handed coil twisted to the left), the coil becomes underwound (negatively supercoiled).

FIGURE 17.14

Supercoiling

The telephone cord resembles a coiled DNA molecule in that it is a right-handed coil that can, through rotations introduced over time, form supercoils.

Chromosomes

DNA is packaged into structures called chromosomes. The term **chromosome** originally referred only to the dense, dark-staining structures visible within eukaryotic cells during meiosis or mitosis. However, this term is now also used to describe the DNA molecules that occur in prokaryotic cells. The physical structure and genetic organization of prokaryotic and eukaryotic chromosomes are significantly different.

PROKARYOTES In prokaryotes such as *E. coli*, a chromosome is a circular DNA molecule that is extensively looped and coiled so that it can be compressed by a factor of about 10,000 times into a

relatively small space (1 μm × 2 μm). Yet the information in this highly condensed molecule must be readily accessible. The *E. coli* chromosome (circumference 1.6 μm) consists of a supercoiled DNA that is complexed with a protein core (**Figure 17.15**).

In this structure, called the *nucleoid*, the chromosome is compacted by several types of nucleoid-associated proteins. Nucleoid-associated proteins are small, positively charged proteins that together serve to compress DNA structure and control chromosome organization by facilitating bending, wrapping, and looping processes.

In addition, the polyamines (polycationic molecules such as spermidine and spermine), found in all living organisms, assist in attaining the chromosome's highly compressed structure. The positively charged polyamines (p. 548) bind to the negatively charged DNA backbone, thus overcoming the charge repulsion between adjacent DNA coils.

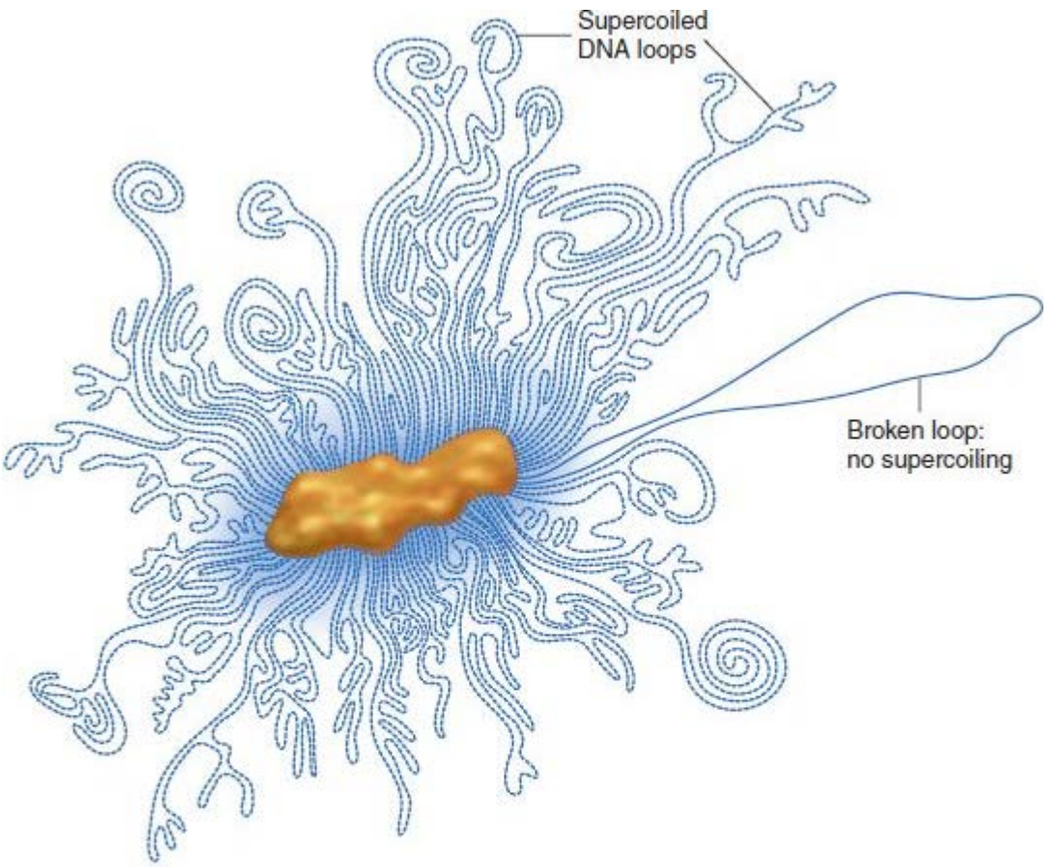


FIGURE 17.15

An *E. coli* Chromosome Removed from a Cell

The circular *E. coli* chromosome has been released from a gently lysed *E. coli* cell. Because the chromosome (3×10^6 bp) is highly supercoiled within the cell, the entire chromosome complex ordinarily measures only 2 μm across. Several types of nucleoid-associated proteins promote compaction through their effects on DNA looping and bending.

3D animation of spermidine

3D animation of spermine

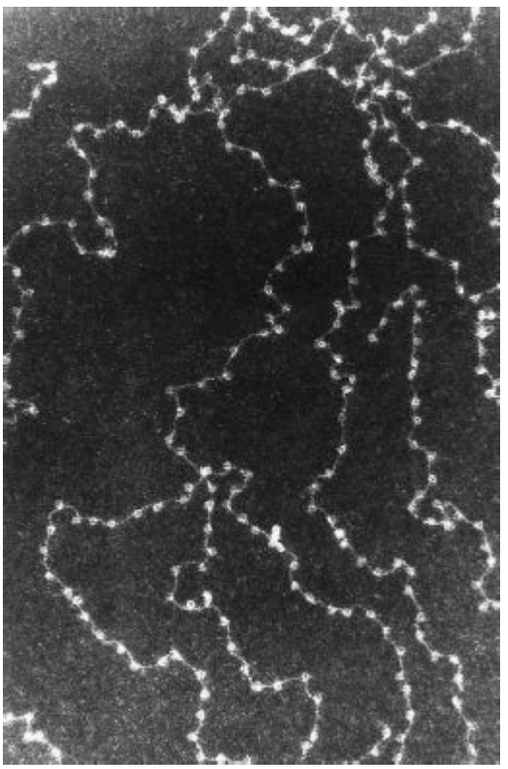


FIGURE 17.16

Electron Micrograph of Chromatin

A chromatin specimen spread out for electron microscopy. Note the “bead on a necklace” structure of this DNA-containing material.

EUKARYOTES In comparison to prokaryotes, the eukaryotes possess extraordinarily large genomes that are encased by the nuclear membrane (p. 56). Depending on species, the chromosomes of eukaryotes vary in both length and number. For example, humans possess 23 pairs of chromosomes and have a haploid genome of approximately 3 billion bp. The fruit fly *Drosophila melanogaster* has four chromosome pairs with 180 million bp, and corn (*Zea mays*) has 10 chromosome pairs with a total of 2.4 billion bp.

DNA molecules in eukaryotes are single linear DNA molecules complexed with histone proteins, referred to as **chromatin,** that condense to form chromosomes during cell division. During interphase of the cell cycle, decondensed chromatin appears as an amorphous mass in which individual chromosomes cannot be distinguished. During this phase, numerous nonhistone proteins are associated with chromatin. In addition to architectural proteins that create cell-specific three-dimensional chromatin structure (e.g., CTCF and cohesin, p. 653), chromatin is also associated with DNA replication and repair enzymes, chromatin remodeling proteins, and a vast

number of transcription factors.

HISTONES The histones are a group of small, basic proteins that bind to DNA, resulting in the formation of **nucleosomes**, the structural units of eukaryotic chromosomes. Consisting of five major classes (H1, H2A, H2B, H3, and H4), the histones are similar in their primary structure among eukaryotic species. They are particularly rich in the basic (i.e., positively charged) amino acids lysine and arginine. In electron micrographs, chromatin has a beaded appearance (**Figure 17.16**). Each of the "beads" is a nucleosome, which is composed of a positively supercoiled segment of DNA forming a toroidal coil around a histone octamer core (two copies each of H2A, H2B, H3, and H4).

Each of the highly conserved core histones (**Figure 17.17a**) contains a common structural feature called the *histone fold*: three α-helices separated by two short unstructured segments. It is the structure of the histone fold that allows the formation of histone octamers. The flexible N-terminal tails of core histones consist of between 25 and 40 amino acid residues that extend away from nucleosomes. Covalent modifications of tail residues of several types (e.g., acetylation and methylation) alter their interaction with nearby nucleosomes to facilitate either the compaction or the unfolding of nearby chromatin or modify the accessibility of DNA to proteins such as transcription factors. Such modifications are referred to as *epigenetic* modifications. (See Biochemistry in Perspective: Epigenetics and the Epigenome, pp. 662–64.)

The histone core formation begins when two sets of H2A and H2B form two head-to-tail heterodimers and H3 and H4 histones form two sets of head-to-tail heterodimers. The H3 H4 heterodimers then associate to form a $H3_2$ $H4_2$ tetramer (**Figure 17.17b**). The assembly of the nucleosome begins when the $H3_2$ $H4_2$ tetramer binds to DNA. When the two H2A H2B dimers associate with the tetramer, nucleosome assembly is complete (**Figure 17.17c**). One molecule of histone H1, a protein that does not contain the histone fold, binds to the nucleosome where the DNA enters and exits and acts as a clamp that prevents nucleosome unraveling (**Figure 17.18**). Approximately 146 bp (1.7 helical turns) are in contact with each histone octamer. Connection between adjacent nucleosomes is by means of linker DNA, which may vary between 20 and 70 bp, depending on species and tissue and even within the same cell.

CHROMATIN In anticipation of cell division, chromatin is compacted (about 10,000-fold) to form chromosomes. The nucleosomes are coiled into a higher order of structure referred to as the *30-nm fiber* (**Figure 17.19**). Despite extensive research, the precise internal structure of the 30-nm chromatin fiber is still unresolved, possibly because of irregularities caused by variation in linker length and histone posttranslational modifications.

In multicellular organisms, the transformation of a fertilized egg through successive mitotic divisions involves intricately regulated programmed gene expression resulting in diverse types of differentiated cells. Each of these gene expression changes involves alterations of chromatin structure.

During interphase of the cell cycle in complex organisms such as mammals, long DNA molecules (e.g., a total of 2 m in humans) are compacted into nuclei with average diameters of 6–10 mm, while remaining accessible to the enzyme complexes that perform vital processes (replication, transcription, and repair). Chromatin packing consists of an elaborate multilevel hierarchy (**Figure 17.20**). Each chromosome occupies its own restricted physical location, referred to as a **chromosome territory** (CT), in a cell type– and differentiation stage–specific, and nonrandom position in relation to other CTs. Gene-rich CTs are typically located near the nuclear interior, while gene-poor CTs tend to occur near the nuclear periphery. Interchromatin channels, which are continuous with nuclear pores, penetrate through and around CTs and serve as

reservoirs of molecules and complexes that perform essential functions. Examples include DNA replication and transcription factors (see pp. 697 and 725), and other regulatory proteins, small molecules such as nucleotides, and soon to be exported RNA–protein complexes. The *interchromatin channel network* also contains membraneless nuclear bodies such as nucleoli, speckles, and Cajal bodies (p. 58) that perform diverse functions (e.g., rRNA synthesis, RNA splicing, and ncRNA processing).

Within each CT, chromatin is organized into two types of compartments. Compartment A chromatin is gene-rich and lightly packed, thereby allowing access to regulatory proteins and DNA-transcribing RNA polymerases. Compartment B, consisting of telomere and centromere repetitive sequences, the majority of long interspersed nuclear elements (LINEs) sequences (p. 660), and cell-specific repressed genes, is tightly packed and, as a result, inaccessible to transcription machinery. The packing density of compartments A and B appears to connect them to the older concepts of **euchromatin** and **heterochromatin**, respectively. Euchromatin is a term that has been used to describe decondensed, interphase chromatin with varying levels of transcriptional activity. **Constitutive heterochromatin** is composed of repetitive sequences that are permanently transcriptionally inactive. **Facultative heterochromatin** consists of tightly packed genes (the repressed genes in compartment B) that can lose their condensed structure and be transcribed in response to a signaling mechanism or are packed as euchromatin in other cell types.

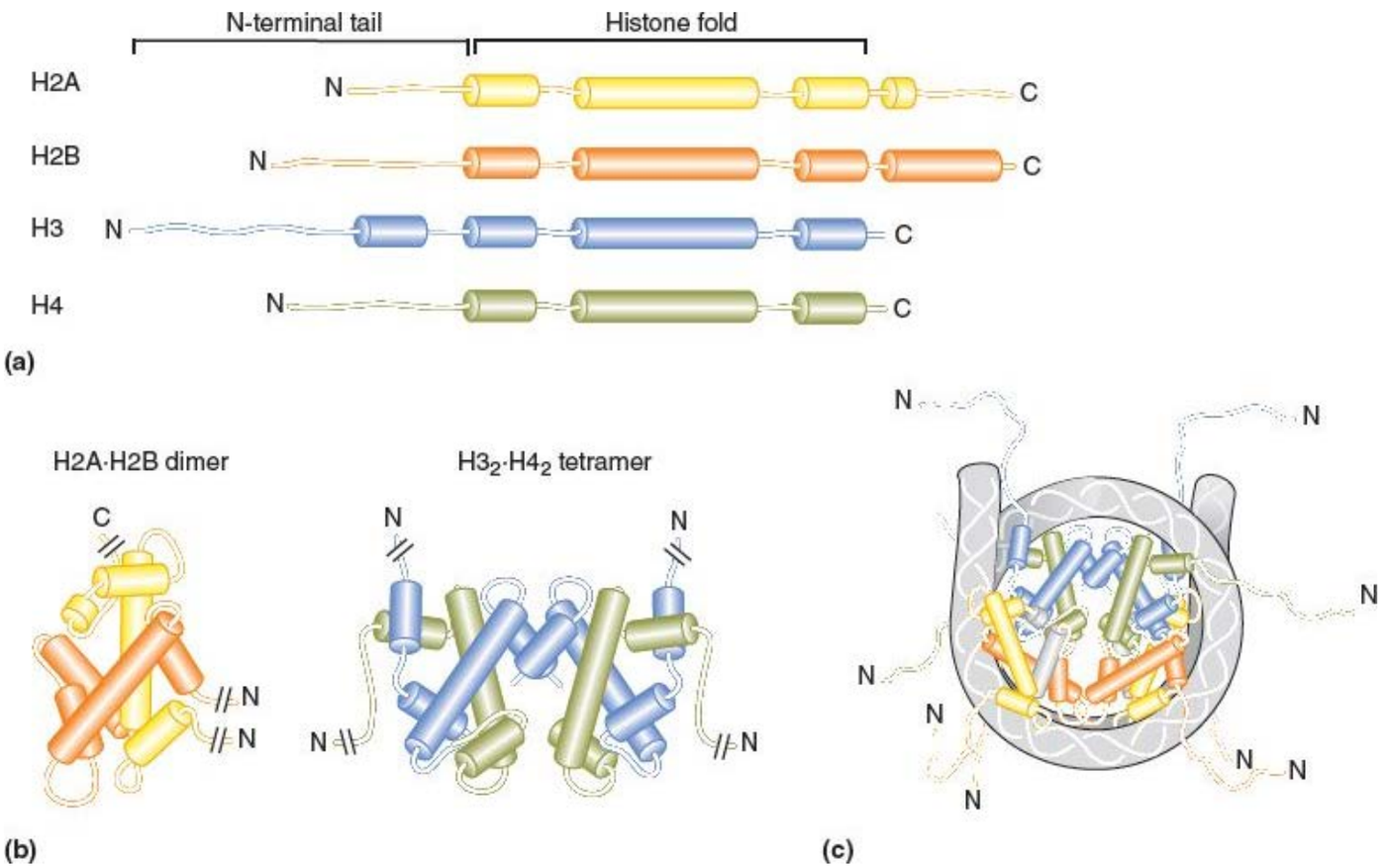


FIGURE 17.17

The Core Histones

Each nucleosome contains eight histone molecules: two each of H2A, H2B, H3, and H4. Each of these molecules contains a globular domain referred to as the histone fold and a long unstructured N-terminal domain. (a) These domains are illustrated as linear molecules with cylinders representing α-helices. (b) The

formation of the histone core begins with the association of two sets of H2A and H2B to form two heterodimers and two molecules each of H3 and H4, combining to form a tetramer. The $H3_2$•$H4_2$ tetramer then binds to DNA. The nucleosome structure is complete (c) when 2 H2A•H2B heterodimers bind to the tetramer.

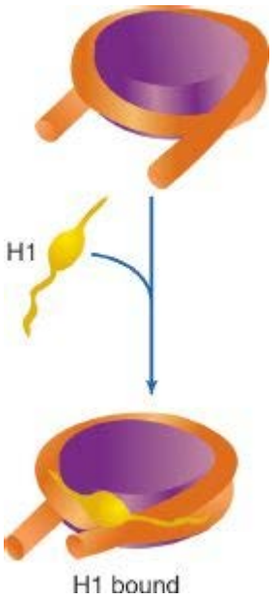


FIGURE 17.18

Histone H1

The binding of H1 to two different sites on the nucleosome DNA stabilizes the wrapping of DNA around the histone octamer.

Each chromosomal compartment is organized into **topologically associated domains** (TADs), with lengths that range from tens of thousands of base pairs to over 1 Mb (million base pairs). TADs, believed to be the basic unit of chromatin folding, are delimited by boundary sequences that are conserved across cell types and often among species (**Figure 17.21**). According to the *loop extrusion model*, TAD formation begins with the binding of CTCF (CCCTC-binding factor) with a cognate DNA sequence. Cohesin, a ring-shaped protein complex, is recruited by CTCF. It then encircles the DNA and begins extruding a DNA loop through the cohesin ring. Loop formation ends when cohesin encounters another CTCF bound to a CTCF sequence with an opposite orientation to that of the first CTCF sequence.

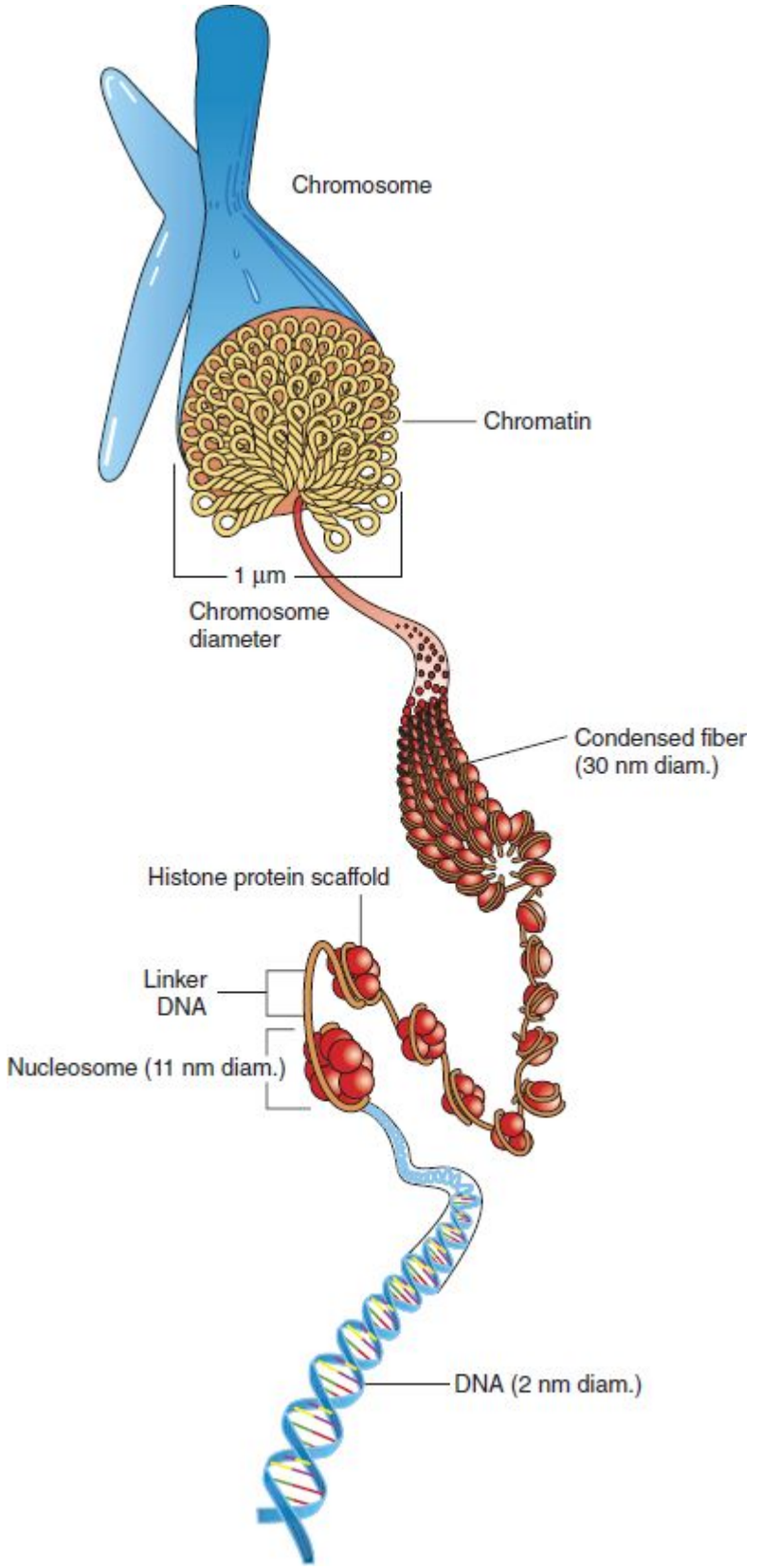


FIGURE 17.19

Chromatin

Nuclear chromatin contains many levels of coiled structure. Despite decades of research efforts, higher-order chromatin structure is largely unresolved.

TADs within euchromatin are believed to stabilize long-range interactions of inducible genes

with their regulatory sequences (i.e., promoters and enhancers; p. 658). Many of the DNA sequences outside of these TAD loops are short interspersed nuclear elements (SINEs; decayed virus sequences, p. 660) or *housekeeping genes*. Housekeeping gene is a term that refers to molecules such as ribosomal proteins, NADH dehydrogenase, and succinate dehydrogenase that are routinely transcribed because they are required for maintaining basic cellular function. Heterochromatic TADs are positioned in close proximity either to the nuclear lamina, in which case they are referred to as *lamina associated domains* (LADs), or to the nucleolus, in which case they are referred to as *nucleolus associated domains* (NADs).

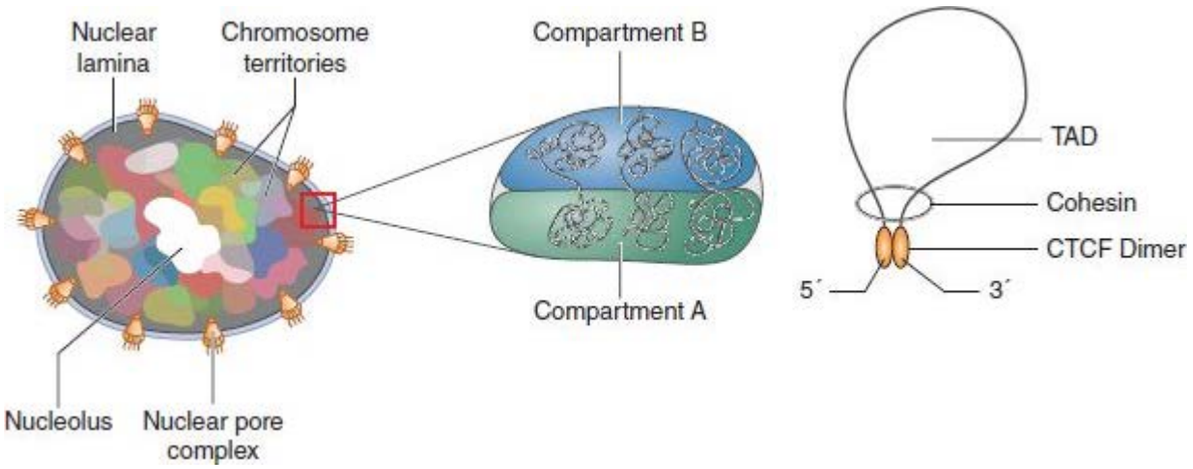


FIGURE 17.20

Hierarchical Nuclear Genome Organization

The interphase nucleus, investigated with high-resolution microscopy and fluorescence in situ hybridization (FISH), has revealed that each chromosome occupies its own territory in a specific location. Each chromosome is divided into one or more compartments—As (euchromatin) and Bs (heterochromatin)—depending on differences in packing density. Topologically associated domains (TADs), another level of chromatin folding, are formed by architectural proteins, most notably CTCF and cohesion. TADs can apparently be hierarchical in structure since mini loops, called sub-TADs (not shown), have been observed.

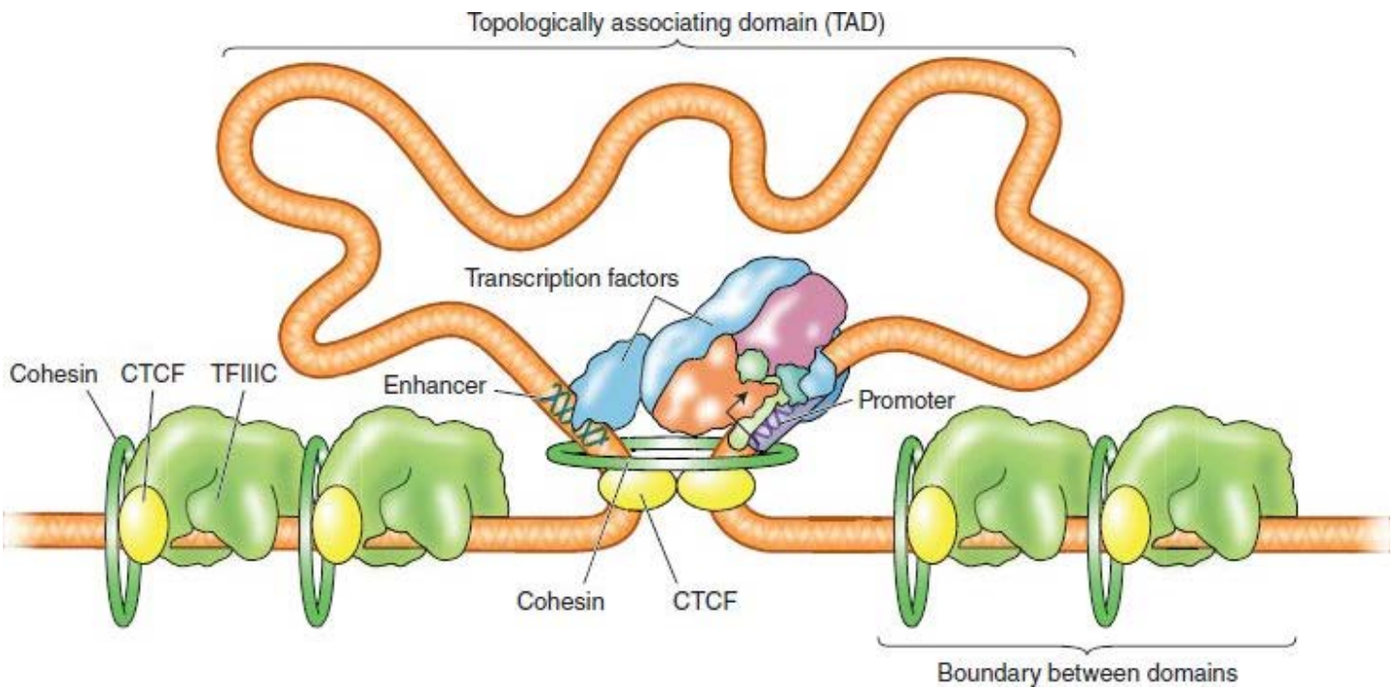


FIGURE 17.21

TAD Chromosomal Boundary Structure

TAD formation, which facilitates the interaction of a gene's promoter and enhancer sequences with transcription factors, is effected in part by CTCF and cohesin. The arrow indicates the gene's transcription start site. CTCF and cohesin in combination with TFIIIC (transcription factor IIIC) establish borders that prevent interactions of nearby genes with a TAD's enhancers.

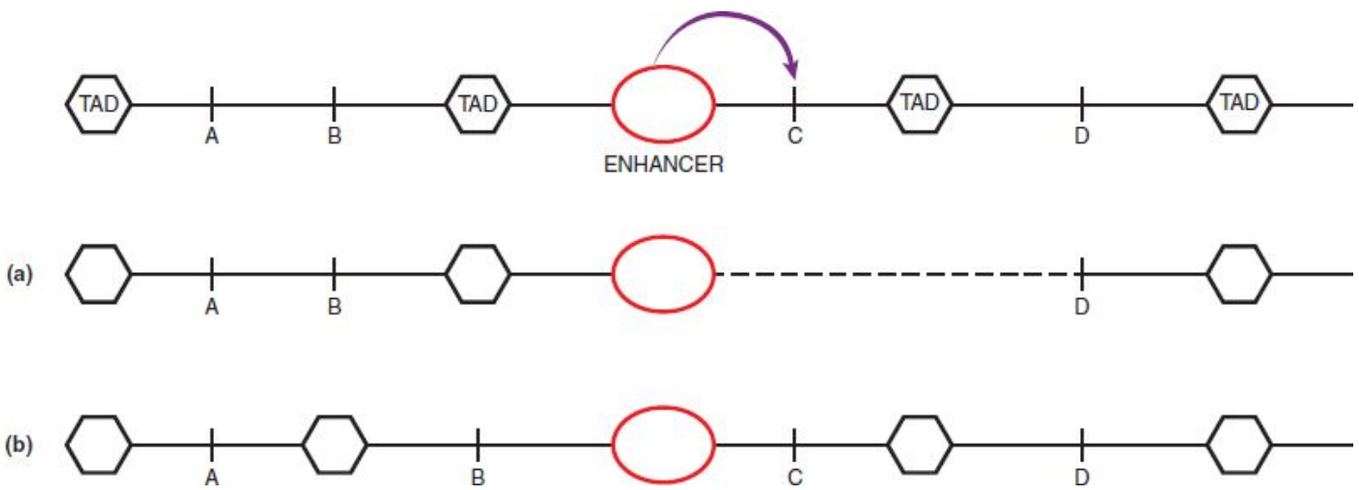
The mechanism by which chromatin reversibly condenses is unresolved, but DNA methylation, several histone covalent modifications, and several chromatin remodeling enzymes are known to play significant roles.

QUESTION 17.5

Explain the hierarchical relationships among the following: genomes, genes, nucleosomes, chromosomes, and chromatin.

QUESTION 17.6

The following diagrams illustrate the relationship between genes that affect limb structure: TAD boundaries (hexagons) and enhancers (ovals). In the normal DNA sequence, the enhancer in the diagram facilitates the transcription of gene C but does not affect transcription of genes *A*, *B*, and *D* because of TAD borders. In (a), a deletion occurs (dotted line) that results in a condition referred to as brachydactyly (shortened fingers and toes). In (b) the inversion of a TAD and gene *B* results in a form of syndactyly (fused fingers and toes). After evaluating the diagrams, give a reason for each syndrome. [gene *A* = WNT6; gene *B* = IHH; gene *C* = EPHA4; gene *D* = PAX3]



ORGANELLE DNA Mitochondria and chloroplasts are semiautonomous organelles; that is, they possess DNA and their own version of protein-synthesizing machinery. Both of these organelles, descended from free-living prokaryotes, have small genomes and require a substantial number of proteins coded for by the nuclear genome. In most mammalian cells, there are between 100 and 10,000 copies of mitochondrial DNA (mtDNA). Each mtDNA has 37 genes that code for 2 rRNAs, 22 tRNAs, and 13 proteins used in either electron transport or ATP synthesis. Chloroplast

DNA (cpDNA) typically has about 150 genes. Among these are genes that code for 3 rRNAs, 31 tRNAs, 4 RNA polymerases, 21 ribosomal proteins, the large rubisco subunit, and diverse thylakoid proteins that are components of PSI, PSII, and ATP synthase.

KEY CONCEPTS

- Each prokaryotic chromosome consists of a supercoiled circular DNA molecule complexed to a protein core.
- Each eukaryotic chromosome consists of a single linear DNA molecule that is complexed with histones and other proteins to form chromatin.

Genome Structure

The genome is an organism's operating system, the full set of inherited instructions required to sustain all living processes. Within each genome are the genes, the units of inheritance that determine the primary structure of gene products (polypeptides and RNA molecules). Genomes differ in size, shape, and sequence complexity. Genome size—the number of base-paired nucleotides, which is loosely related to organismal complexity—varies over an enormous range from less than 10^6 bp in some species of *Mycoplasma* (the smallest known bacteria) to greater than 10^{10} bp in certain plants. Most prokaryotic genomes are smaller than those of eukaryotes. In contrast to prokaryotic genomes, which typically consist of circular DNA molecules, eukaryotic genomes are divided into two or more linear DNA molecules. The most significant difference between prokaryotic and eukaryotic genomes, however, is the vastly larger information-coding capacity of eukaryotic DNA. Amazingly, the majority of eukaryotic sequences do not code for gene products. For this reason, each type of genome will be considered separately. Short segments of the genomes of several eukaryotes are compared with those of *E. coli* in **Figure 17.22**.

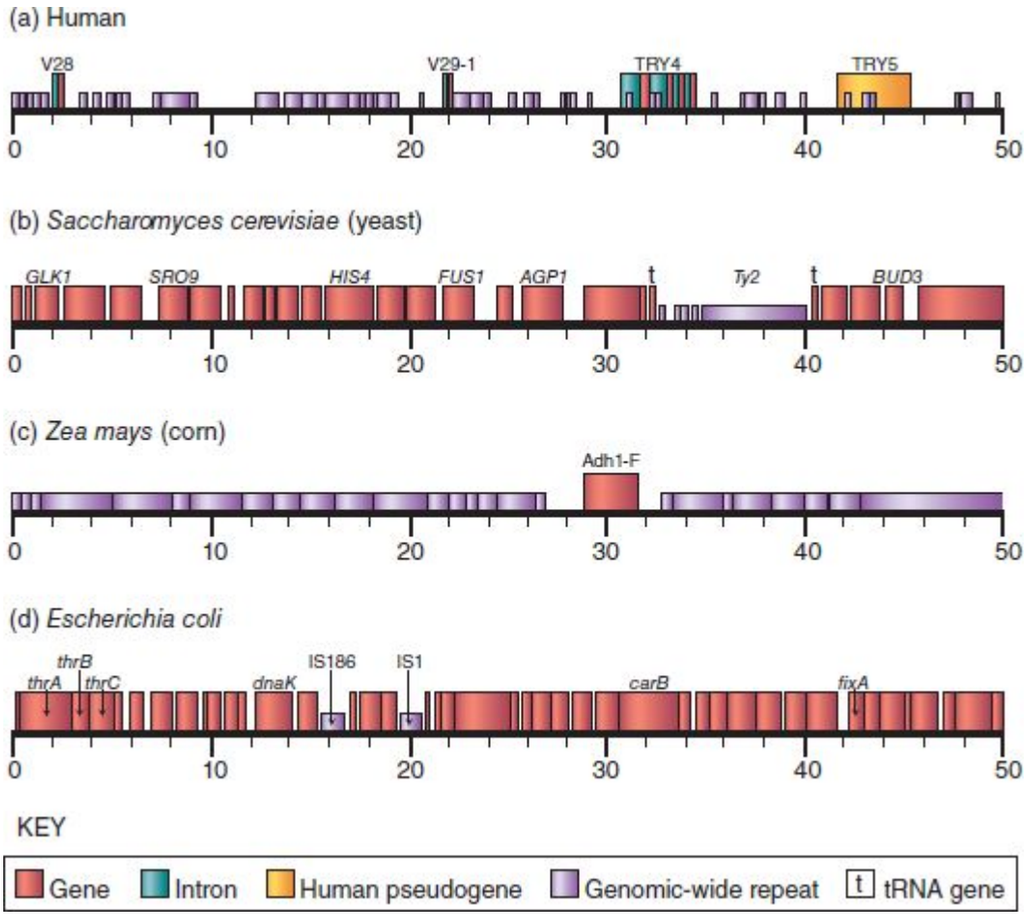


FIGURE 17.22

Comparison of 50-kb Segments of the Genomes of Selected Eukaryotes with the Prokaryote *E. coli* Genome

As indicated, the genomes of organisms such as (a) humans, (b) *Saccharomyces cerevisiae*, (c) maize, and (d) *E. coli* can vary considerably in their complexity and gene density. Genes are indicated by letters and/or numbers. Humans and other complex eukaryotes have genes that are interrupted with sequences such as introns and nonfunctional sequences called pseudogenes that resemble true genes. Bacteria have few, if any, genome-wide repeats (repetitive, noncoding segments).

PROKARYOTIC GENOMES Investigations of prokaryotic chromosomes, especially those of several strains of *E. coli*, have revealed the following features.

1. **Genome size**. As described, most prokaryotic genomes are relatively small, having considerably fewer genes than those of eukaryotes. The K12 *E. coli* chromosome contains about 4.6 megabases, which code for 4377 protein-coding genes and at least 109 ncRNAs [1 megabase (Mb) = 1×10^6 bases].
2. **Coding capacity**. The genes of prokaryotes are compact and continuous; that is, they contain only about 15% noncoding DNA sequences. This is in sharp contrast to human DNA, in which more than 80% of DNA can be in noncoding form. Prokaryotic genomes do contain numerous ncRNA-coding sequences, which function as regulators of cellular processes. During nutrient deprivation, for example, the small RNA (sRNA) molecule 6S binds to and inhibits RNA polymerase, the DNA-dependent enzyme complex that converts DNA sequences into RNA. Other bacterial ncRNAs function as riboswitches (p. 735), noncoding segments of mRNAs that control gene expression in response to cell conditions.
3. **Gene expression**. The organization of functionally related genes into operons enhances

their regulation. An **operon** is a set of functionally linked genes, the transcription of which is regulated as a unit. About one-fourth of the genes of *E. coli* are organized into operons (see pp. 733–34).

Recall that prokaryotes also often possess additional small pieces of DNA (see p. 44) called *plasmids*, which are usually, but not always, circular. Plasmids typically have genes that are not present on the main chromosome and are seldom essential for bacterial growth and survival. They may, however, code for biomolecules that provide the cell with a growth or survival advantage: antibiotic resistance, unique metabolic capacities (e.g., nitrogen fixation or degradation of unique energy sources such as aromatic compounds) or virulence (e.g., toxins or other factors that undermine host defense mechanisms).

EUKARYOTIC GENOMES The organization of genetic information in eukaryotic chromosomes is substantially more complex than that observed in prokaryotes. Eukaryotic nuclear genomes possess the following unique features:

1. **Genome size**. Although eukaryotic genomes tend to be substantially larger than those of prokaryotes, genome size alone is not necessarily a measure of the complexity of the organism. For example, the haploid genome of humans is 3200 Mb. The genomes of peas and the salamander are 4800 and 40,000 Mb, respectively.
2. **Coding capacity**. Although eukaryotic genomes have relatively enormous capacity, only a small fraction is devoted to coding for proteins (e.g., only 1.5% of DNA sequences in the human genome). Until recently, much of the remaining DNA sequences were thought to be nonfunctional "junk." ENCODE (*Enc*yclopedia *o*f *D*NA *E*lements), a decade-long research project funded by the U.S. Human Research Institute, revealed that 80% of the human DNA sequences have biological functions. In addition to DNA sequences that directly control protein-coding genes, an enormous number of DNA sequences code for diverse types of ncRNAs that regulate every facet of genome function.
3. **Coding continuity**. Most eukaryotic genes are discontinuous. Noncoding sequences (called **introns** or intervening sequences) are interspersed between sequences called **exons** (expressed sequences), which code for part of a gene product (e.g., a polypeptide, or a ncRNA). Intron sequences, which may together be significantly longer than the exons in a gene, are removed from primary RNA transcripts by a splicing mechanism (Section 18.2) to produce a functional RNA molecule.

The existence of exons and introns enables eukaryotes to produce more than one polypeptide from each protein-coding gene. By utilizing a process called *alternative splicing* (p. 736), various combinations of exons can be joined together to form a series of mRNAs. For example, the random rearrangements of gene sequences that encode the antigen receptors of the immunoglobulins have a major role in generating the millions of antibodies produced by the mammalian immune system.

DNA REGULATORY SEQUENCES The major types of DNA regulatory sequences are promoters, enhancers, silencers, and insulators. **Promoters** are DNA sequences (100–1000 bp) in close proximity to a transcription start site of a specific gene. For a gene to be transcribed, certain transcription factors bind to the gene's promoter and then recruit an RNA polymerase complex to the site. An **enhancer** is a DNA sequence (50–1500 bp) that, when bound to an activator transcription factor interacts with and stimulates the activity of an RNA polymerase complex. When it is bound to a repressor protein, a **silencer** sequence inhibits the transcription of its gene by preventing an RNA polymerase from binding to the promoter. An **insulator** sequence binds to

an insulator-binding protein (e.g., CTCF), which in turn blocks the interaction between enhancers and the promoters of neighboring genes. Insulators also prevent the spread of heterochromatin to active genes.

PSEUDOGENES A **pseudogene** is defined as a nonfunctional DNA sequence that is homologous to a known protein or *RNA* gene. It is estimated that humans have at least 17,000 pseudogenes, many of which we share with other primates. There are three principal types of pseudogenes: nonprocessed, processed, and disabled. A nonprocessed pseudogene is the result of a gene duplication, followed by a series of disabling mutations that eventually render it nonfunctional. A processed pseudogene is the result of retrotransposition, an event in which a segment of an mRNA is reverse transcribed into a DNA copy that is then inserted into a chromosome. Processed pseudogenes are so called because they have the characteristics of processed mRNAs (i.e., the absence of promoter sequences and introns and the presence of a poly(A) tail). A disabled pseudogene is a gene that is not expressed because of disabling mutations but has not been duplicated. The classic example is the human gene for the enzyme L-gulono-γ-lactone oxidase, which catalyzes the conversion of L-gulono-γ–lactone to L-ascorbate (p. 253).

Genome-wide investigations have revealed that some pseudogenes are actively transcribed, and there are several examples of pseudogene transcripts with biological activity. For example, PTENP1 is a processed pseudogene of the tumor suppressor gene *PTEN*, which codes for a cell cycle phosphatase that prevents rapid cell division by negatively regulating the PKB (Akt) signaling pathway (p. 612). PTENP1 and PTEN transcripts can competitively bind to a specific set of *oncomirs* (cancer-promoting miRNAs; see p. 738) in a process that protects PTEN mRNA from degradation. Reduced synthesis or elimination of PTENP1 transcripts has been observed in numerous cancers.

REPETITIVE DNA The term *repetitive DNA* refers to patterns of DNA that occur in multiple copies throughout a genome. There are two general classes: tandem repeats and interspersed genome-wide repeats.

Tandem repeats are DNA sequences in which one or more nucleotides are repeated with multiple copies arranged next to each other. These sequences were originally referred to as **satellite DNA** because they form a separate, or "satellite," band when genomic DNA is broken into pieces and centrifuged to separate the fragments by density gradient centrifugation (see Biochemistry in the Lab: Nucleic Acid Methods, pp. 666–70). Total lengths of the tandem repeats often vary between 10^5 and 10^7 bp.

Certain types of tandem repeats apparently play structural roles in **centromeres** (a specialized DNA sequence that contains a kinetochore, which attaches a chromosome to the mitotic spindle during mitosis and meiosis) and **telomeres** (structures at the ends of chromosomes that buffer the loss of critical coding sequences after a round of DNA replication). Two relatively small repetitive sequence types are referred to as minisatellites and microsatellites. **Minisatellites** have tandemly repeated sequences of 10 to 100 bp, with total lengths between 10^2 and 10^5 bp. Telomeres, which have a high mutation rate, contain minisatellite clusters. In **microsatellites**, also referred to as single-sequence repeats, there is a core sequence of 1 to 4 bp that is tandemly repeated from 10 to 100 times. The functions of these repetitive sequences are for the most part unknown. Because of their large number in genomes and because they are pleomorphic (i.e., they vary with each individual organism), human minisatellites and microsatellites are used as markers in genetic disease diagnosis, in kinship and population studies, and in forensic investigations (see Biochemistry in Perspective: Forensic Investigations, pp. 671–72).

As their name implies, **interspersed genome-wide repeats** are repetitive sequences that are

scattered around the genome. Most of these sequences are the result of **transposition** (Section 18.1), a mechanism whereby certain DNA sequences, referred to as **mobile genetic elements** or **transposons**, can be duplicated and enabled to move within the genome. Transposons are a significant feature of eukaryotes, where they have expanded and often rewired gene regulatory networks on an evolutionary time scale. Often classified as junk DNA in the recent past, transposable elements are now recognized as a major feature of animal genomes. For example, transposons are a significant component of mammalian genomes (e.g., almost 50% of the human genome) where they exert a diverse variety of regulatory influences.

There are two types of eukaryotic mobile genetic elements: DNA transposons and RNA transposons. A DNA transposon can move around genomes using its own transposase enzyme to excise itself from one location and then insert into a new location (p. 709). Active DNA transposons occur in bacterial and numerous eukaryotic genomes but are considered to be fossils in humans where they make up 3% of the genome. Unlike DNA transposons, the transposition mechanism of **RNA transposons** or **retrotransposons** involves an RNA transcript intermediate. Retrotransposons can be classified into two groups based on the presence or absence of long terminal repeats (LTRs), sequences that are involved in reverse transcription.

LTR retrotransposons, believed to be decayed viruses, are often referred to as **endogenous retroviruses** (ERVs). The yellow and obese Agouti mouse (p. 664), for example, is the consequence of the insertion of an LTR retrotransposon 100 kb upstream of the Agouti gene. This gene codes for AgRP (pp. 624 and 664), a polypeptide that interferes with melanin synthesis and appetite suppression. When the retrotransposon is active (unmethylated), a cryptic promoter in its 5′LTR permits the unrestricted transcription of the Agouti gene. Humans also possess endogenous retroviruses (HERVs), which compose about 8% of the genome and are believed to be the result of ancient infections of germ cells (eggs and sperm). Although HERVs are in general inactive, several functional HERV sequences have been identified. For example, the fusion protein syncytin 1, a product of a member of the HERV-W family, is expressed during development of the placenta. Inadequate synthesis of syncytin 1 is one of several factors that cause *preeclampsia*—pregnancy-induced hypertension that may lead to maternal mortality.

Preeclampsia

Non-LTR retrotransposons with lengths greater than 5 kb are called **LINEs**. LINE sequences contain a strong promoter, an integration sequence (a base sequence required for insertion into another DNA molecule), and the coding sequences for transposition enzymes. LINEs, which constitute 21% of the human genome, have undergone duplication and mutation over time, and only a small percentage of them (about 0.1%) are at all functional. One in every 1200 mutations in humans is estimated to be the result of a LINE insertion. One example is hemophilia A (a blood-clotting disorder), which results when a LINE sequence inserts into the gene for clotting factor VIII.

SINEs are non-LTR retrotransposons with less than 500 bp. Although SINEs contain integration sequences, they cannot undergo transposition without the aid of a functional LINE sequence. SINEs have greatly expanded over time to comprise almost 14% of the human genome with more than 1.7 million copies. *Alu* elements (300 bp in length) are primate-specific SINEs. Of the 1.2 million Alus in the human genome, 7000 Alu insertions are uniquely human. Alus have been observed to affect mRNA splicing (removal of introns, p. 658) and translation. A significant

percentage of Alus occur within introns. Alus also mediate chromosome rearrangements, insertions, deletions, and recombinations. Mutations mediated by the Alu subfamily of SINEs have accounted for more than 20 different human genetic diseases. Alu-mediated insertions have been observed to cause hemophilia B (defective clotting factor IX), and unequal recombinations resulting from Alu-mediated insertions have been linked to Lesch–Nyhan disease (p. 559) and Tay−Sachs disease (p. 415). *Recombination* (Section 18.1) is the shuffling of DNA sequences involving crossing over of homologous chromosomes in the cells that give rise to egg or sperm cells. Unequal recombination is an aberrant process that causes insertion or deletion mutations.

KEY CONCEPTS

- In each organism's genome, the information required to direct living processes is organized for efficient storage and use.
- Genomes from different types of organism differ in their sizes and levels of complexity.

QUESTION 17.7

Compare the sizes and coding capacity of prokaryotic genomes with those of eukaryotes. What other features distinguish them?

THE HUMAN GENOME Of the approximately 3200 Mb of the human genome, 1.5% codes for approximately 20,000 protein-coding genes. The remainder of the genome (98%) is referred to as noncoding DNA (ncDNA). The functions of almost one-quarter of known protein-coding genes (**Figure 17.23**) are related to DNA synthesis and repair and gene expression. Signal transduction proteins are coded for by about 21% of genes and about 17% code for general biochemical functions of cells (i.e., metabolic enzymes). The remaining genes, about 38%, code for an array of proteins involved in transport processes (e.g., ion channels), protein folding (molecular chaperones and proteosomal subunits), structural proteins (e.g., actin, myosin, tubulin, and cell adhesion proteins), and immunological proteins (e.g., antibodies that modulate inflammatory responses). The functions of almost one-quarter of protein-coding genes are as yet unknown. Protein-coding genes are not distributed evenly among the chromosomes. For example, chromosomes 1, 11, and 19 contain about 22% of all human protein-coding genes.

About 80 to 90% of the human genome consists of intergenic or noncoding sequences. Noncoding DNA sequences are diverse in their functions. Examples include ncRNA genes (pp. 673–75), introns, and UTRs (untranslated regions before and after the coding sequence in an mRNA). Noncoding DNA also includes regulatory DNA sequences (promoter, enhancer, silencer, and insulator sequences), pseudogenes, and repetitive DNA.

17.2 RNA

RNA is an extraordinarily large and diverse group of molecules that are synthesized by the transcription of DNA sequences. In addition to their well-known roles in protein synthesis, RNAs are amazingly versatile molecules that perform functions previously thought to be exclusively performed by proteins. Examples include gene expression regulation, cell differentiation, and catalysis. The primary structure of polyribonucleotides is similar to that of their DNA

counterparts, but there are several differences.

1. The sugar moiety of RNA is ribose instead of deoxyribose in DNA. The presence of the 2′-OH group of ribose makes RNA more reactive than DNA.
2. The nitrogenous bases in RNA differ somewhat from those observed in DNA. Instead of thymine, RNA molecules use uracil. In addition, the bases in some RNA molecules are modified by a variety of enzymes (e.g., methylases, thiolases, and deaminases).

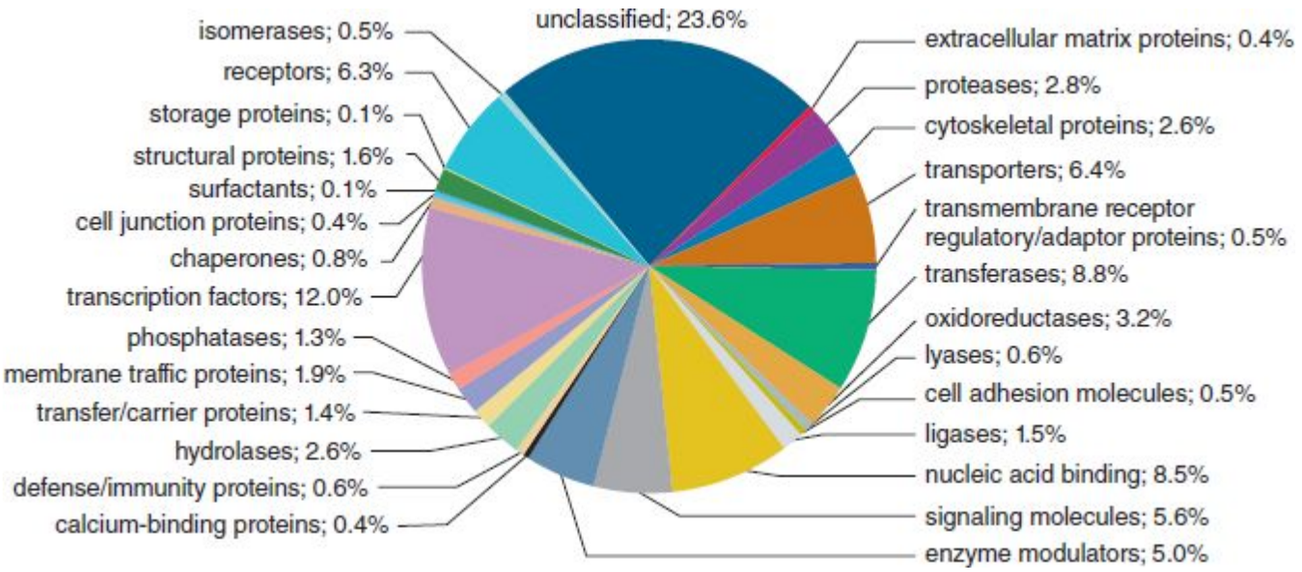


FIGURE 17.23

Human Protein-Coding Genome

Human genes are annotated for function and their percentage of all protein-coding genes.

Biochemistry IN PERSPECTIVE

Epigenetics and the Epigenome: Genetic Inheritance beyond DNA Base Sequences

How do covalent modifications of DNA and histones affect the functions of multicellular organisms? How do the more than 200 cell types in humans arise from a fertilized egg? Life scientists have known for many years that the transformation of a single cell into a multicellular organism is the result of cell specialization effected by gene expression changes that occur during the developmental process. Early signal mechanisms must “instruct” cells, each with an identical genetic blueprint, to progress down separate developmental pathways to yield terminally differentiated cells such as red blood cells, neurons, or skeletal muscle cells. In recent years, it has become apparent that this process, the result of sequential, programmed changes in the pattern of expressed and silenced genes in each cell type, does not depend on

genetic information (DNA base sequences) alone. Rather, development is the result of chromatin remodeling that is accomplished by two mechanisms: DNA methylation and histone covalent modifications. Because covalent modification–induced gene activations and repressions are heritable but do not change DNA base sequences, this phenomenon is referred to as **epigenetics** [*epi* (Gk) = over or above]. Epigenetic modifications convert affected DNA sequences within facultative heterochromatin into transcriptionally active euchromatin or vice versa. Each differentiated cell type has unique epigenetic modifications that are referred to as its **epigenome**. After a brief description of epigenetic modifications, the role of epigenetics is discussed as an interface between genomes and the environment.

DNA Methylation

In DNA methylation reactions, a methyl group is donated by SAM (p. 546) to carbon-5 of cytosine residues (**Figure 17A**). In mammals, methylated cytosines occur predominantly in 5′-CG-3′ sequences, which are referred to as CpG dinucleotides or **CpGs**. The C-5 methyl groups of methylated cytosine residues (5-MeCpG) protrude into the major groove where they prevent binding of certain DNA-binding proteins (i.e., transcription factors). They also enable the binding of proteins, called **methyl-CpG-binding proteins** (MeCPs) that promote heterochromatin formation. CpGs are relatively rare in mammalian genomes. However, there are CpG-rich regions, called **CpG islands**, in which CpGs are typically about 50% of bases. CpG islands are found in or near about 40% of mammalian promoter sequences. They often occur upstream of constitutively expressed (continuously produced) genes and some regulated genes. The methylation of CpG islands represses gene expression.

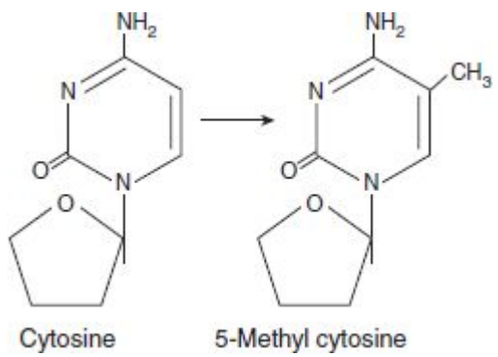


FIGURE 17A

Cytosine Methylation

Cytosine residues in CpG dinucleotides are methylated by specific methyltransferases.

There are two classes of CpG methylating enzymes: maintenance methyltransferases and de novo methyltransferases. *Maintenance methyltransferases* recognize methylated CpGs in the parental DNA strand and then catalyze the methylation of cytosines in the corresponding CpGs in the newly synthesized strand. It is this process that is responsible for the stable inheritance of DNA methylation patterns between cell generations. The addition of methyl groups to previously unmodified CpGs is catalyzed by de novo *methyltansferases*, usually in response to various signal transduction mechanisms. Cytosine demethylation is catalyzed by several methyl dioxygenases to yield 5-carboxycytosine, which is then removed and replaced with cytosine by a DNA repair enzyme.

Histone Modifications

Histones have a featured role in epigenetic gene expression regulation. According to the *histone code hypothesis*, the pattern of histone modifications within each DNA sequence is a major factor in gene expression regulation. Covalent modification of histone N-terminal tails (**Figure 17B**) can occur at specific amino acid residues because the unstructured tails protrude outward from the nucleosome where they are accessible to modifying enzymes (**Figure 17Ca**). The most commonly observed modifications are methylation, acetylation and ubiquitination of lysine, methylation of arginine, and phosphorylation of serine. Histone modifications are designated by histone type followed by a one-letter symbol of the modified amino acid (**Table 5.1**, p. 138) and an abbreviation of the modification type. For example, mono- and dimethyl modifications of lysine 4 on histone 3 are referred to as H3K4me and H3K4me2, respectively. Enzymes that acetylate or methylate histone proteins can be considered as code *writers*; those that deacetylate or demethylate are *erasers*. Histone modifications, since they serve as platforms for the binding of specific accessory proteins, are *readers*. Once each accessory protein complex binds to modified histones, processes are initiated that alter chromatin structure such that transcription is inhibited or facilitated.

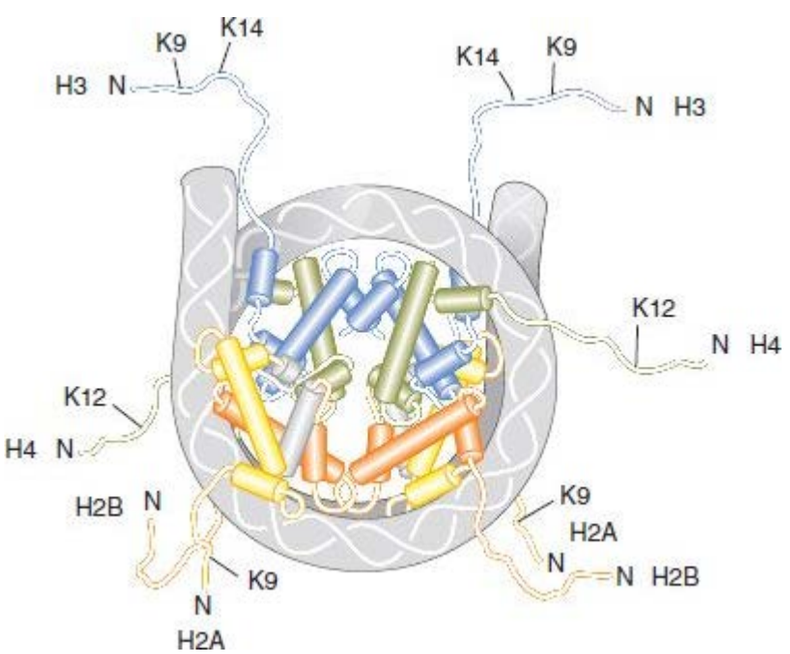


FIGURE 17B

Histone Modifications in a Nucleosome

The acetylation of the N-terminal tail lysine residues (K), for example, in combination with several other types of covalent modifications triggers chromatin remodeling that results in gene transcription.

Acetylation and methylation are the best characterized histone modifications. The acetylation of certain lysine residues in histone tails (e.g., H3K9ac) by *histone acetyltransferases* (HAT) (**Figure 17Cb**) promotes transcription by facilitating access of DNA to transcription factors. Note that acetylation of lysine residues neutralizes its positive charge, thereby reducing their electrostatic attraction to the negatively charged DNA. The DNA wrap around the histone octamer is loosened, and there are reduced interactions

between adjacent nucleosomes. Lysine acetylation, in combination with other histone modifications, also creates binding sites for proteins that initiate chromatin remodeling. Lysine deacetylation, catalyzed by *histone deacetylases* (HDAC) (**Figure 17Cc**), impedes transcription because deacetylated histones have an increased affinity for DNA, a circumstance that results in tighter chromatin coiling. Most, but not all, histone methylation reactions repress transcription. Depending on the specific pattern of histone methylations, transcription can be inhibited (e.g., H3K9me3 and H3K27me2) or promoted (e.g., H3K4me3 and H3K27me). Certain long noncoding (lnc) RNAs (p. 675) can increase or decrease transcription of a gene by facilitating the binding of chromatin-modifying proteins (**Figure 17Cd**). MeCPs mediate gene silencing by preferentially binding to 5-MeCpGs and recruiting HDAC to the site along with histone methylases.

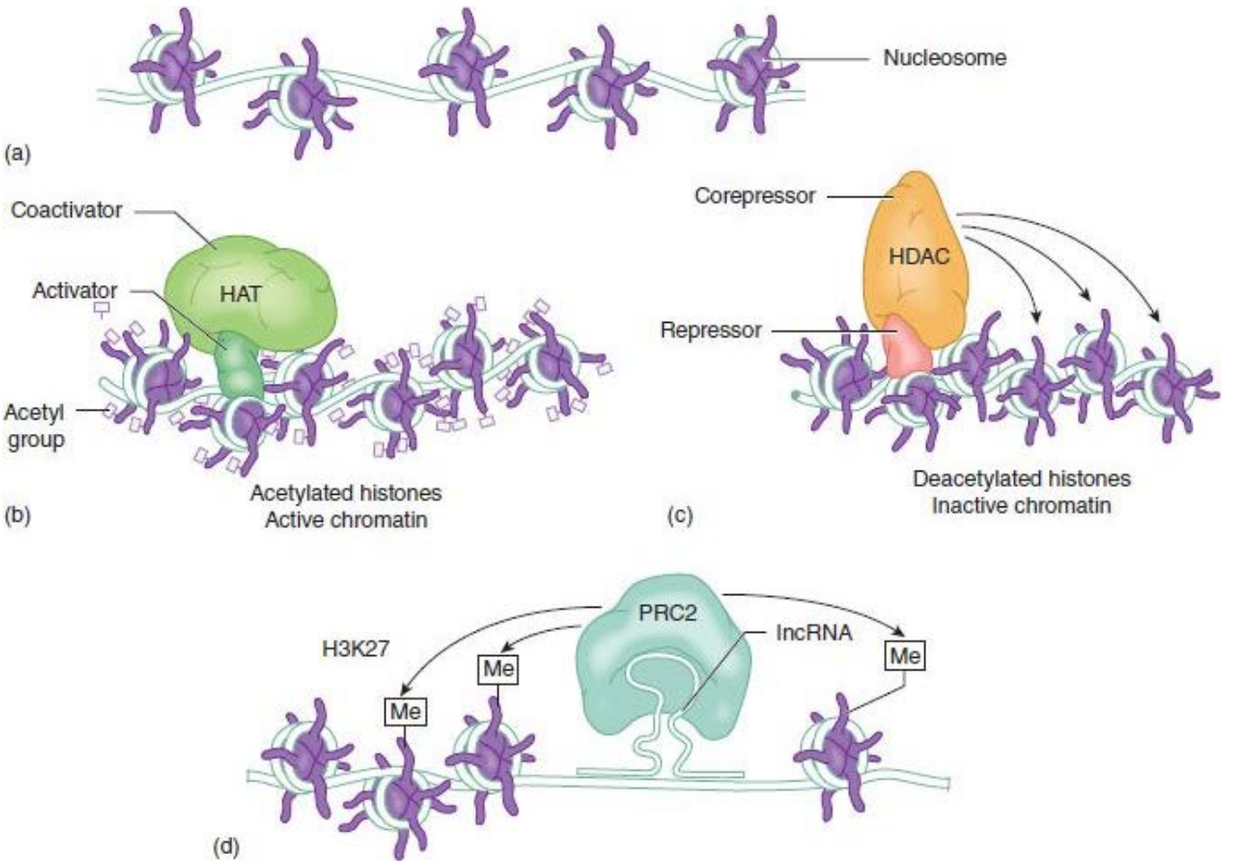


FIGURE 17C

Histone Modifications

Chemical modification of histones within nucleosomes (a) is an important chromatin regulatory process. (b) Acetylation by HAT of lysine residues in the N-terminal tails of histones facilitates chromatin activation. HATs are recruited by the binding of transcription factors (activators). (c) Deacetylation by HDAC, recruited by repressive transcription factors, promotes chromatin inactivation. (d) Numerous lncRNAs recruit chromatin-modifying proteins to specific DNA sequences. In this example, a lncRNA recruits polycomb repressive complex 2 (PRC2) to a DNA sequence. PRC has a methyltransferase activity that initiates chromatin inactivation by methylating H3 lysine residue 27 (H3K27me3).

Epigenetics: The Connection between Gene Expression and the Environment

Life scientists have long suspected that environmental factors can influence heredity. For example, monozygotic (identical) twins often have different susceptibilities to diseases such as diabetes or cancer. In 2000, a genetic experiment involving agouti mice unambiguously linked inheritance to the environment. Agouti mice are yellow, obese, and prone to diabetes and cancer. A seemingly subtle change in diet was observed to cause a profound alteration in the health and appearance of agouti offspring.

Agouti mice are obese animals with a yellow coat color because they overexpress the agouti gene that codes for AgRP (p. 624). AgRP is an antagonist of melanocortin receptors (MCRs). MCRs are G-protein-coupled receptors that bind α-MSH, a peptide hormone that in some cells stimulates the synthesis of melanin pigment. Yellow coat color results from AgRP inhibition of the hair follicle MCR. Obesity in agouti mice is caused by MCR inhibition in the hypothalamic appetite control center (p. 623). When pregnant yellow, obese agouti mice were fed a diet supplemented with methyl donors such as folic acid, methionine, and choline, their offspring were slender with a dark coat color (**Figure 17D**). Subsequent research revealed that agouti mice have a spontaneous insertion of a virus-like genetic element near the AgRP gene sequence. Hypomethylation of the inserted genetic element results in continuous synthesis of AgRP. Diet-induced methylation of the genetic element results in vastly reduced AgRP synthesis.

FIGURE 17D

Agouti Mice and Epigenetics

Both mice in this photograph have identical DNA sequences. The mouse on the left has yellow color and obesity, both typical agouti traits. The mouse on the right is the offspring of a yellow, obese mother fed a diet enriched in methyl donors such as folic acid and methionine.

Over the past decade, research efforts have revealed numerous examples of abnormal epigenome changes, referred to as **epimutations**, linked to environmental factors such as diet, toxins, pathogens, and behavior. Folate-deficient diets, which cause global and gene-specific DNA hypomethylation, have been observed to increase risk for cardiovascular disease and several forms of cancer (e.g., colorectal cancer). Toxins such as those in cigarette smoke cause epimutations that result in several cancers. Stomach cancer, induced by hypermethylation of a tumor suppressor gene, is caused by infection by the bacterium

Heliobacter pylori.

SUMMARY Heritable covalent modifications of cytosine bases and histone tails contribute to a sophisticated mechanism that regulates gene expression by determining the accessibility of DNA sequences to transcription machinery.

3. In contrast to the double helix of DNA, RNA exists as a single strand. For this reason, RNA can coil back on itself and form unique and often quite complex three-dimensional structures (**Figure 17.24**). The shape of these structures is determined by complementary base pairing by specific RNA sequences, as well as by base stacking and interactions between double-stranded regions (formed from single-stranded sequences of the same molecule or between single-stranded sequences of neighboring molecules) and free loops of RNA. Base-pairing rules apply in double-stranded regions where A-U, G-U, and G-C pairing occurs. Loop or single-stranded RNA (ssRNA) sequences contain a number of modified bases in mature non-mRNA molecules. The base composition of RNA does not follow Chargaff's rules because RNA molecules are single-stranded.
4. RNA molecules have catalytic properties because of the complex three-dimensional structures with binding pockets that they can form. The majority of catalytic RNA molecules, referred to as **ribozymes**, catalyze self-cleavages or the cleavage of other RNAs. The most notable example of ribozyme activity, however, is peptide bond formation within ribosomes. A magnesium ion cofactor is usually required in RNA catalysis because Mg^{2+} stabilizes transition states.

The most prominent types of RNA directly involved in protein synthesis are transfer RNA, ribosomal RNA, and messenger RNA. The structure and function of each of these molecules are discussed next. Examples of several other types of ncRNAs, are also provided.

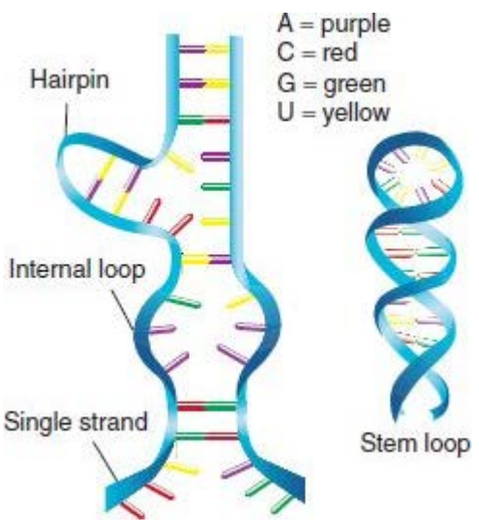


FIGURE 17.24

Secondary Structure of RNA

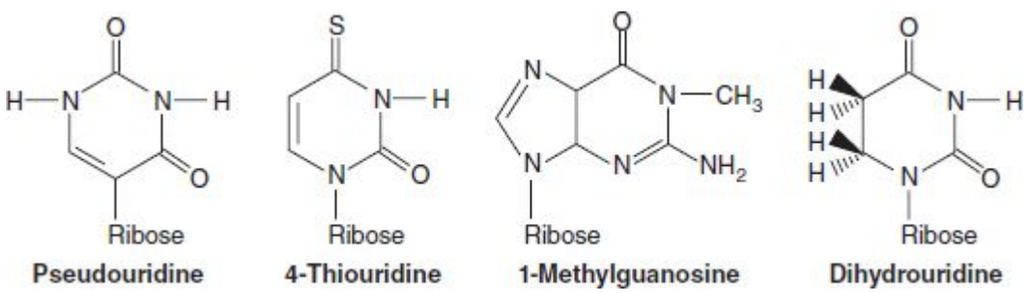
Three of the many different types of secondary structure that occur in RNA molecules are illustrated. Hairpin

loops form when there are at least four bases that form base pairs. Internal loops form when both sides of a double-stranded segment cannot form base pairs. Stable stem loops are between four and eight bases long. Note that RNA hairpins and stem loops form because of the inverted repeat sequences of DNA palindromes (see p. 648).

Transfer RNA

Transfer RNA (tRNA) molecules transport amino acids to ribosomes for assembly into proteins. The length of tRNA molecules varies from 75 to more than 90 nucleotides. Each type of tRNA molecule binds to a specific amino acid. Consequently, cells possess at least one type of tRNA for each of the 20 standard amino acids. The three-dimensional structure of tRNA molecules, which resembles a warped cloverleaf (**Figure 17.25**), results primarily from extensive intrachain base pairing.

Eukaryotic tRNAs typically have at least 50 different modifications that serve to stabilize their structures and facilitate function. Examples of tRNA modified nucleotide bases include pseudouridine, 4-thiouridine, 1-methylguanosine, and dihydrouridine:



The structure of tRNA allows it to perform two critical functions involving the anticodon loop and the 3′-terminus. The *anticodon loop* contains a three-base sequence (the anticodon) that is complementary to the mRNA triplet code for the specific amino acid. The 3′-terminus forms a covalent bond to a specific amino acid. (This specificity is achieved because the set of enzymes called the *aminoacyl-tRNA synthetases* link each amino acid to its cognate tRNA.) The conformational relationship between the 3′-terminus and the anticodon loop allows the tRNA to align its attached amino acid within the ribosome's active site during protein synthesis.

tRNAs also possess three other prominent structural features, referred to as the D loop, the TψC loop, and the variable loop. (The Greek letter psi, ψ, stands for the modified base pseudouridine.) These structures facilitate specific binding of a tRNA to the appropriate aminoacyl-tRNA synthetase and the appropriate alignment of the aminoacyl-tRNA within the nucleoprotein scaffold of the ribosome. The *D loop* is so named because it contains dihydrouridine. Similarly, the TψC loop contains the base sequence thymine, pseudouridine, and cytosine. tRNAs can be classified on the basis of the length of their *variable loop*. The majority (approximately 80%) of tRNAs have variable loops with four to five nucleotides, whereas the others have variable loops with as many as 20 nucleotides.

Biochemistry IN THE LAB

Nucleic Acid Methods

The techniques used in the isolation, purification, and characterization of biomolecules take advantage of their physical and chemical properties. Most of the techniques used in nucleic acid research are based on differences in molecular weight or shape, base sequences, or complementary base pairing. Techniques such as chromatography, electrophoresis, and ultracentrifugation, which have been used successfully in protein research, have also been adapted to use with nucleic acids. In addition, other techniques have been developed that exploit the unique properties of nucleic acids. For example, under certain conditions, DNA duplexes reversibly melt (separate) and reanneal (base-pair to form a duplex again). One of several techniques that exploit this phenomenon, called *Southern blotting*, is often used to locate specific (and often rare) nucleic acid sequences. After brief descriptions of several techniques used to purify and characterize nucleic acids, the common method for determining DNA sequences is outlined. More complex techniques are described in Chapter 18 in Biochemistry in the Lab: Genomics (p. 712).

Techniques Adapted from Use with Other Biomolecules

Many of the techniques used in protein purification procedures have also been adapted for use with nucleic acids. For example, several types of chromatography (e.g., ion-exchange, gel filtration, and affinity) have been used in several stages of nucleic acid purification and in the isolation of individual nucleic acid sequences. Because of its speed, HPLC (high-performance liquid chromatography, p. 186) has replaced many slower chromatographic separation techniques applied to small samples.

The movement of nucleic acid molecules in an electric field depends on both their molecular weight and their three-dimensional structure. However, because DNA molecules often have relatively high molecular weights, their capacity to penetrate some gel preparations (e.g., polyacrylamide) is limited. Although DNA sequences with less than 500 bp can be separated by polyacrylamide gels with especially large pore sizes, more porous gels must be used with larger DNA molecules. Agarose gels, which are composed of a cross-linked polysaccharide, are used to separate DNA molecules with lengths between 500 bp and approximately 150 kilobases (kb). Larger sequences are now isolated by pulsed-field electrophoresis (PFGE), a variation of agarose gel electrophoresis in which two electric fields (perpendicular to each other) are alternately turned on and off. DNA molecules reorient themselves each time the electric field alternates, resulting in a very efficient and precise separation of heterogeneous groups of DNA molecules. The reorientation times decrease with the size of the DNA.

Density gradient centrifugation with cesium chloride (CsCl) has been widely used in nucleic acid research (see Biochemistry in the Lab: Cell Technology, p. 68, in Chapter 2). At high speeds, a linear gradient of CsCl is established. Mixtures of DNA, RNA, and protein migrating through this gradient separate into discrete bands at positions where their densities are equal to the density of the CsCl. DNA molecules with high guanine and cytosine content are more dense than those with a higher proportion of adenine and thymine. This difference helps separate heterogeneous mixtures of DNA fragments.

Techniques That Exploit the Unique Structural Features of the Nucleic Acids

Several unique properties of the nucleic acids are exploited in nucleic acid research. These properties include absorption of UV light at specific wavelengths and the tendency of nucleic acids

to reversibly form double-stranded complexes.

Because of their aromatic structures, the purine and pyrimidine bases absorb UV light. At pH 7 this absorption is especially strong at 260 nm. However, when the nitrogenous bases in single-stranded polynucleotides are incorporated into the base pairs of double-stranded polynucleotides, various noncovalent forces promote interactions between them that decrease their absorption of UV light. This **hypochromic effect** is an invaluable aid in studies involving nucleic acid. For example, absorption changes are routinely used to detect the disruption of the double-stranded structure of DNA or the hydrolytic cleavage of polynucleotide strands by enzymes.

The binding forces that hold the complementary strands of DNA together can be disrupted. This process, referred to as **denaturation** (**Figure 17E**), is promoted by heat, low-salt concentrations, and extremes in pH. (Because it is easily controlled, heating is the most common denaturing method in nucleic acid investigations.) When a DNA solution is slowly heated, absorption at 260 nm remains constant until a threshold temperature is reached. Then the sample's absorbance increases. The absorbance change is caused by the unstacking of bases and the disruption of base pairing. The temperature at which one-half of a DNA sample is denatured, referred to as the melting temperature (T_m), varies among DNA molecules according to their base compositions. (Recall that DNA stability is affected by the number of hydrogen bonds between GC and AT pairs and base stacking interactions [see p. 640]. More energy is required, therefore, to "melt" DNA molecules with high G and C content.) If the separated DNA strands are held at a temperature approximately 25ºC below the T_m for an extended time, renaturation is possible. Renaturation, or reannealing, does not occur instantaneously because the strands explore various configurations until they achieve the most stable one (i.e., the one having paired complementary regions).

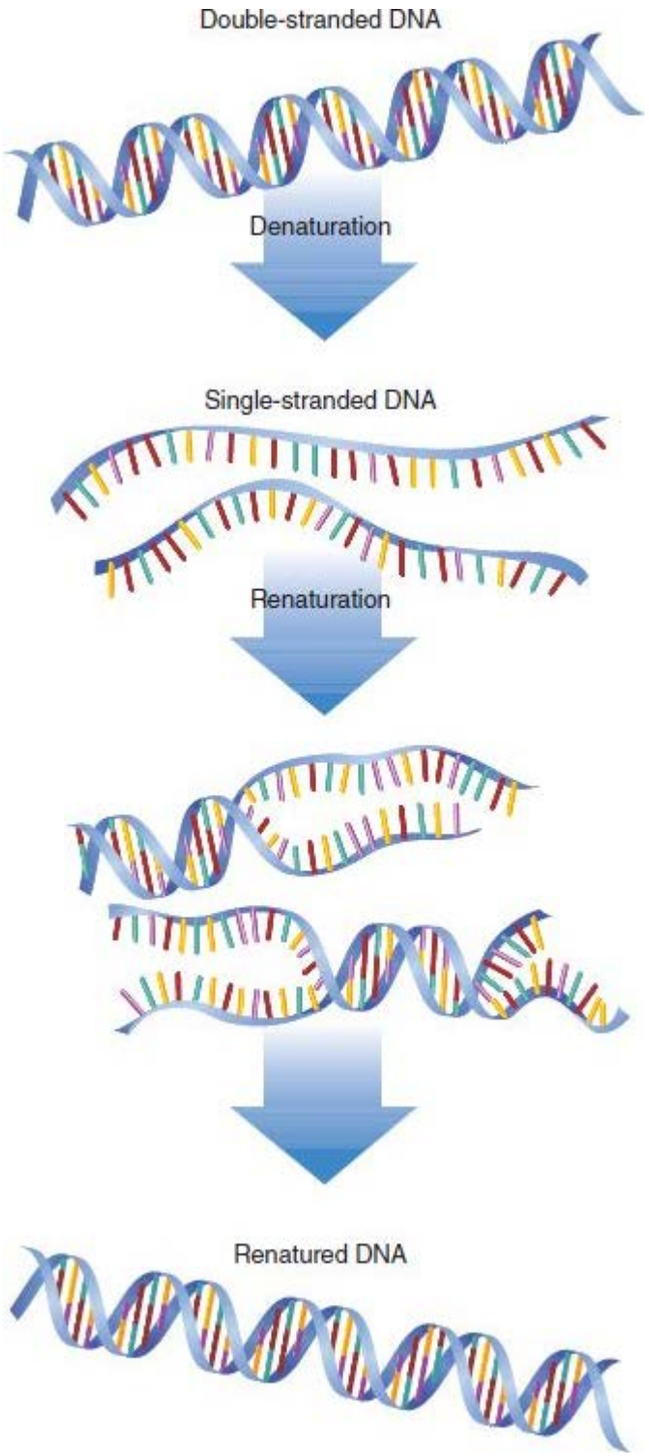


FIGURE 17E

Denaturation and Renaturation of DNA

Under appropriate conditions, DNA that has been denatured can renature; that is, strands with complementary sequences will re-form into a double helix.

Hybridization, a phenomenon in which single-stranded DNA or RNA anneal to complementary DNA or RNA single strands, can also be used to locate and/or identify specific genes or other DNA sequences. For example, ssDNA from two different sources (e.g., tumor cells and normal cells) can be screened for sequence differences. If one set of ssDNA is biotinylated, then the double-stranded hybrids bind to an avidin column. (Avidin is a protein that binds with high affinity to biotin.) If any unhybridized sequence is present, it passes through the column. Then it can be

isolated and identified.

In **Southern blotting** (**Figure 17F**), radioactively labeled DNA or RNA probes (sequences with known identities) locate a complementary sequence in the midst of a DNA digest, which typically contains a large number of heterogeneous DNA fragments. A DNA digest is obtained by treating a DNA sample with restriction enzymes that cut at specific nucleotide sequences (**Figure 17G**). (Produced by bacterial cells, restriction enzymes protect bacteria against viral infection by cleaving viral DNA at specific sequences.) Once the DNA sample has been digested, the fragments are separated by agarose gel electrophoresis according to their sizes. After the gel has been soaked in 0.5 M NaOH, a process that converts dsDNA to ssDNA, the DNA fragments are transferred to nitrocellulose filter paper by placing them on a wet sponge in a tray with a high-salt buffer. (Nitrocellulose has the unique property of binding strongly to ssDNA.) Absorbent dry filter paper is placed in direct contact with the nitrocelluose filter/agarose gel sandwich. As buffer is drawn through the gel and filter paper by capillary action, the DNA is transferred and becomes permanently bound to the nitrocellulose filter. (The transfer of DNA to the filter is the "blotting" referred to in the name of this technique.) Subsequently, the nitrocellulose filter is exposed to the radioactively labeled probe, which binds to any ssDNA with a complementary sequence. For example, an mRNA that codes for β-globin binds specifically to the β-globin gene, even though β-globin mRNA lacks the introns present in the gene. Apparently, there is sufficient base pairing between the two single-stranded molecules to locate the gene.

Detection techniques other than autoradiography are also used. Nucleic acid probes can be labeled with ethidium bromide or luminol, for example. The intercalating agent ethidium bromide (p. 645) fluoresces when the gel is exposed to UV light. Probes covalently linked to luminol, a chemiluminescent molecule, emit a blue light when luminol is exposed to an oxidizing agent (H_2O_2) and a catalyst (iron atoms).

DNA Sequencing

The determination of DNA nucleotide sequences has provided valuable insights in biochemistry, medical science, and evolutionary biology. The analysis of long DNA sequences may be accomplished with the use of multiple primers, or it may begin with the formation of smaller fragments by means of one type of restriction enzyme. (A primer is a short single-stranded nucleic acid sequence that provides a starting point for the synthesis of a complementary DNA strand.) Each fragment is then sequenced independently by the chain-terminating method. As with protein primary structure determinations, these steps are repeated with a different set of polynucleotide fragments (generated by another type of restriction enzyme) that overlap the first set. Sequence information from both sets of experiments then orders the fragments into a complete sequence.

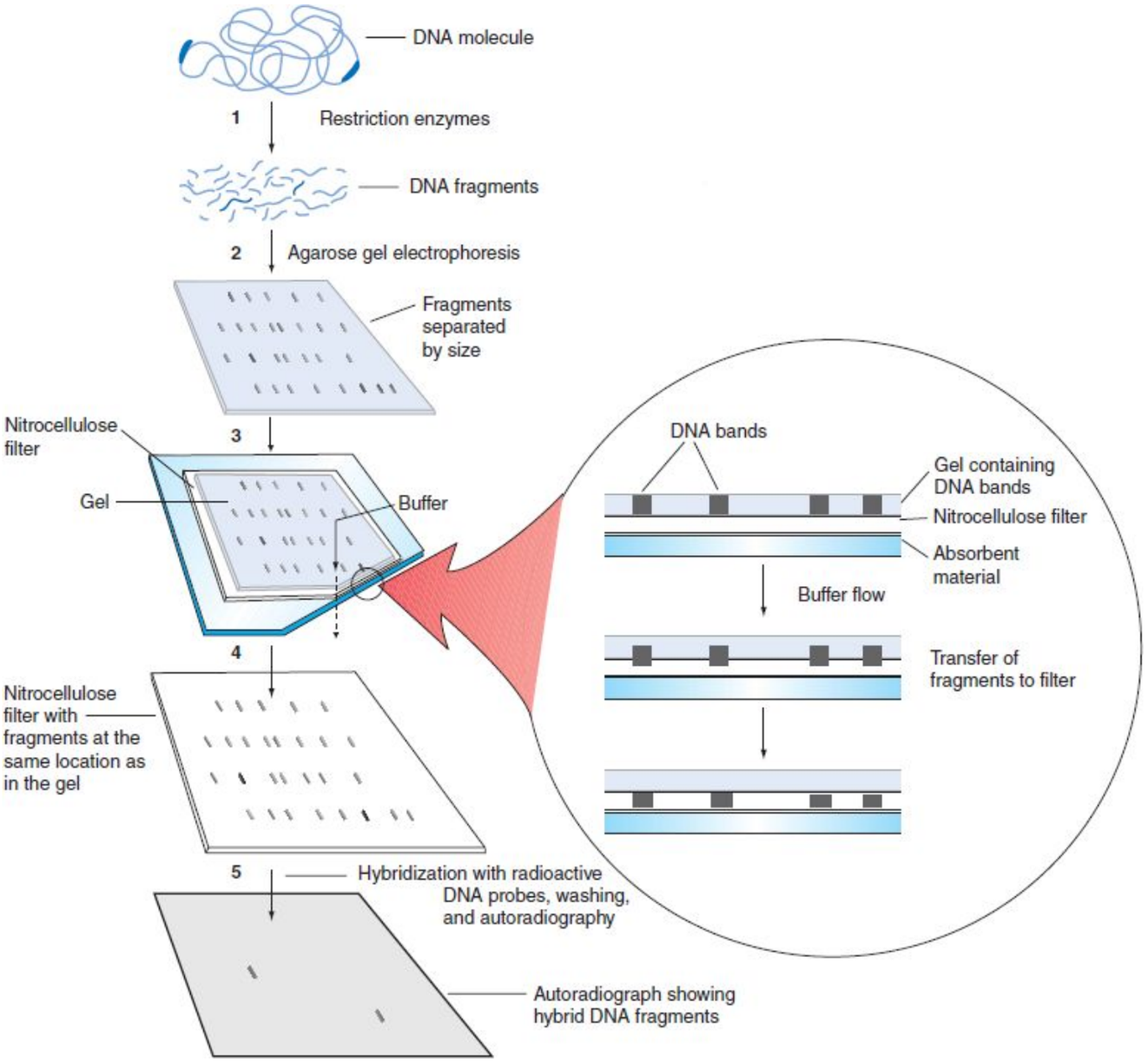


FIGURE 17F

Southern Blotting

(1) DNA analysis begins with its digestion by a restriction enzyme. (2) DNA fragments are separated by agarose gel electrophoresis. (3) The DNA fragments are transferred to nitrocellulose filter paper under denaturing conditions. (4) The ssDNA on the nitrocellulose filter paper is hybridized with a radioactively labeled ssDNA probe. (5) Any hybridized DNA can be visualized by autoradiography or other detection techniques. Southern blotting is used in diverse DNA investigations.

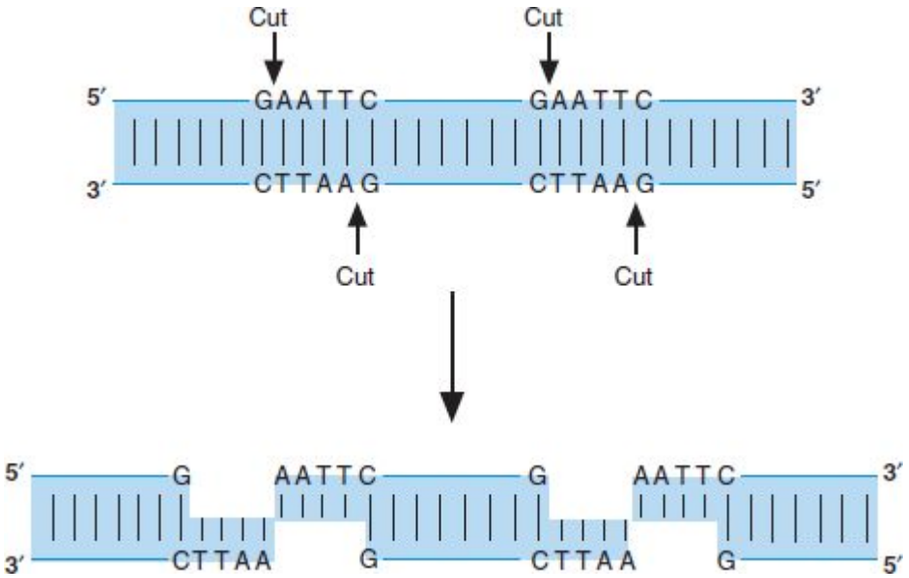


FIGURE 17G

Restriction Enzymes

Restriction endonucleases are enzymes isolated from bacteria that cut DNA at specific sequences. In this example, the enzyme EcoRI (obtained from *E. coli*) makes staggered cuts that result in the formation of "sticky ends." A "sticky end" is a single-stranded terminus on a double-stranded DNA fragment. Sticky ends facilitate the formation of recombinant DNA because a single-stranded segment on one DNA fragment can anneal with a complementary sticky end on another DNA fragment. Some restriction enzymes make what are called "blunt cuts." For a discussion of recombinant DNA, see Biochemistry in the Lab: Genomics, pp. 712–18 in Chapter 18.

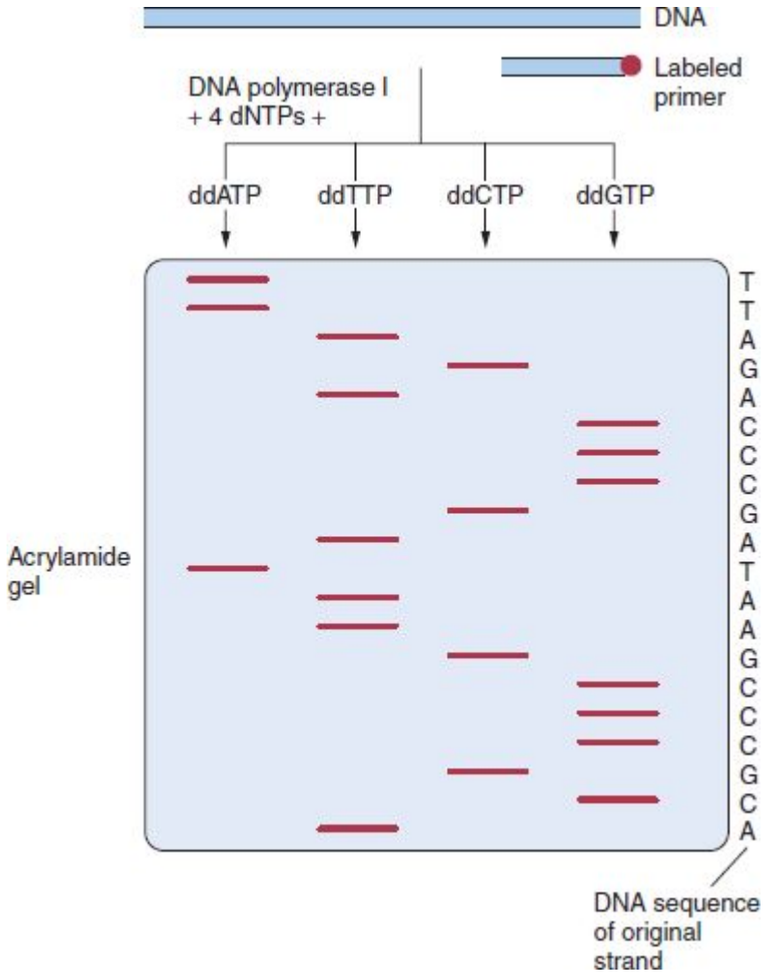


FIGURE 17H

The Sanger Chain Termination Method

A specific ^{32}P-labeled primer is chosen so that DNA synthesis will begin at the point of interest on the template strand to be replaced. DNA synthesis continues in the presence of a mixture of dNTPs and a trace of one of four dideoxynucleotides (ddNTPs). Chain termination occurs when the ddNTP is incorporated into the growing DNA chain. The mixture of truncated fragments is separated by gel electrophoresis and analyzed by autoradiography. Four samples are run in separate lanes, each with a different ddNTP. The DNA sequence is "read" from the bottom to the top of the gel.

In DNA sequencing by the **chain-terminating method** (**Figure 17H**), developed by Frederick Sanger, restriction enzymes cleave large DNA segments into smaller fragments. Each fragment is separated into two strands, one of which is used as a template to produce a complementary copy. The sample is further divided into four test tubes. To each tube is added the substances required for DNA synthesis (e.g., the enzyme DNA polymerase and the four deoxyribonucleotide triphosphates and a ^{32}P-labeled primer). The investigator determines the site at which sequencing is to start by selecting the appropriate primer, usually beginning with the recognition sequence of the restriction enzyme used to generate the DNA fragment of interest.

In addition to a template, DNA synthesis materials, and a primer, each of the four tubes contains a different 2′, 3′-dideoxynucleotide derivative. (The dideoxy derivatives are synthetic nucleotide analogues in which the hydroxy groups on the 2′- and 3′-carbons have been replaced with hydrogens.) Dideoxynucleotides can be incorporated into a growing polynucleotide chain, but they cannot form a phosphodiester linkage with another nucleotide. Consequently, when dideoxynucleotides are incorporated, they terminate the chain. Because small amounts of the

dideoxynucleotides are used, they are randomly incorporated into growing polynucleotide strands. Each tube, therefore, contains a mixture of DNA fragments containing strands of different lengths. Each newly synthesized strand ends in a dideoxynucleotide residue. The reaction products in each tube are separated by gel electrophoresis and analyzed together by autoradiography. Each band in the autoradiogram corresponds to a polynucleotide that differs in length by one nucleotide from the one that precedes it in any of the four lanes of the autoradiogram. Note that the smallest polynucleotide appears on the bottom of the gel because it moves more quickly than larger molecules.

Eventually, a more rapid and efficient automated version of the Sanger method that used fluorescent-tagged dideoxynucleotides was introduced (**Figure 17I**). However, in the mid-2000s, the demand for low-cost sequencing led to the development of a series of next-generation high-throughput methods such as Illumina sequencing and ion torrent sequencing. In Illumina sequencing, which uses a reversible dye terminator method, DNA molecules and primers are attached to a slide and then amplified by a polymerase to produce DNA clusters. First, four types of fluorescently labeled base (adenine, guanine, cytosine, and thymine) are introduced, each with a blocking group attached. Next, each DNA cluster is monitored by a laser-scanning confocal microscope, which acquires images point by point and then reconstructs them via computer. The four bases, each with a specific fluorescent color, compete with each other for binding sites on the template DNA molecules. During each sequencing cycle, a single nucleotide is added to the DNA chain. After nonincorporated bases are washed away, the fluorescent dye bound to the incorporated base is identified. The cycle ends with the removal of the terminal blocking group. The cycles are repeated until the DNA molecule is completely sequenced. With an estimated cost of $0.10 per million bases, Illumina sequencing is significantly cheaper than Sanger sequencing ($2400/million bases).

In ion torrent sequencing, DNA synthesis monitoring by computer is used in combination with a high-density array of micromachined wells in a semiconductor chip. The ion torrent chip is so designed that a hypersensitive ion-sensitive layer below the wells acts as a miniature pH meter. Each well contains many copies of a different DNA template and DNA polymerase. All four unmodified nucleotides are then sequentially introduced into the wells. A cycle begins with adding one type of nucleotide to the wells. If an introduced nucleotide is incorporated (i.e., there is a covalent bond formed), the resulting release of a hydrogen ion is detected by the ion sensor. In the absence of the incorporating reaction, there is no biochemical reaction or hydrogen ion release. Each cycle ends with the removal of unattached nucleotides. The cost per million bases for ion torrent sequencing is approximately $1.

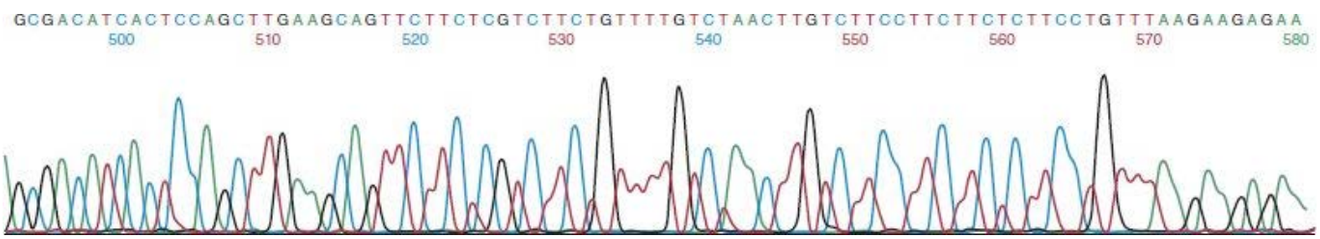


FIGURE 17I

Automated Sanger DNA Sequencing

With the use of fluorescent tags on the dideoxynucleotides, a detector can scan a gel quickly and determine the sequence from the order of the colors in the bands.

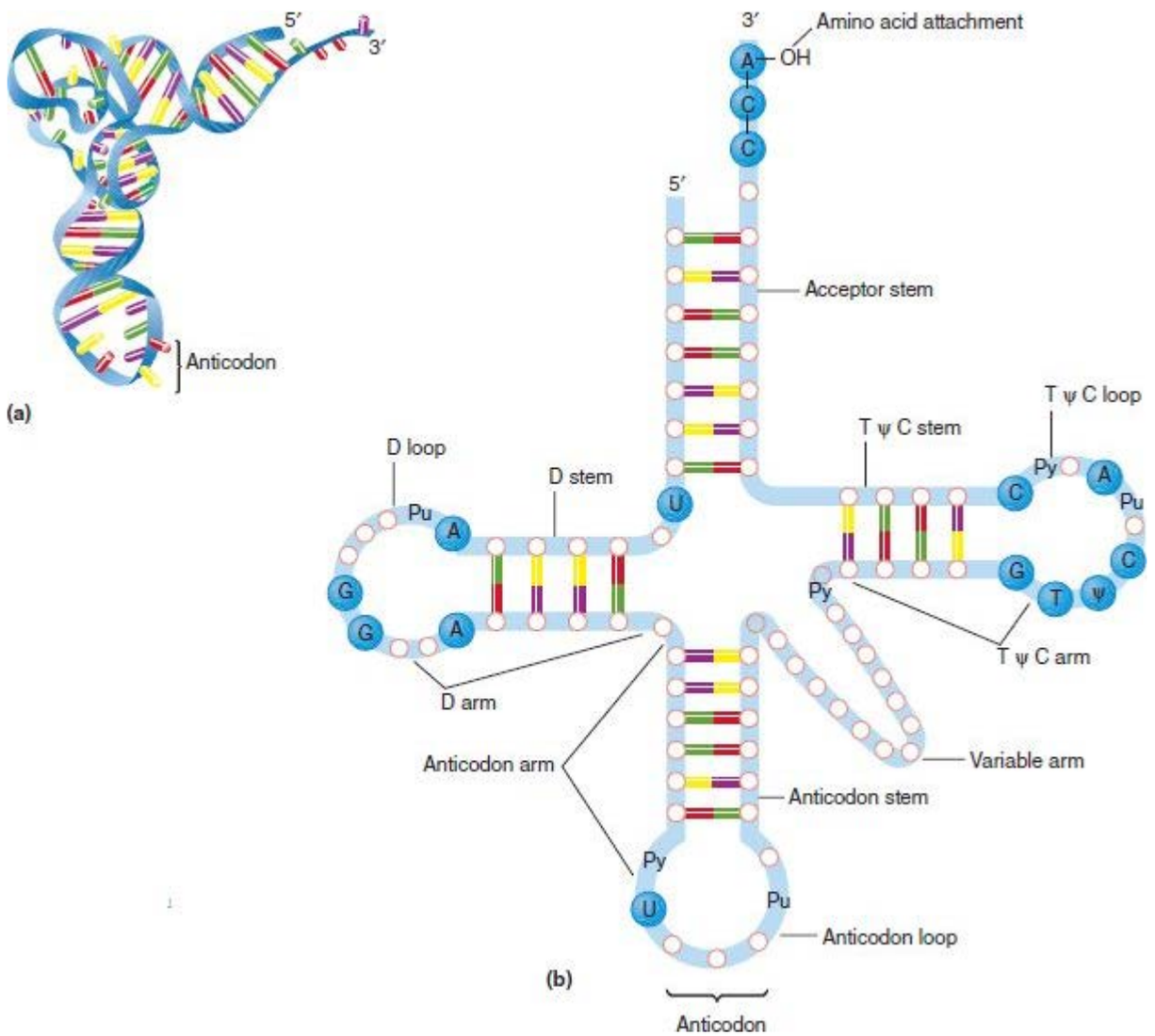


FIGURE 17.25

Transfer RNA

(a) Three-dimensional structure of a tRNA molecule. (b) Schematic view of a tRNA molecule. The three-base anticodon sequence within the anticodon arm forms complementary base pairs with a codon triplet in mRNA. The 3’ CCA sequence of a tRNA is the site that will be covalently bonded to its cognate amino acid by an aminoacyl-tRNA synthetase. The positions of invariant bases and bases that seldom vary are indicated.

Biochemistry IN PERSPECTIVE

Forensic Investigations

How is DNA analysis used in the investigation of violent crime? DNA persists for many years in dried biological specimens (e.g., blood, saliva, hair, and semen) and in bone.

Consequently, DNA can be used as evidence in any type of forensic investigation in which such specimens are available. DNA analysis techniques that are typically used to ascertain the identity of victims and/or perpetrators of violent crimes are referred to as **DNA typing**, or profiling. DNA typing involves the analysis of variable sequences called markers. Markers include tandem repeat variations (p. 659) and single nucleotide polymorphisms (SNPs). SNPs are single nucleotide variations or point mutations that occur in at least 1% of the human population. With several million SNPs identified in both coding and noncoding DNA, they represent the majority of genetic variation among humans. Although 99.9% of the DNA humans share is identical, variations in the remaining 0.1% allow investigators to generate identifying genetic profiles for each individual human.

In numerous court cases since the 1990s, DNA typing has provided decisive information concerning defendants' presence at a crime scene or their absence. **DNA fingerprinting**, introduced in 1985 by the British geneticist Alec Jeffreys, is a variation of Southern blotting. In this technique, referred to as restriction fragment polymorphisms, or RFLPs, the banding characteristics of DNA minisatellites (p. 659) from different individuals are compared—for example, crime scene specimen DNA with samples from suspects. When the quantity of DNA extracted from a crime scene sample is too minute to analyze, it is amplified by means of the *polymerase chain reaction* (PCR), a technique that is used to amplify the number of copies of DNA in a tiny sample. (Refer to p. 714.) Up to 10^9 copies can be obtained. Consequently, DNA from a single cell is now sufficient for DNA fingerprint analysis. The entire genome in each sample is isolated and treated with a restriction enzyme (see p. 668).

Although RFLP testing is an accurate method, it does have limitations. Among these are the substantial amounts of time (6–8 weeks), labor, and expertise required to obtain DNA profiles. A newer methodology that analyzes **short tandem repeats (STRs)**, (DNA sequences with between 2 and 4 bp repeats, called microsatellites; see p. 659), has significantly greater discriminating power than RFLP and is relatively rapid (several hours).

Moreover, STR sequences are sufficiently robust that STR analysis can often be successfully used to analyze degraded specimens. After DNA has been extracted from a specimen, several target STR sequences are amplified by PCR and linked to fluorescent dye molecules.

In the United States, 13 core polymorphic (highly variable) DNA markers, called *loci*, are used to generate genetic profiles and to distinguish between individuals. A **DNA profile** (**Figure 17J**), which results when the PCR products are separated in an electrophoretic gel, consists of the pattern and number of repeats of each target sequence visible on the gel. Fluorescence detection increases the sensitivity of the technique. Unlike RFLP, STR-based DNA typing is easily automated. If the DNA profiles from individual samples are compared and determined to be identical, the samples are said to be a match. If compared profiles are not identical, they are said to have come from different sources. The results are reported in terms of the probabilities of a random match (the chance that a randomly selected person from the population will have a DNA profile identical to that of the specimen of interest such as that left at a crime scene). The use of multiple markers and the sensitivity of the methodology reduce the random match probability to at least one in several billion.

When DNA is too degraded for nuclear DNA STR analysis, mitochondrial DNA (mtDNA) and SNPs can often yield results. Both types of DNA analysis are used as a last resort because they are expensive and labor intensive. MtDNA analysis, which involves sequencing of the entire mtDNA, can be successful when nuclear DNA analysis fails because most cells have hundreds to thousands of mitochondria. SNP markers are small enough that they can often be identified in degraded samples, but their discriminating power is not as high as that used in STR analysis. In contrast to

the 13 loci used in STR analysis, SNP analysis requires at least 50 loci to discriminate between individuals.

FIGURE 17J

A DNA STR Profile

STR DNA analysis is often used to analyze evidence collected at a crime scene.

SUMMARY Forensic scientists use PCR and other technologies to amplify crime-scene DNA to generate the unique genetic profile that distinguishes one individual from all others.

Transfer RNAs have functions other than their well-known role in protein synthesis. tRNA fragmentation is the best researched example. In mammalian cells under stress (nutrient deficit, hypoxia, oxidative stress, or heat shock), ribonucleases such as angiogenin cleave tRNAs near the anticodon to yield 5′- and 3′-tRNA fragments (tiRNAs). 5′-tiRNA^{cys} and 5’-tiRNA^{ala} induce the formation of stress granules. Stress granules are aggregates of RNA and proteins such as failed translation initiation complexes and a diverse array of other proteins that include molecules that stabilize mRNAs, molecular chaperones, and small ribosomal subunits (p. 768). During the recovery from stress, ATP-dependent remodeling complexes disassemble stress granules, thereby liberating preassembled translation complexes that proceed with polypeptide synthesis. tiRNAs also protect cells from programmed cell death under stressful conditions. Both 5′- and 3′-tiRNAs are capable of binding to cytochrome c, the molecule released from damaged mitochondria that initiates apoptosis.

Ribosomal RNA

Ribosomal RNA (rRNA), the most abundant form of RNA in living cells, is a component of ribosomes. Since rRNA is present in all living organisms, its sequences have been used to

characterize the evolutionary relationships between organisms. rRNA has an extraordinarily complex structure (**Figure 17.26**). Although there are species differences in the primary nucleotide sequences of rRNA, the overall three-dimensional structure of this class of molecules is conserved.

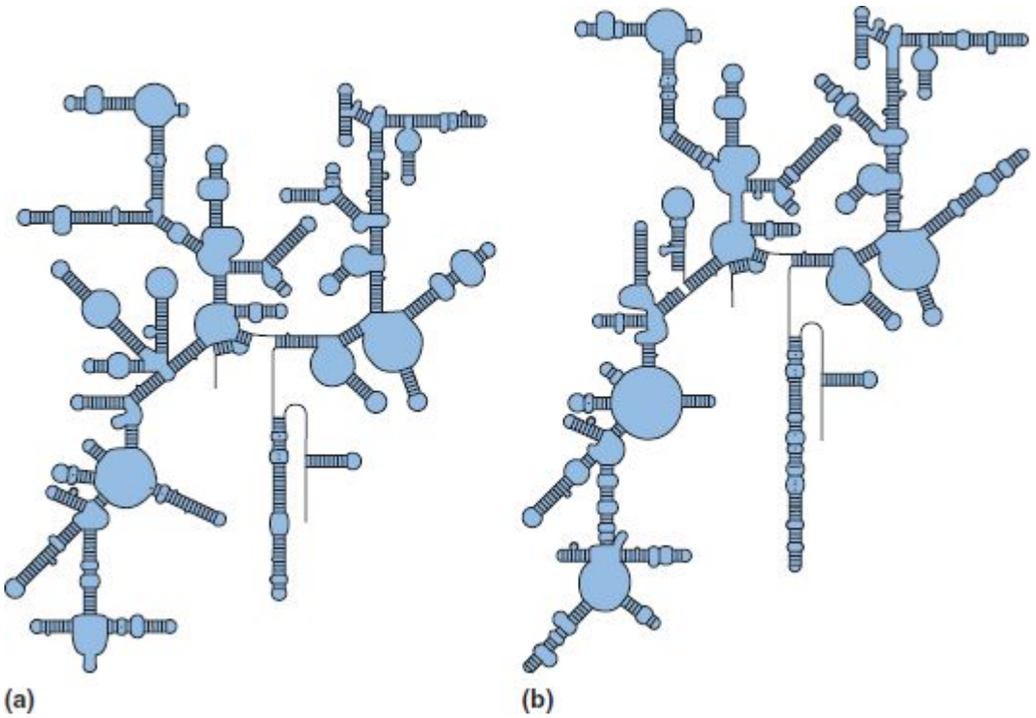


FIGURE 17.26

rRNA Structure

Although their sequences differ, the three-dimensional structure of these rRNAs from (a) *E. coli* and (b) *S. cerevisiae* (yeast) appears remarkably similar.

Ribosomes are the cytoplasmic ribonucleoprotein complexes that synthesize proteins. There are as many as 50,000 ribosomes in bacterial cells such as those of *E. coli*. Eukaryotic cells may possess up to 10 million ribosomes. Prokaryotic and eukaryotic ribosomes are similar in shape and function, although they differ in size and in their chemical composition. Both types of ribosome consist of two subunits of unequal size, which are usually referred to in terms of their S values. (The Svedberg unit S is a measure of sedimentation velocity in a centrifuge. Because sedimentation velocity depends on both the molecular weight and the shape of a particle, S values are not necessarily additive.) Prokaryotic ribosomes (70S) are composed of a 50S subunit and a 30S subunit, whereas ribosomes of eukaryotes (80S) contain a 60S subunit and a 40S subunit.

Several different kinds of rRNA and protein are found in each type of ribosomal subunit. The large ribosomal subunit of *E. coli*, for example, contains 5S and 23S rRNAs and 34 polypeptides. The small ribosomal subunit of *E. coli* contains a 16S rRNA and 21 polypeptides. The bacterial rRNA genes are organized in an operon in the following order: 3′-16S, 23S, 5S-5′. *E. coli* has seven copies in its genome. A typical large eukaryotic ribosomal subunit contains three rRNAs (5S, 5.8S, and 28S) and 49 polypeptides; the small subunit contains an 18S rRNA and approximately 30 polypeptides. Eukaryotic rRNA gene copy numbers vary from 50 to more than 5000. Humans are estimated to have about 350 copies present in clusters on chromosomes 13, 14, 15, 21, and 22. The rRNA serves as a scaffold for the self-assembly of proteins to form the native ribosomal subunit. Peptidyl transferase, the enzymatic activity responsible for peptide bond formation, resides in the 23S rRNA in prokaryotic ribosomes and 28S rRNA in eukaryotic

ribosomes.

Messenger RNA

As its name suggests, **messenger RNA** (mRNA) is the carrier of genetic information from DNA for the synthesis of polypeptides. Messenger RNA molecules contain three-base sequences, called *codons*, which dictate specific amino acids in the subsequently synthesized protein. The base sequence within an mRNA that codes for a polypeptide is called an **open reading frame** (**ORF**). An ORF begins with an initiation codon and ends with a termination or stop codon. ORFs are flanked by untranslated sequence regions (UTRs). The untranslated region upstream of the ORF is referred to as the 5′UTR. The 3′UTR follows the ORF. As a result of varying lengths of both ORFs and UTRs, mRNAs can vary widely in length. For example, mRNA length in *E. coli* varies from 500 to 6000 nt. The average mRNA lengths in budding yeast (*S. cerevisiae*) and mammals are 1250 nt and 2200 nt, respectively.

Prokaryotic mRNA and eukaryotic mRNA differ in several respects. First, many prokaryotic mRNAs are *polycistronic*; that is, they contain coding information for several polypeptide chains. In contrast, eukaryotic mRNA typically codes for a single polypeptide and is therefore referred to as *monocistronic*. (A **cistron** is a DNA sequence that contains polypeptide coding information and several signals that are required for ribosome function.) Second, prokaryotic and eukaryotic mRNAs are processed differently. In contrast to prokaryotic mRNAs, which are translated into protein by ribosomes during or immediately after they are synthesized, eukaryotic mRNAs are modified extensively before translation. These modifications include capping (linkage of 7-methylguanosine to the 5′-terminal residue), splicing (removal of introns), and the 3’ attachment of an adenylate polymer referred to as a poly(A) tail. (Each of these processes is described in Chapter 18.)

Noncoding RNA

RNAs that are not directly involved in polypeptide synthesis are called **noncoding RNAs** (**ncRNAs**). Once believed to be transcriptional noise, these molecules have roles in numerous cell processes that include DNA replication, gene expression, transcription, translation, stress management, genome structure and defense, and epigenetic regulation. Although both prokaryotes and eukaryotes possess ncRNAs, only eukaryotic molecules are described in the following discussion.

Noncoding RNAs are classified according to both length and function. Small ncRNAs (sncRNAs) have lengths less than 200 nt. Long noncoding RNAs (lncRNAs) have lengths greater than 200 nt. There are two classifications of ncRNA function: housekeeping and regulatory. Housekeeping ncRNAs are constitutively expressed; in other words, they perform basic cellular functions such as RNA splicing (snRNAs), or chemical modifications of nucleotide bases (snoRNAs). Regulatory RNAs, as their name suggests, have roles in the activation or inhibition of gene expression. Regulatory ncRNAs can be further classified according to length. Short regulatory ncRNAs include microRNAs (miRNAs) and short interfering RNAs (siRNAs), which play important roles in *RNA interference* (p. 738), and Piwi-interacting RNAs (piRNAs), which are involved in transposon silencing. LncRNAs have diverse regulatory functions. Note that ncRNAs often perform their functions as components of ribonucleoprotein complexes. Each of these ncRNA classes is briefly described next.

The primary function of the **small nuclear RNAs** (snRNAs) is the processing of pre-mRNA. Composed of an average of 150 nt, the snRNAs U1, U2, U4, U5, and U6 combine with several

proteins to form **small nuclear ribonucleoproteins (snRNPs)**, often called “snurps.” Together with several other proteins, the snRNPs form a molecular machine called the spliceosome because of its function: splicing is a key step in the processing of eukaryotic mRNAs. **Spliceosomes** excise introns from pre-mRNA and then join exons together. Other snRNAs have been shown to regulate the activity of transcription factors and of RNA polymerase II (p. 725).

Small nucleolar RNAs (**snoRNAs**) are single-stranded RNAs containing 70 to 300 nucleotides; they facilitate chemical modifications of rRNA, tRNA, and snRNA within the nucleolus. Encoded within the introns of rRNA genes, snoRNAs are a component of the small nucleolar ribonucleoprotein (snoRNP). The function of snoRNAs (several hundred in humans) is to guide the snoRNP via base pairing to the specific sequence site on a target rRNA. Modifications that occur during rRNA processing include methylation of the 2′OH of ribose and isomerization of uridine to form pseudouridine (**Figure 17.27**). Several snoRNAs are involved in certain tRNA and snRNA base modifications.

The **microRNAs** (**miRNAs**), between 22 and 26 nt in length, are involved in gene expression regulation. After binding to several proteins to form the **RNA-induced silencing complex** (**RISC**), each type of miRNA binds to complementary base sequences on the 3′UTR of target mRNAs, thereby preventing their translation or enhancing their degradation. Each miRNA (at least 1000 in humans) that is expressed in a cell may target as many as 200 mRNAs. MiRNAs are believed to regulate about 60% of human protein-coding genes.

Small interfering RNAs (siRNAs) are 21- to 23-nt double-stranded RNAs (dsRNAs) with 2-nt-long 3′ overhangs that play a crucial role in *RNA interference* (RNAi) (see p. 738). RNAi is an RNA-degrading process involving RISC that defends cells from RNA-containing viruses and any inadvertently transcribed transposons. siRNAs are also used to interfere with the expression of specific genes by initiating the degradation of mRNAs.

Piwi-interacting RNAs (**piRNAs**) are ncRNAs that interact with Piwi proteins, a family of RNA-binding proteins within the gene-silencing Argonaute superfamily (p. 738). PiRNAs are a large class of small noncoding single-stranded RNAs (26–31 nt in length) that are transcribed from clusters unevenly distributed across the genome. PiRNAs were once believed to be primarily involved in preserving genomic integrity by recognizing and silencing retrotransposons in the germ line genome in the mammalian testes. They have also been found in somatic cells of mammals and other animals (e.g., *Drosophila melanogaster* and *Caenorhabditis elegans*), where they have been linked to euchromatin remodeling and developmental process regulation.

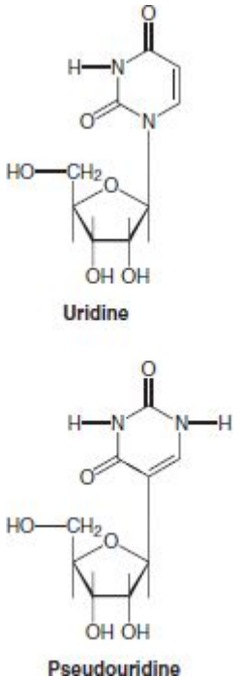


FIGURE 17.27

Structures of uridine and pseudouridine.

3D animation of uridine

3D animation of pseudouridine

Long ncRNAs (**lncRNAs**) are expressed at low levels and are often tissue-specific. Many lncRNAs have structures resembling mRNAs because they are spliced, polyadenylated, and 5′-capped. The roles of most of the approximately 28,000 lncRNAs in humans are not known. Those whose functions have been defined are involved in regulation of transcription and posttranscriptional modification reactions, cell cycle control, and apoptosis. LncRNAs have also been shown to encode miRNAs. Cell processes involving lncRNAs include cell cycle regulation and epigenetic regulation such as imprinting (the differential expression of a gene, depending on whether it was paternally or maternally inherited).

LncRNAs facilitate a diverse range of regulatory mechanisms because their structures allow them to achieve temporal and spatial gene regulation by binding directly to DNA or indirectly via RNA-protein interactions. Synthesized as a result of signaling mechanisms, diverse lncRNAs can impact transcriptional activation or repression by serving as *guides* that recruit proteins to specific chromatin locations, or acting as *scaffolds* that assemble combinations of regulatory proteins into protein complexes (e.g., chromatin-modifying complexes or a transcription preinitiation complex; p. 725). Certain lncRNAs can also act as *decoys*, that is, they bind to targets such as transcription factors, DNA sequences, or other RNA molecules to inhibit their functions. For example, when certain lncRNAs bind to their target mRNA, its translation is inhibited.

KEY CONCEPTS

- RNA is a nucleic acid that is involved in various aspects of protein synthesis and in the regulation of gene expression.
- RNAs directly involved in protein synthesis are transfer RNA, ribosomal RNA, and messenger RNA.
- Important noncoding RNAs include snRNAs snoRNAs, siRNAs, miRNAs, piRNAs, and lncRNAs.

QUESTION 17.8

When a gene is transcribed, only one DNA strand acts as a template for the synthesis of the RNA molecule. This strand is referred to as an **antisense** (or non-coding) strand; the nontranscribed DNA strand is called the **sense** (or coding) strand. The base sequence of the sense strand is the DNA version of the mRNA used to synthesize the polypeptide product of the gene. The antisense RNA (the transcript of the antisense DNA strand) plays a role in transcriptional and translational regulation. An antisense RNA can anneal specifically to a corresponding mRNA and prevent translation.

Because mRNA-antisense RNA binding is so specific, antisense molecules are considered promising research tools. Numerous investigators are using antisense RNA molecules to study eukaryotic function by selectively turning on and off the activities of specific genes. This so-called *reverse genetics* is also useful in medical research. Although serious problems have been encountered in antisense research (e.g., the inefficiency of inserting oligonucleotides into living cells and high manufacturing costs), antisense technology has already provided valuable insight into the mechanisms of cancer and viral infections. Consider the following sense DNA sequence:

5′-GCATTCGAATTGCAGACTCCTGCAATTCGGCAAT-3′

Determine the sequence of its complementary strand. Then determine the mRNA and antisense RNA sequences. (Recall that in RNA structure, U is substituted for T. So A in a DNA strand is paired with a U as RNA is synthesized.)

17.3 VIRUSES

Viruses lack most of the properties that distinguish life from nonlife, the most important of which is the inability to carry on metabolic processes. Yet under specific conditions they can wreak havoc on living organisms. Often described as obligate intracellular parasites, viruses can also be viewed as mobile genetic elements because of their structure, a piece of nucleic acid enclosed in a protective coat. Once a virus has infected a host cell, its nucleic acid can hijack the cell's nucleic acid and protein-synthesizing machinery. As viral components accumulate, complete new viral particles are produced and then released from the host cell. In many circumstances, so many new viruses are produced that the host cell lyses (ruptures). Alternatively, the viral nucleic acid may insert itself into a host chromosome, resulting in a latent infection. A latent virus may or may not be subsequently reactivated.

Viruses occur in a bewildering array of sizes and shapes. Virions (complete viral particles) range from 10 nm to approximately 400 nm in diameter. Although most viruses are too small to be seen with the light microscope, a few (e.g., the pox viruses) can be visualized because they are as large

as the smallest bacteria.

Simple virions (e.g., bacteriophage T-4) are composed of a *capsid* (a protein coat made of interlocking protein molecules called capsomeres), which encloses nucleic acid. (The term *nucleocapsid* is often used to describe the complex formed by the capsid and the nucleic acid.) Most capsids are either helical or icosahedral (20-sided structures composed of triangular capsomeres). The nucleic acid component of virions is either DNA or RNA.

Most viruses possess double-stranded DNA (dsDNA) or single-stranded RNA (ssRNA). However, viruses with ssDNA and dsRNA genomes have also been observed. There are two types of ssRNA genome. A *positive-sense* RNA genome [(+)-ssRNA] acts as a giant mRNA; that is, it directs the synthesis of a long polypeptide that is cleaved and processed into smaller molecules. **Retroviruses** such as human immunodeficiency virus (HIV) are single-stranded positive sense viruses with a life cycle that involves a DNA intermediate. They use their own reverse transcriptase to synthesize DNA from an RNA genome. The new DNA molecule may then be inserted into the host cell genome where it may lie dormant (a latent infection) as a provirus. Once the provirus has been reactivated its DNA is transcribed yielding mRNAs that directs the synthesis of viral proteins, and genome copies that will be incorporated into new virions. *Negative-sense* single-stranded RNA viruses such as Ebola virus have genomes [(-)-ssRNA] that must be converted to positive-sense mRNAs required for viral protein synthesis. New virions form using newly synthesized viral proteins and negative-sense RNAs.

In many viruses, the nucleocapsid is surrounded by a membrane envelope, which usually arises from the host cell nuclear or plasma membranes. Envelope proteins, coded for by the viral genome, are inserted into the envelope membrane during virion assembly. Proteins that protrude from the surface of the envelope, called spikes, are believed to mediate the attachment of the virus to the host cell. HIV and ebola virus are examples of enveloped viruses.

KEY CONCEPTS

- Viruses are composed of nucleic acid enclosed in a protective coat. The nucleic acid may be a single- or double-stranded DNA or RNA.
- In simple viruses, the protective coat, called a capsid, is composed of protein.
- In more complex viruses, the nucleocapsid, composed of nucleic acid and protein, is surrounded by a membranous envelope derived from host cell membrane.

Visit the companion website at **www.oup.com/us/mckee** to read the Biochemistry in Perspective essay on HIV infection.

Bacteriophage T4: A Viral Lifestyle

The T4 bacteriophage (**Figure 17.28**) is a large virus with an icosahedral head and a long, complex tail. The head contains dsDNA, and the tail attaches to the host cell and injects the viral DNA into the host cell.

The life cycle of T4 begins with the virion adsorbing on to the surface of an *E. coli* cell. The entire virion cannot penetrate into the cell's interior because the bacterial cell wall is rigid. Instead, the virion injects its DNA into the cell by flexing and constricting the tail apparatus. Once the viral DNA has entered the cell, the infective process is complete, and the next phase (replication) begins.

Within 2 minutes after the injection of T4 phage DNA into an *E. coli* cell, synthesis of host

DNA, RNA, and protein stops and phage mRNA synthesis begins. Phage mRNA codes for the synthesis of capsid proteins and some of the enzymes required for replication of the viral genome and the assembly of virion components. In addition, other enzymes are synthesized that weaken the cell wall of the host, allowing release of the new phage for new rounds of infection. Approximately 22 minutes after the injection of viral DNA (vDNA), the host cell, now filled with several hundred new virions, lyses. Upon release, the virions attach to nearby bacteria, thus initiating new infections.

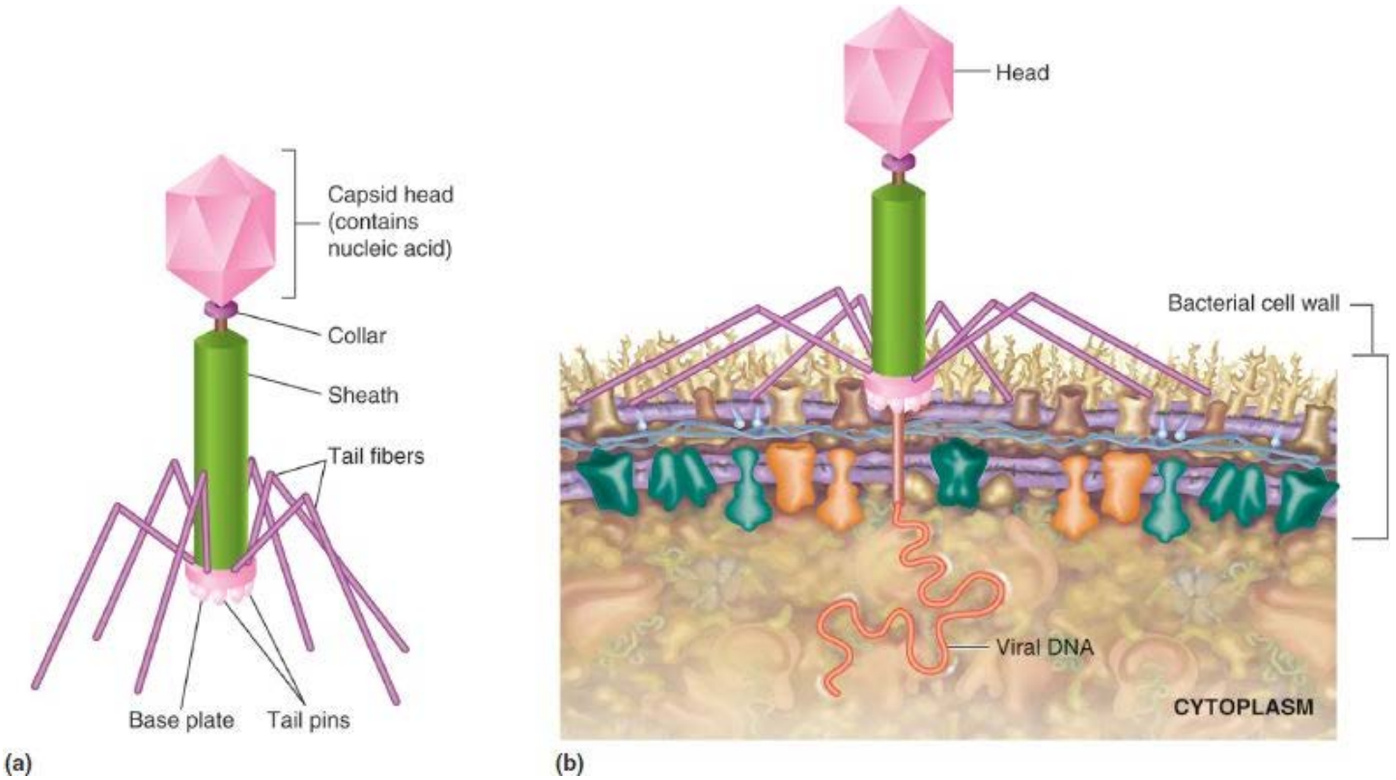


FIGURE 17.28

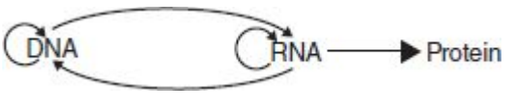
The T4 Bacteriophage

(a) The structure of an intact T4 bacteriophage. (b) Penetration of the cell wall and injection of viral DNA (vDNA) into the host cell by the bacteriophage. The vDNA directs the host cell to synthesize about 30 proteins that facilitate new virion synthesis.

Bacteriophage that initiate this so-called **lytic cycle** are referred to as *virulent* because they destroy their host cells. Many phage, however, do not initially kill their hosts. So-called *temperate* or *lysogenic* phage integrate their genome into that of the host cell. (The term **lysogeny** describes a condition in which the phage genome is integrated into a host chromosome.) The integrated viral genome, called the **prophage**, is copied along with host DNA during cell division for an indefinite time. Occasionally, lysogenic phage enter a *lytic* phase. Certain external conditions, such as UV or ionizing radiation, activate the prophage, which then directs the synthesis of new virions. Sometimes, a lysing bacterial cell releases a few virions that contain some bacterial DNA along with the phage DNA. When such a virion infects a new host cell, this DNA is introduced into the host genome. This process is referred to as **transduction**.

QUESTION 17.9

Recall that according to the central dogma, the flow of genetic information is from DNA to RNA and then to protein. Retroviruses are an exception to this rule. The alterations of the central dogma that are observed in retroviruses and other RNA viruses can be illustrated as follows:



Compare this with the original central dogma (p. 638). Describe in your own words the implications of each component of these figures.

Biochemistry IN PERSPECTIVE

Ebola Virus

What is Ebola, and why is it so dangerous? Sir Peter Medawar (1915–1987), a Nobel Prize-winning immunologist, described viruses as "a piece of bad news wrapped in a protein coat." Although only a small fraction of the millions of viruses known to exist have affected humans directly, these vanishingly small parasites have had an enormous impact on our history. Epidemics of viral diseases such as smallpox, poliomyelitis, measles, and influenza, for example, have killed or maimed millions of people in the twentieth century alone. Vaccines have tamed these viruses, but several newly emerging (recently discovered) viruses are now recognized as major threats to human health. Among these is Ebola virus, which provides a stark example of our vulnerability.

First observed in 1977, Ebola virus (**Figure 17K**) causes a severe hemorrhagic fever (internal and external bleeding) with an exceptionally high fatality rate. Ebola Zaire, one of three strains of the virus, kills up to 90% of the individuals that it infects; the Sudan strain kills about 71%. Ebola is a filamentous enveloped virus with a single-stranded negative-sense RNA genome (18,960 nt in length). The virus has three layers: a nucleocapsid and a matrix space that is surrounded by a membrane envelope. The RNA genome codes for seven proteins (**Figure 17L**). The nucleocapsid is a protein complex formed by NP (nucleoprotein), VP24, and VP35 that encapsulates the viral genome. VP30, also found in the nucleocapsid, has a role as a transcription factor and in ssRNA packaging into the nucleocapsid of new virions. VP40 codes for a matrix protein that exists in different conformational and oligomeric states, and has roles in viral assembly and the budding process that releases the virus from infected cells. The L protein codes for an RNA-dependent RNA polymerase. The GP (glycoprotein) gene transcript undergoes transcriptional editing, which results in several protein products, including GP1 and GP2 subunits of the viral surface protein and soluble GP (sGP).

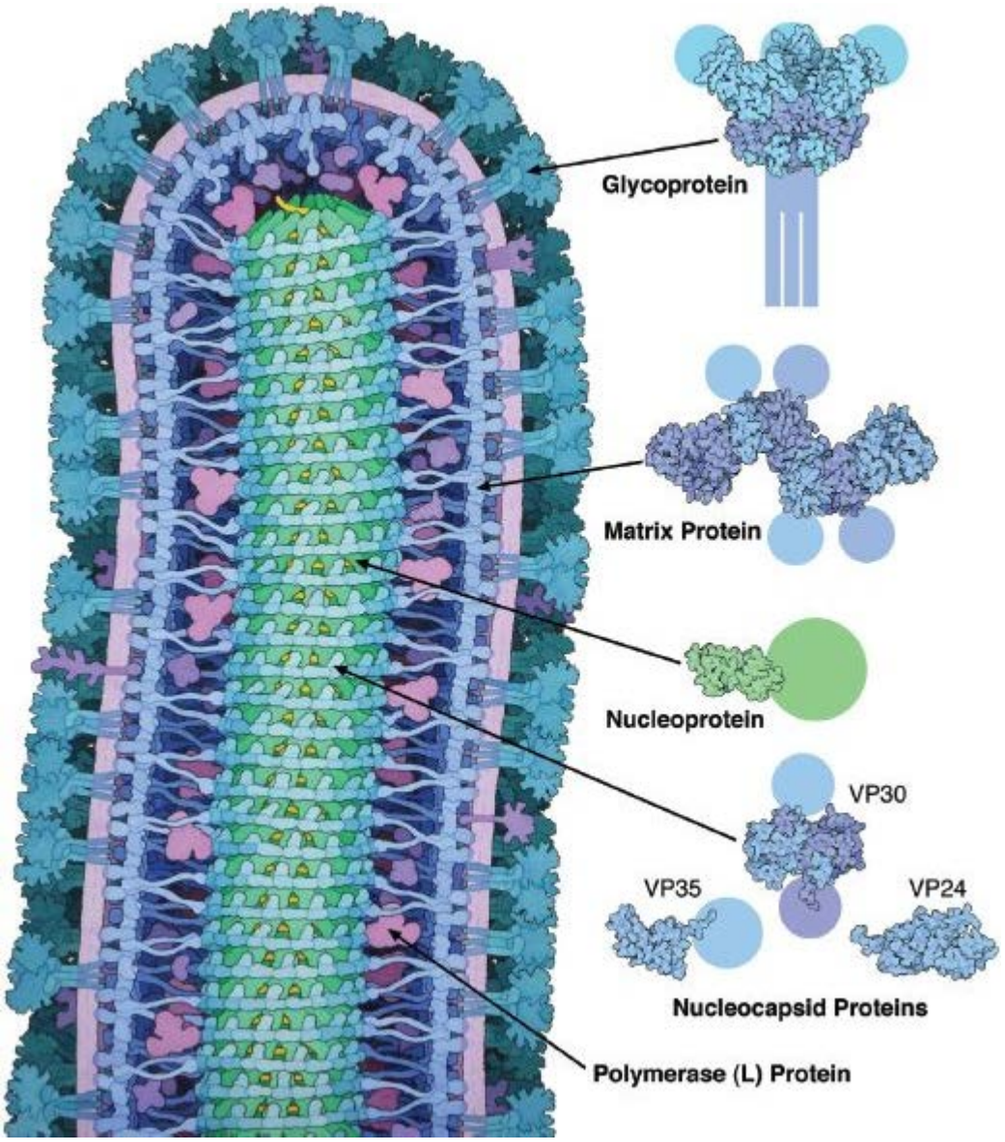


FIGURE 17K

Ebola Virus Structure

The Ebola virus is a filamentous virus with a single-strand negative-sense RNA that is encased in a nucleocapsid composed of nucleoprotein (NP), VP24, VP30, and VP35. The Ebola matrix protein (VP40) contributes to the virus's shape and is involved in the budding process. The viral glycoprotein (BP) initiates entry into host cells by binding to cell receptors and then initiating fusion with endosomal membranes. The L RNA polymerase synthesizes the positive-sense copy of the viral RNA genome (i.e., an mRNA) that is used by host cell ribosomes to synthesize viral proteins. The L RNA polymerase also uses the positive-sense RNA to synthesize the negative-sense copies to be incorporated into a new virus.

The Ebola viruses enter a cell through a nonspecific pinching off of the cell membrane called *macropinocytosis*. Once inside, the binding of the virus's GP1 subunit to NPC1 (a cholesterol transport regulatory protein) and TPC2 (a Ca^{2+} permeable channel), receptors on endosomal membrane, triggers a conformational change that exposes the GP2 subunit fusion loop, thereby initiating insertion into the membrane. Once the virus enters an endosome, the nucleocapsid is released, and the negative-sense RNA genome is converted to an mRNA by the L RNA polymerase. Subsequently, host cell ribosomes synthesize viral proteins using the viral mRNA. The L RNA polymerase synthesizes new copies of the negative-sense RNA genome that are subsequently combined with newly synthesized viral proteins to form the nucleocapsid. The new viral particles exit the cell by a budding process that envelops them in host cell membrane in which

GP has been inserted.

Ebola virus enters the body through mucous membranes or breaks in the skin when in direct contact with blood or body fluids. Patients with Ebola virus disease (EVD) develop a fever within 8 to 12 days after exposure. The virus, which infects a variety of cell types, migrates from the initial infection site to the nearest lymph nodes, where it begins its assault on immune system cells. The virus is shielded from the immune system by large amounts of sGP that bind to GP antibodies, making the body's antibody response ineffective. The virus then spreads to most organs and tissues (e.g., liver, spleen, adrenal gland, brain, intestines, and eyes). The damaged immune system eventually initiates a cytokine storm, an excessive and uncontrolled release of inflammatory proteins (e.g., interleukins and TNFβ) and ROS that further damage organs and blood vessels. Death is caused by massive organ failure or hypovolemic shock (severe blood loss caused by impaired blood clotting).

Currently, EVD patients are treated with symptom reduction therapies such as intravenous fluids and electrolytes, and medications that support blood pressure and reduce GI symptoms (e.g., vomiting), fever, and pain. There are no antiviral medications for EVD, but there are currently two promising vaccine trials supported by the National Institute of Allergy and Infectious Disease (NIAID).

FIGURE 17L

The Ebola Virus Genome

The single-stranded negative-sense RNA genome is composed of seven genes: nucleoprotein (NP), VP35, VP40, glycoprotein (GP), VP30, VP24, and the RNA-directed RNA polymerase (L).

SUMMARY The Ebola virus is a single-stranded negative-sense RNA virus that attacks virtually every organ. In an astonishingly short time, medical researchers have identified and largely characterized the Ebola virus and its devastating effects on the human body.

Chapter Summary

1. The information required for directing all living processes is stored in the nucleotide sequences of DNA. DNA is composed of two antiparallel polynucleotide strands that wind around each other to form a right-handed double helix. The deoxyribose-phosphodiester bonds form the backbones of the double helix, and the nucleotide bases project to its interior. The nucleotide base pairs form because of hydrogen bonding between the following bases: adenine with thymine and cytosine with guanine.
2. Several types of noncovalent interaction contribute to the stability of DNA's structure: hydrophobic and van der Waals's interactions between stacked heterocyclic bases, hydrogen bonds between GC and AT base pairs, and hydration with water molecules. Decoding of the genetic information in DNA requires molecular machinery largely composed of proteins.
3. Mutations are changes in DNA structure, which may be caused by collisions with solvent molecules, thermal fluctuations, ROS, radiation, or xenobiotics. In a transition mutation, one pyrimidine base is substituted for another pyrimidine base, or a purine base is substituted for another purine. In a transversion mutation, a pyrimidine base is substituted for a purine base or vice versa. Other mutation

types include indels (insertions and deletions), inversions, translocations, and duplications.

4. DNA can have several conformations, depending on the nucleotide sequence and the degree of hydration of the double helix. In addition to the classical structure determined by Watson and Crick (B-DNA), A-DNA and Z-DNA have also been observed. DNA supercoiling is a critical feature of several biological processes, such as DNA packaging, replication, and transcription.
5. Each eukaryotic chromosome is composed of nucleohistone, a complex formed by winding a DNA molecule around histone octamers to form nucleosomes. DNA methylation and several types of histone covalent modification (e.g., acetylation and methylation) alter chromatin structure and gene expression. The DNA of mitochondria and chloroplasts is similar to the chromosomes found in prokaryotes.
6. A genome is the full set of the inherited instructions required to sustain an organism's living processes. Although there are some similarities, the genomes of prokaryotes and eukaryotes are substantially different in size, coding capacity, coding continuity, and gene expression regulation mechanisms.
7. The majority of the DNA sequences in humans do not code for proteins or functional RNAs. There are two general classes of intergenic sequences: tandem repeats and interspersed genome-wide repeats. Mobile genetic elements can be duplicated and then move within the genome. Retrotransposons can cause disease by inserting into genes or regulatory sequences.
8. RNA differs from DNA in that it contains ribose (instead of deoxyribose), has a somewhat different base composition, and is usually single-stranded. The forms of RNA involved in protein synthesis are transfer, ribosomal, and messenger RNAs. Transfer RNA molecules have specific amino acids attached to them by specific enzymes and transport these to the ribosome for incorporation into newly synthesized protein, where they are properly aligned during protein synthesis. The ribosomal RNAs are components of ribosomes, where they are the sites of catalytic activity. Messenger RNA contains within its nucleotide sequence the coding instructions for synthesizing a specific polypeptide. There are several classes of noncoding RNAs that have diverse roles in genome regulation and protection. Examples include snRNAs, snoRNAs, miRNAs, siRNAs, and piRNAs. Lnc RNAs are a group of ncRNAs that are involved in diverse regulatory processes that include transcription, cell cycles and epigenetic DNA and histone modifications.
9. Viruses are obligate intracellular parasites. Although they are acellular and cannot carry out metabolic activities on their own, viruses can wreak havoc on living organisms. Each type of virus infects a specific type of host (or small set of hosts). This is possible because a virus can either inject its genome into the host cell or gain entrance for the entire virion by fusing to host cell membranes. Each virus has the capacity to use the host cell's metabolic processes to manufacture new virions. Viruses possess dsDNA, ssDNA, dsRNA, or ssRNA genomes.
10. Retroviruses are a class of RNA viruses that possess a reverse transcriptase activity that converts their RNA genome to a DNA molecule. The retrovirus HIV causes AIDS.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on nucleic acids to help you prepare for exams.

Chapter 17 Review Quiz

Suggested Readings

Belair C, et al. 2018. Noncoding RNA surveillance: the end justifies the means. Chem Rev 118(8):4422–47.

Brickner J. 2017. Genetic and epigenetic control of the spatial organization of the genome. Mol Biol Cell 28(3)364–9.

Carey N. 2013. The epigenetics revolution: how modern biology is rewiring our understanding of

genetics, disease and inheritance. New York (NY): Columbia University Press.
Carey N. 2017. Junk DNA: a journey through the dark matter of the genome. New York (NY): Columbia University Press.
Eagen KP. 2018. Principles of chromosome architecture revealed by Hi-C. Biochem Sci 43(6):469–78.
Flaus A, Owen-Hughes T. 2017. Unlocking the nucleosome: both DNA and histones flex to undergo remodeling. Science 355:245–6.
Fyodorov DV, et al. 2018. Emerging roles of linker histones in regulating chromatin structure and function. Nat Rev Mol Cell Biol 19(3):192–206.
Hansen AS, et al. 2018. Recent evidence that TADs and chromatin loops are dynamic structures. Nucleus 9(1):20–32.
Hentze MW, et al. 2018. A brave new world of RNA binding proteins. Nat Rev Mol Cell Biol 19(5):327–41.
Reno PL. 2017. Missing links: our big brains, upright gait, and style of love may exist because we shed key pieces of DNA. Sci Amer 316(5):42–7.
Solovei I, et al. 2016. How to rule the nucleus: divide et impera. Curr Opin Cell Biol 40:47–59.
Zuccato C, Cattaneo E. 2016. The Huntington's paradox. Sci Amer 315(2):56–61.

Key Words

A-DNA, *646*
alkylating agent, *644*
antisense strand, *675*
base analogues, *644*
B-DNA, *646*
centromere, *659*
chain-terminating method, *669*
Chargaff's rules, *646*
chromatin, *651*
chromosome, *650*
cistron, *673*
CpG, *662*
CpG island, *662*
denaturation, *666*
DNA fingerprinting,*671*
DNA profile, *671*
DNA typing, *671*
endogenous retrovirus, *660*
epigenetics, *662*
epigenome, *662*
epimutation, *664*
euchromatin, *653*
exon, *658*

gene duplication, *643*
genes, *635*
genetics, *635*
genome, *636*
heterochromatin, *653*
hybridization, *667*
hypochromic effect, *666*
indel, *642*
intercalating agent, *645*
interspersed genome-wide repeat, *659*
intron, *658*
inversion mutation, *643*
inverted repeat, *648*
LINE, *660*
long noncoding RNA, *674*
lysogeny, *677*
lytic cycle, *677*
messenger RNA, *673*
metabolome, *637*
methyl-CpG-binding protein, *662*
microRNA, *674*
microsatellite, *659*
minisatellite, *659*
missense mutation, *642*
mobile genetic element, *659*
molecular biology, *646*
nonalkylating agent, *645*
noncoding RNA, *673*
nonsense mutation, *642*
nucleosomes, *651*
open reading frame, *673*
operon, *657*
piwi-interacting RNA, *674*
point mutation, *642*
prophage, *677*
proteome, *637*
replication, *636*
retrotransposon, *660*
retrovirus, *676*
ribosomal RNA, *672*
ribozyme, *665*
RNA-induced silencing complex, *674*

RNA transposons, *660*
satellite DNA, *659*
sense strand, *675*
short tandem repeats, *671*
silencer, *658*
silent mutation, *642*
SINE, *660*
single nucleotide polymorphism, *642*
small interfering RNA, *674*
small nuclear ribonucleoprotein, *674*
small nuclear RNA, *674*
small nucleolar RNA, *674*
Southern blotting, *667*
spliceosome, *674*
tandem repeats, *659*
telomere, *659*
topologically associated domain, *653*
transcript, *636*
transcription, *636*
transcription factor, *637*
transcriptome, *636*
transduction, *677*
transfer RNA, *665*
transition mutation, *642*
translation, *637*
translocation, *643*
transposition, *659*
transposon, *659*
transversion mutation, *642*
Z-DNA, *647*

Review Questions

SECTION 17.1

Comprehension Questions

1. Define the following terms:
    a. molecular biology
    b. genetics
    c. replication
    d. transcription

e. transcriptome

2. Define the following terms:

   a. transcript
   b. proteome
   c. metabolome
   d. double helix
   e. base stacking

3. Define the following terms:

   a. point mutation
   b. transition mutation
   c. transverse mutation
   d. silent mutation
   e. missense mutation

4. Define the following terms:

   a. single nucleotide polymorphism
   b. nonsense mutation
   c. indel
   d. inversion
   e. translocation

5. Define the following terms:

   a. alkylating agents
   b. base analogue
   c. nonalkylating agent
   d. intercalating agent
   e. ethidium bromide

6. Define the following terms:

   a. Chargaff’s rules
   b. constitutive heterochromatin
   c. bacteriophage
   d. satellite DNA
   e. LINE

7. Define the following terms:

   a. A-DNA
   b. B-DNA
   c. pseudogene
   d. cruciform
   e. intron

8. Define the following terms:

   a. positive supercoiling
   b. SINE
   c. polyamines
   d. euchromatin

e. nucleosome

9. Define the following terms:

   a. histone
   b. heterochromatin
   c. spermine
   d. intergenic sequences
   e. tandem repeats

10. Define the following terms:

   a. epigenetics
   b. CpG island
   c. epimutation
   d. DNA methylation
   e. histone acetylation

11. Define the following terms:

   a. DNA typing
   b. short tandem repeats
   c. DNA profile
   d. nucleosome
   e. retrotransposon

Fill in the Blanks

12. The term _____ refers to the sum total of all the low-molecular-weight metabolite molecules produced by a cell.
13. The set of proteins produced by a cell is called its _____.
14. An inappropriate base change that has no discernible effect is called a _____ mutation.
15. A _____ is a point mutation that converts a codon for an amino acid to a premature stop signal.
16. According to the _____ model, TAD formation begins with the binding of CTCF with a cognate DNA sequence
17. _____ are structures at the ends of eukaryotic chromosomes that buffer DNA from critical loss of coding sequences during a round of replication.
18. A noncoding DNA sequence within a gene is called a(n) _____.

Short-Answer Questions

19. List three differences between eukaryotic and prokaryotic DNA.
20. List three biochemical processes facilitated by supercoiling.
21. Describe the structure of the nucleosome.
22. Describe the role of promoters, enhancers, silencers and insulators in gene expression.
23. Z-DNA derives its name from the zig-zag conformation of phosphate groups. What features of the DNA molecule allow this structure to form?
24. There is one base pair for every 0.34 nm of DNA, and the total contour length of all DNA in a single human cell is approximately 2 m. With this information, calculate the number of base pairs in a single cell. Assuming that there are 10^{14} cells in the human body, calculate the total length of DNA. How does this estimate compare to the distance from the Earth to the sun

(149.6×10^6 km)?

25. A DNA sample contains 21% adenine. What is the complete percentage base composition?
26. The melting temperature of a DNA molecule increases as the GC content increases. Explain.
27. Describe the impact of the 1953 publication of the Watson–Crick paper on genetic research?
28. What error would cause more damage to a cell: A DNA replication or a DNA transcription error?
29. When an aromatic hydrocarbon intercalates between two stacked base pairs, what effect may there be on DNA structure?
30. Chargaff's rules apply to DNA, but not RNA. Explain
31. Describe how water affects DNA structure.
32. Describe the types of noncovalent interactions that stabilize DNA structure.
33. Mitochondrial DNA is especially vulnerable to ROS. Suggest several DNA-protecting mechanisms that you would expect to find in mitochondria.
34. Describe the evidence that Watson and Crick used to construct their DNA model.
35. Compare the size and coding capacity of prokaryotic genomes with those of eukaryotes. What other features distinguish them?
36. The xenobiotic molecule ethyl chloride (CH_3CH_2Cl) is mutagenic. To what class of mutagenic substances does it belong?
37. What are Alu elements? How can they adversely affect human health?
38. The polyamines spermine and spermidine (see p. 651) have numerous effects on both prokaryotic and eukaryotic cells. Examples include chromatin condensation, transcription, translation, and apoptosis. They are best known for their role in promoting DNA stability. Explain how polyamines enhance DNA stability and promote supercoiling.

Critical-Thinking Questions

39. Preeclampsia is the most common dangerous complication of pregnancy. Characterized by high blood pressure and the occurrence of protein in the urine, preeclampsia occurs most commonly after the 32nd week of pregnancy. Without treatment, patients with preeclampsia may progress to eclampsia, a life-threatening disorder, characterized by placental rupture, seizures, coma, and the risk of maternal death. The causes of preeclampsia and eclampsia remain uncertain. However, there may be a connection between these disorders and an endogenous retrovirus sequence located on chromosome 7. Explain.
40. What structural features of DNA cause the major groove and minor groove to form?
41. Jerome Vinograd found that when circular DNA from a polyoma virus is subjected to a cesium chloride gradient centrifugation, it separates into two distinct bands, one consisting of supercoiled DNA and the other of relaxed DNA. Explain how you would identify each band.
42. 5-Bromouracil is an analogue of thymine that usually pairs with adenine. However, 5-bromouracil frequently pairs with guanine. Explain.
43. Unlike linker DNA and deproteinized DNA, DNA segments wrapped around histone cores are relatively resistant to the hydrolytic actions of nucleases. Explain.
44. Some living organisms are under considerable pressure to streamline their genomes for the sake of more efficient operation. As a result, the mitochondria of eukaryotic species have lost, to one degree or another, the overwhelming majority of their genes. During this process, several hundred mitochondrial genes were transferred to the nuclear genome. Yet mitochondria still retain a genome, with the capacity to produce several electron transport proteins. Review mitochondrial electron transport and suggest a reason why these energy-generating organelles retained the genes to produce this set of molecules.
45. It has been estimated that each phosphate group in B-DNA can form hydrogen bonds with six

water molecules. Draw a diagram of a phosphodiester linkage with its associated water molecules.

46. The suggestion has been made that DNA extracted from ancient fossils could be extracted, cloned, and used to resurrect an extinct species. Comment on the practicality of this idea.
47. Fluorouracil is a structural analogue of thymine. The fluorine promotes enolization. Suggest a reason why this effect used in the treatment of cancer?

SECTION 17.2

Comprehension Questions

48. Define the following terms:
    a. noncoding RNA
    b. miRNA
    c. siRNA
    d. snoRNA
    e. lncRNA
49. Define the following terms:
    a. ribozyme
    b. stem loop
    c. tRNA
    d. mRNA
    e. rRNA
50. Define the following terms:
    a. spliceosome
    b. RNA interference
    c. ORF
    d. cistron
    e. piRNA
51. Define the following terms:
    a. tiRNA
    b. monocistronic
    c. snRNPs
    d. antisense DNA strand
    e. sense DNA strand

Fill in the Blanks

52. Abnormal epigenome changes are referred to as _____.
53. When nitrogenous bases in single-stranded polynucleotides are incorporated into double-stranded molecules, base pair formation results in altered UV light absorption. This is called the _____ effect.
54. The base sequence within mRNA molecules that codes for a polypeptide is called a (n) _____.

55. RNA that is not directly involved in polypeptide synthesis is called _____.
56. _____ are molecular machines that excise introns from pre-mRNA and then join exons together.

Short-Answer Questions

57. Provide the complementary strand and the RNA transcription product for the following DNA template segment: 5′ AGGGGCCGTTATCGTT-3′.
58. Describe the structural features that stabilize RNA molecules.
59. Describe how DNA methylation patterns are retained from one cell generation to the next.
60. Describe in general terms the histone code hypothesis.
61. Describe the steps in Southern blotting.
62. Identical twin brothers begin life with identical genomes and epigenomes. How will this circumstance change with age? Suggest how these changes could be used as a forensic tool.
63. In ion torrent DNA sequencing, covalent bond formation is detected by an ion sensor sensitive to protons. What is the source of the protons?

Critical-Thinking Questions

64. An experiment was conducted with two groups of genetically identical female agouti mice (A^{vy}/a) that were mated with normal (a/a) male mice . Both female groups ate lab chow with bisphenol A (BPA) (50mg/Kg). (BPA, used in the manufacture of polycarbonate plastics, has estrogenic effects that include decreased CpG methylation.) The lab chow of one group was supplemented with methyl donors (e.g., folate, choline, and vitamin B_{12}). The effects of BPA were abolished in the A^{vy}a offspring of the mice fed with the methyl donor–supplemented chow (i.e., the coat color distribution observed was shifted toward brown). Explain.
65. The transmission of information is from DNA to RNA to proteins. Would it not be more efficient for DNA to code for proteins directly?
66. How and why do prokaryotic mRNA and eukaryotic mRNA differ?
67. The set of RNAs within in a cell changes over time. Explain why this circumstance is an advantage.

SECTION 17. 3

Comprehension Questions

68. Define the following terms:
    a. virus
    b. virion
    c. capsid
    d. retrovirus
    e. enveloped virus
69. Define the following terms:
    a. lytic cycle
    b. prophage
    c. transduction
    d. lysogeny

e. nuclocapsid

70. Define the following terms:

    a. negative sense RNA virus
    b. vDNA
    c. proviral DNA
    d. macropinocytosis
    e. temperate phage

Fill in the Blanks

71. A class of RNA viruses that possess reverse transcriptase activity are called _____.
72. A complex formed of a capsid and a nucleic acid is called a _____.
73. _____ is the process in which viral DNA is incorporated into the host cell genome.
74. The viral genome that is integrated into the bacterial genome is called the _____.

Short-Answer Questions

75. Can viruses be infected by other viruses.
76. Describe the infective process of a bacteriophage as it attacks a cell.
77. Describe the difference between lytic and lysogenic bacteriophage.

Critical-Thinking Questions

78. In Europe, for several centuries, measles has been a childhood infection that rarely kills. When European settlers came to the New World, however, the measles virus they brought with them killed a large portion of native populations. Explain.
79. Describe why Ebola virus infection is so easily spread and often fatal.

MCAT Study Questions

80. A mutation caused by a base deamination or a tautomerization is called a

    a. silent mutation
    b. transition mutation
    c. nonsense mutation
    d. missense mutation

81. The structural unit of the eukaryotic genome is called a(n):

    a. nucleosome
    b. histone
    c. chromatin
    d. exon

82. A small nonessential circular DNA molecule in prokaryotes is called a:

    a. plasmid
    b. telomere
    c. bacteriophage
    d. prophage

83. Promoters are DNA sequences

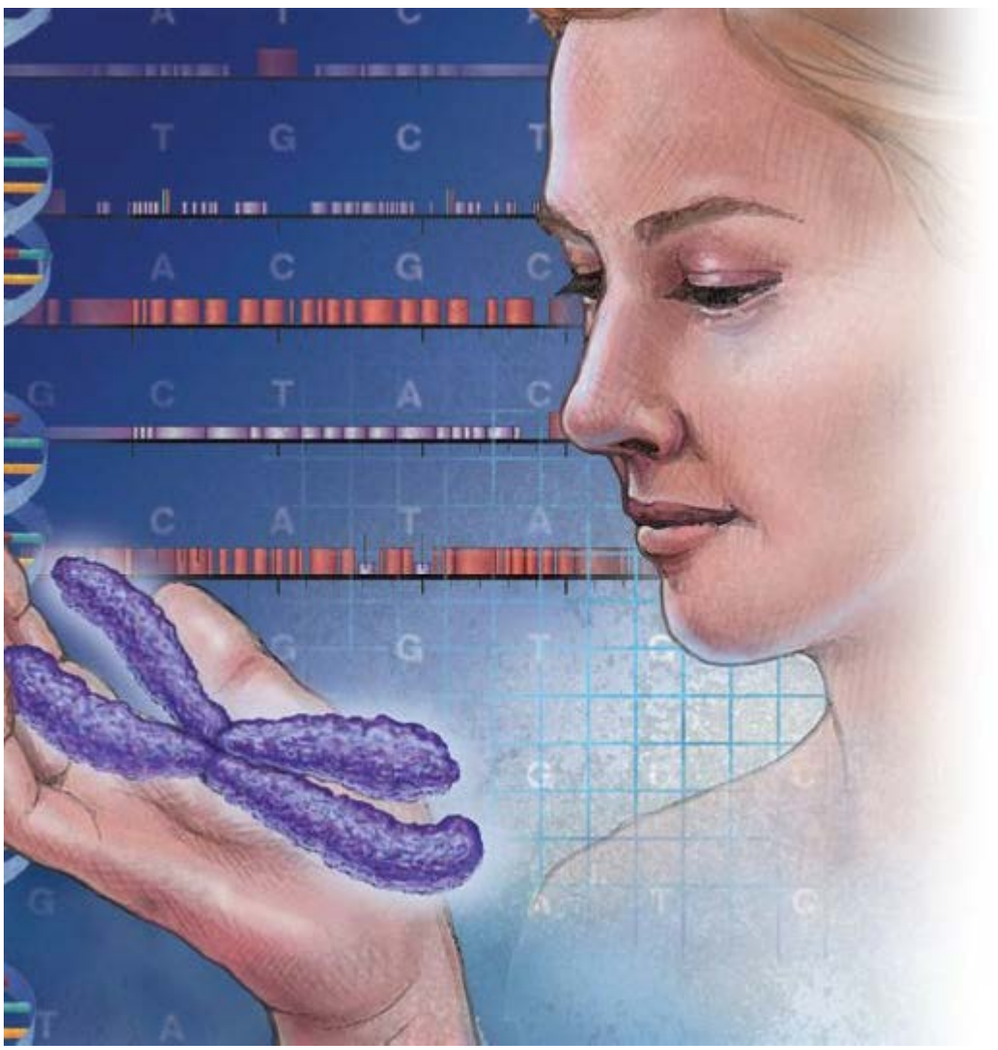
    a. near a transcription start site
    b. bound to a repressor protein

c. that inhibit transcription of a gene
d. that stimulate ncRNA activity

84. Chromatin remodeling is affected by two processes, DNA methylation and

a. base substitution
b. histone covalent modifications
c. ultraviolet radiation
d. transcription factor binding

CHAPTER 18 Genetic Information



The Human Genome and ENCODE (*Enc*yclopedia of *D*NA *E*lements) Projects The Human Genome Project, performed by thousands of scientists in public and private laboratories around the world, took more than 15 years to accomplish. Researchers are just beginning to interpret and utilize the resulting tidal wave of biological information to solve the medical and biological problems of humans. The ENCODE (*Enc*yclopedia *O*f *D*NA *E*lements) project, intended as a follow-up to the Human Genome Project, aims to identify and analyze all functional elements in the human genome.

OUTLINE

DNA AND CHIMERAS: A BIOLOGICAL AND LEGAL MYSTERY

18.1 GENETIC INFORMATION: REPLICATION, REPAIR, AND RECOMBINATION

DNA Replication
DNA Repair
DNA Recombination

18.2 TRANSCRIPTION
Transcription in Prokaryotes
RNAP and the Prokaryotic Transcription Process
Transcription in Eukaryotes

18.3 GENE EXPRESSION
Gene Expression in Prokaryotes
Gene Expression in Eukaryotes

Biochemistry in the Lab
Genomics

Biochemistry in the Lab
CRISPR

Biochemistry in Perspective
Carcinogenesis

AVAILABLE ONLINE
Biochemistry in Perspective
The Meselson–Stahl Experiment

DNA and Chimeras: A Biological and Legal Mystery

In 2002, the lives of two women intersected in a very surprising and unexpected way when their status as mothers was challenged. Living on opposite coasts of the United States, both women, L.F. in Washington State and K.K. in Boston, were leading ordinary lives until they underwent genetic testing. In December 2002, L.F., the pregnant mother of two other children, received the results of the DNA testing performed in support of her welfare application. Separated from her children's father and currently unemployed, L.F. could only receive temporary public support if the father's paternity was confirmed. To her utter astonishment, the test confirmed the father's paternity but revealed that she could not be her children's mother. Despite hospital birth records and assurances from family members and her obstetrician who witnessed the births, state prosecutors proceeded, solely on the basis of DNA test results, to charge L.F. with welfare fraud and threaten her with the loss of her children.

DNA profile analysis, the result of research spanning several decades, is considered the most accurate means of identifying individuals. Used in parental testing and forensic investigations, DNA profiling technology takes advantage of two facts: DNA is inherited from parents and, with the exception of identical twins, the DNA profile of each human is unique. How could well-accepted DNA profiling science be reconciled with the puzzling DNA profile results of L.F. and her children?

As a consequence of L.F.'s vehement protests, a trial judge appointed a court representative who witnessed the birth of the third child and the taking of blood samples from L.F. and the newborn infant. Subsequent DNA testing yielded a bewildering result: L.F. was not the mother of this child either. Fortunately, a prosecution lawyer found a precedent for this remarkable phenomenon in a recently published article in the *New England Journal of Medicine*, which reported a similar case.

K.K., a Boston teacher, was revealed to be a *chimera*, an individual with two genetically distinct,

intermingled cell lines. This discovery was the result of research triggered by a 1998 search for a suitable kidney donor that involved the testing of K.K. and her family members for major histocompatibility complex (MHC) antigens. MHC antigens, the cell-surface proteins coded for by genes on chromosome 6, allow immune system T-cells to distinguish between self and nonself (foreign cells). Although organ rejection is suppressed by antirejection drugs, transplant success is enhanced if the MHC antigens of the donor and the recipient are either identical or a close match. Genetic testing of blood samples from K.K. and her family (husband and three sons) revealed that K.K. was not the biological mother of two of her three sons.

Several years of intense research by genetic investigators at a Boston medical center eventually revealed that K.K. is a *tetragametic chimera*, an individual who developed from the fusion of two nonidentical zygotes (fertilized eggs) formed from four genetically distinct gametes (two eggs and two sperm). If fusion had not occurred, the two zygotes would have developed into fraternal female twins. The genetic markers found in two of her sons (and later in her brother's genetic profile) that differed from those in her blood and saliva were subsequently found in several other tissues (e.g., skin and thyroid cells). The researchers concluded that K.K. had two separate egg cell types in her ovaries.

As a result of the 2002 paper detailing the chimera research, several tissues of L.F. were tested. She was subsequently declared to be the mother of her children when the DNA markers missing in her blood were located in cervical epithelial cells.

Overview

ALL LIVING ORGANISMS ARE INFORMATION-PROCESSING SYSTEMS. THEIR ULTIMATE SOURCE OF INFORMATION IS ENCODED IN THE NUCLEOTIDE base sequence of DNA. As biochemists have searched ever more deeply into the mysteries of genetic information storage and transmission, of how DNA is replicated and gene expression is controlled, they have transformed all of the life sciences. The knowledge and technologies acquired during this pursuit have provided us with an understanding of the intricacies of living processes that is still unfolding.

In any successful information-based system, the instructions required to produce a certain type of organization (e.g., for building a house or for reproducing a living organism) must be stably stored to safeguard their accuracy and availability for use. Information must also be converted into a form that can be utilized. Living organisms have partitioned these functions as follows. DNA is a relatively stable molecule with structural features that maximize information storage and facilitate duplication. RNA molecules, more reactive than DNA, have numerous roles in protein synthesis and gene expression regulation. Finally, it is the many thousands of diverse proteins with their individual structural, catalytic, and machine-like properties that together sustain the living state.

DNA contains the genetic information that drives living processes; it does not directly control cellular processes. Decoding DNA base sequences requires molecular machinery, largely composed of proteins and ncRNAs and powered by cellular energy resources. These machines bend, twist, unwind, and unzip DNA during replication and transcription. At first glance, the seemingly repetitious and regular structure of DNA makes it an unlikely partner for the productive binding of specific base sequences with appropriate proteins. However, numerous contacts (often about 20 or so) involving hydrophobic interactions, hydrogen bonds, and ionic bonds between

amino acid residues and the edges of bases within the major groove (and to a lesser extent the minor groove) of DNA result in highly specific DNA-protein binding. The three-dimensional structures of most DNA-binding proteins analyzed thus far have surprisingly similar features. In addition to usually possessing a twofold axis of symmetry, many of these molecules can be separated into families (**Figure 18.1**) on the basis of the following structures: (1) helix-turn-helix, (2) helix-loop-helix, (3) leucine zipper, and (4) zinc finger. DNA-binding proteins, many of which are transcription factors, often form dimers. For example, a variety of transcription factors with leucine zipper motifs form dimers as their leucine-containing α-helices associated via van der Waals interactions.

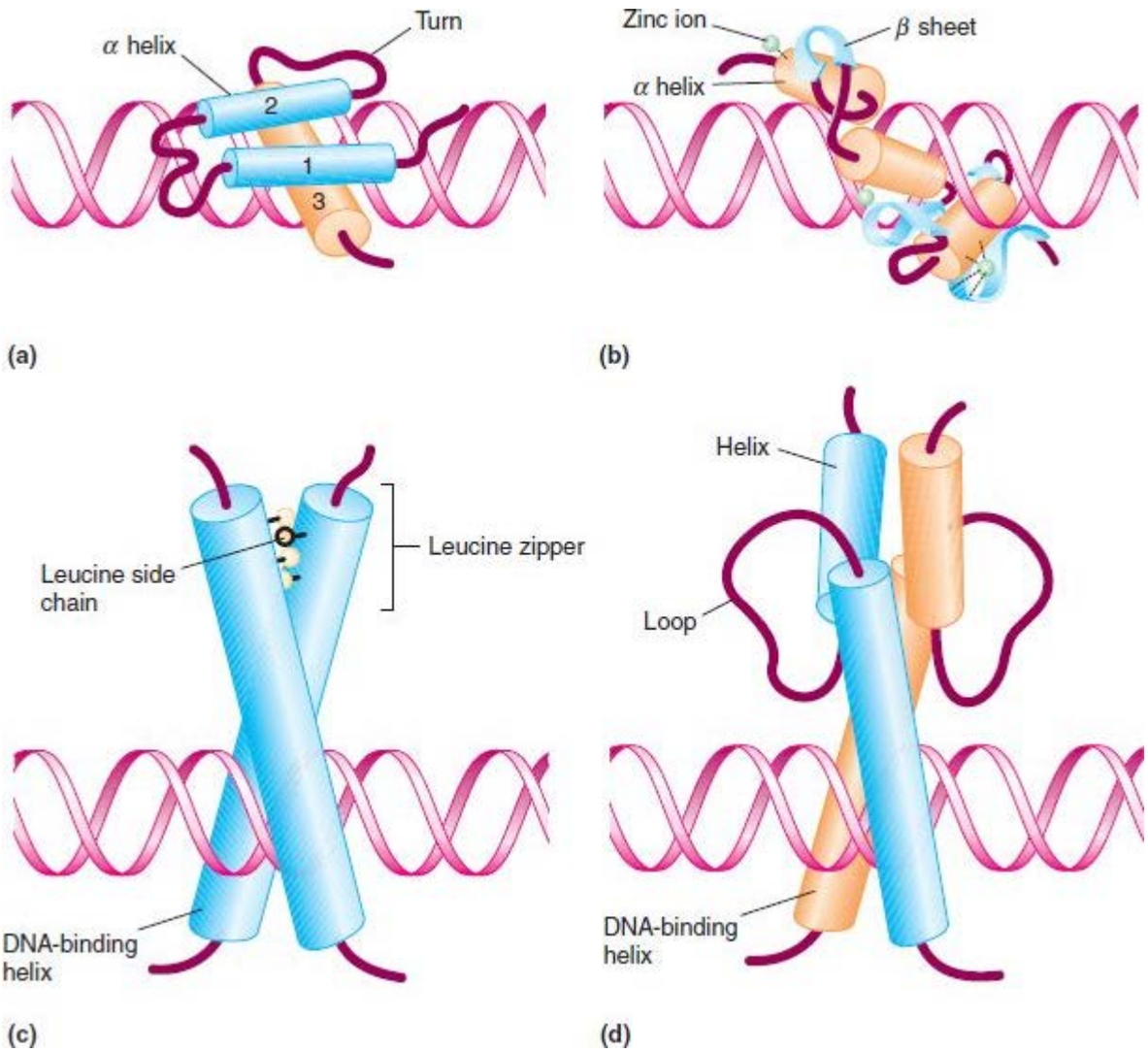


FIGURE 18.1

DNA-Protein Interactions

DNA-binding proteins contain specific structural motifs for interacting with DNA: (a) helix-turn-helix, (b) zinc fingers, (c) leucine zipper, and (d) helix-loop-helix.

Chapter 18 provides an overview of the mechanisms that living organisms use to synthesize the nucleic acids DNA and RNA that direct cellular processes. The chapter begins with a discussion of several aspects of DNA replication, repair, and **recombination** (the reassortment of DNA sequences). This is followed by descriptions of RNA synthesis and processing. Also included is an overview of several of the basic tools of biotechnology that biochemists use to investigate living

processes. The chapter ends with a section devoted to gene expression, the mechanisms cells use to produce gene products in an orderly and timely manner.

18.1 GENETIC INFORMATION: REPLICATION, REPAIR, AND RECOMBINATION

All viable living organisms possess the following features: rapid and accurate DNA synthesis and genetic stability maintained by effective DNA repair mechanisms. Paradoxically, the long-term survival of species also depends on genetic variations that allow adaptation to changing environments. In most species, these variations arise predominantly from genetic recombination, although mutation also plays a role. The following sections describe the mechanisms that prokaryotes and eukaryotes use to achieve these goals.

DNA Replication

DNA **replication** occurs before every cell division. The mechanism by which DNA copies are synthesized is similar in all living organisms. After the two strands have separated, each serves as a template for the synthesis of a complementary strand (**Figure 18.2**). (In other words, each of the two new DNA molecules contains one old strand and one new strand.) This process, referred to as **semiconservative replication**, was first demonstrated in an elegant experiment reported in 1958 by Matthew Meselson and Franklin Stahl. Refer to the online Biochemistry in Perspective essay The Meselson–Stahl Experiment for a description of the historic work proving that DNA replication is a semiconservative process.

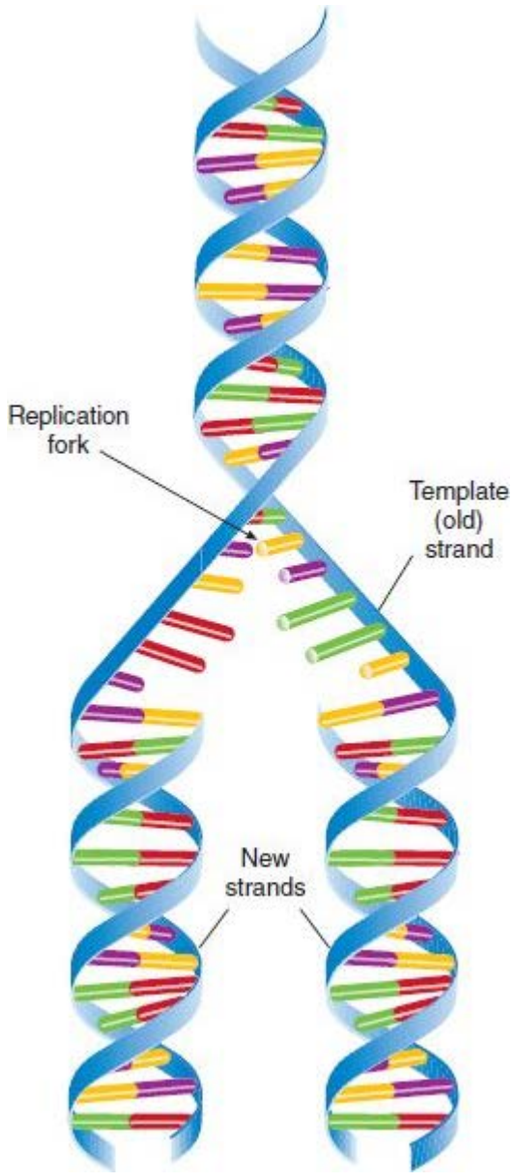


FIGURE 18.2

Semiconservative DNA Replication

As the double helix unwinds at the replication fork, each old strand serves as a template for the synthesis of a new strand.

In the years since the Meselson and Stahl experiment, many of the details of DNA replication have been discovered. Until recently, it was assumed that DNA replication machinery moved along a DNA "track" that is for the most part stationary. Recent research efforts have revealed that DNA replication occurs in specific nuclear or nucleoid compartments called **replication factories**, which are relatively stationary during the process of replication. The replication machinery within these factories performs as an energy-driven DNA pump. In prokaryotes, replication factories are attached to the cell membrane. In eukaryotes, replication factories assemble during the synthesis phase of the cell cycle (S. phase: see later, **Figure 18.12**) in association with the nuclear matrix.

DNA SYNTHESIS IN PROKARYOTES DNA replication in *E. coli* consists of several basic steps, each of which requires enzymatic activities associated with the following processes: DNA unwinding, primer synthesis, and DNA polynucleotide synthesis. DNA fragment ligation and supercoiling control are also described.

DNA Unwinding. Helicases, as the name implies, use ATP hydrolysis to mechanically break the hydrogen bonds between the annealed strands of duplex DNA. The principal helicase in *E. coli* is DnaB, a ring-shaped hexamer that opens the double helix during replication.

Primer Synthesis. The formation of short RNA segments complementary to a single-stranded DNA template called **primers** is required for the initiation of DNA replication. Primer synthesis is catalyzed by **primase**, an RNA polymerase. Primase is a 60-kDa polypeptide product of the *dnaG* gene. A multienzyme complex containing primase, helicase and several other auxiliary proteins is called the **primosome**.

DNA Synthesis. DNA is synthesized by a DNA polymerase in the 5′ → 3′ direction by forming phosphodiester linkages between nucleotides base-paired to a template strand (**Figure 18.3**). In the current model for the catalytic mechanism of DNA polymerases, illustrated in **Figure 18.4**, the 3′-hydroxyl oxygen is a nucleophile that attacks the α-phosphate of the incoming nucleotide to form a new P—O bond. With a polymerization rate of 1000 nt/s, DNA polymerase III (pol III) is the major DNA-synthesizing enzyme in prokaryotes. The pol III holoenzyme (**Figure 18.5**) is composed of at least 10 subunits (**Table 18.1**). The core polymerase is formed from three subunits: α, ε, and θ. The β-protein (also called the sliding clamp protein, or the β_2-clamp) is composed of two subunits that form a donut-shaped ring around the template DNA strand. The γ-complex is composed of τ, γ, δ, δ', χ, and ψ. Of these, τ, γ, δ, and δ' contain a motor ATPase domain that uses the energy released by ATP hydrolysis to catalyze DNA clamp loading. The γ-complex recognizes single DNA strands with primer and, acting as a **clamp loader**, transfers the β_2-clamp dimer to the core polymerase, where it forms a closed ring around the DNA strand (**Figure 18.6**). The inside diameter of the β_2-clamp is about 3.5 Å larger than that of dsDNA, large enough for hydrated DNA strands to slide through easily. The β_2-clamp promotes **processivity**; that is, it prevents frequent dissociation of the polymerase from the DNA template. As a result, POL III incorporates approximately 50 kb per binding event. Once the γ-complex is ejected in an ATP hydrolysis-driven process, the Pol III holoenzyme proceeds to replicate DNA. Note that τ subunits allow two core enzyme complexes to form a dimer, which also improves processivity. The DNA replicating machine, called the **replisome**, is composed of two copies of the pol III holoenzyme, the primosome, and DNA unwinding proteins.

E. coli also possesses four other DNA polymerases: I, II, IV, and V. DNA polymerase I (pol I), the first DNA polymerase to be discovered (Arthur Kornberg, Nobel Prize in Physiology or Medicine, 1959), is a versatile enzyme with several roles in DNA replication and repair. It has three distinctly different catalytic activities: 5′ → 3′ exonuclease activity; 5′ → 3′ template-directed polymerase activity; and 3′ → 5′ exonuclease activity. (An **exonuclease** is an enzyme that removes nucleotides from the end of a polynucleotide strand.) Pol I's 5′ → 3′ exonuclease activity immediately removes mispaired segments of DNA during replication. This same activity also removes RNA primers. Once the RNA primer is removed, the 5′ → 3′ template directed polymerase activity of pol I replaces ribonucleotides with deoxyribonucleotides. [Note that pol I is a slow enzyme (18 nt/s) with low processivity, in contrast to pol III, the major replicating enzyme.] As pol I synthesizes short DNA segments, it also uses its proofreading 3′ → 5′ exonuclease activity to ensure accuracy. Pol I also has important roles in several types of postreplication damage repair (pp. 699–704).

Pol II, even slower in its polymerase activity than pol I, also has proofreading 3′ → 5′ exonuclease and primase activities. A high-fidelity enzyme, pol II has roles in proofreading pol III mismatched bases, most often on the lagging strand (p. 693), and replication restarts after DNA damage repair. Pols IV and V are repair enzymes normally present in cells where they repair

various types of DNA damage.

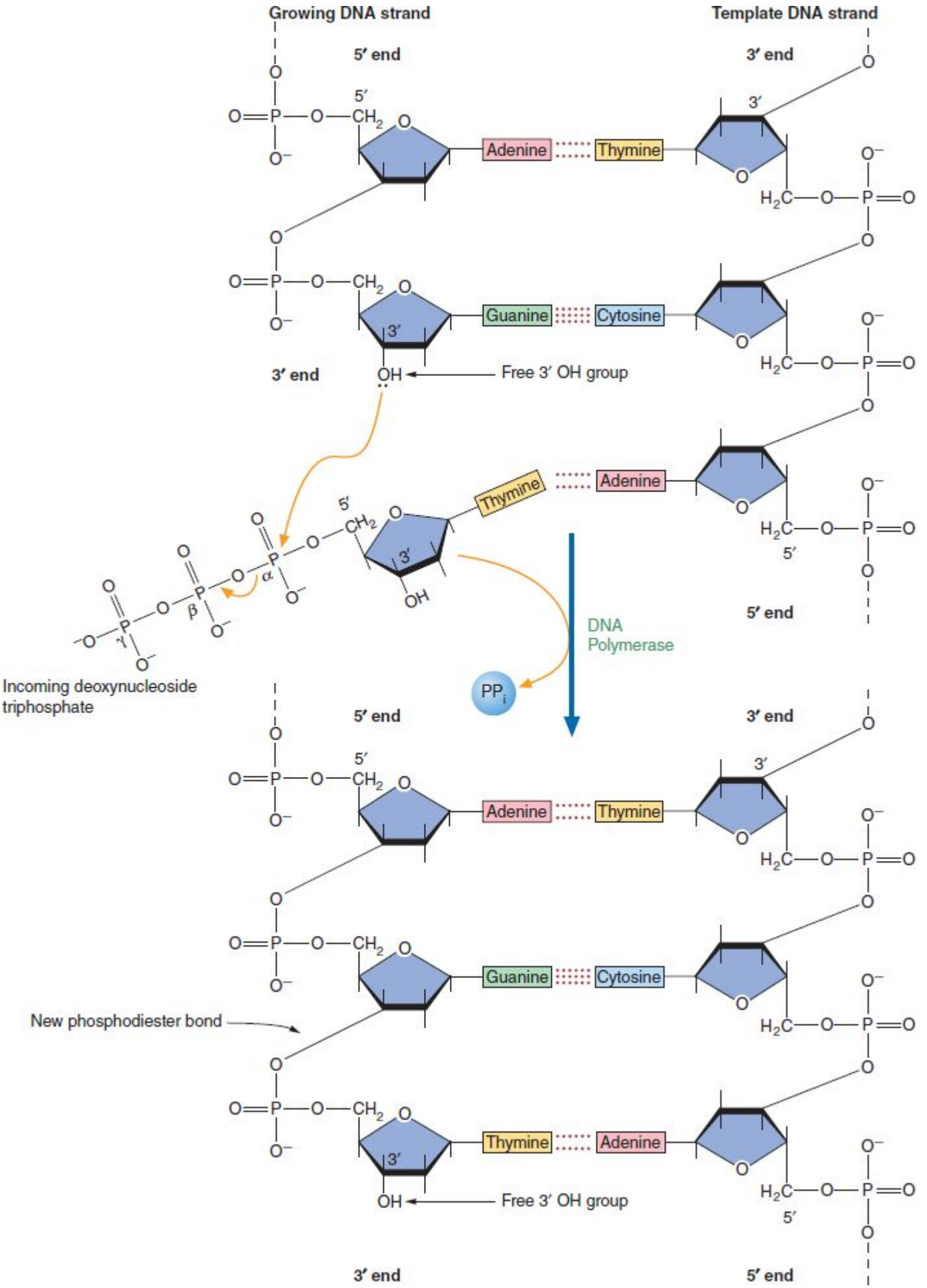

Growing DNA strand
Template DNA strand
5′ end
3′ end
Adenine
Thymine
Guanine
Cytosine
3′ end
Free 3′ OH group
Thymine
Adenine
DNA Polymerase
Incoming deoxynucleoside triphosphate
5′ end
3′ end
New phosphodiester bond
Thymine
Adenine
Free 3′ OH group
3′ end
5′ end

FIGURE 18.3

The DNA Polymerase Reaction

The essential feature of DNA synthesis is the formation of a phosphodiester linkage between a growing 5′ → 3′ DNA strand and an incoming dNTP (deoxyribonucleoside triphosphate). The bond is created by a nucleophilic attack of the 3′-hydroxyl group of the terminal residue on the α-phosphate of the incoming dNTP. Pyrophosphate (PP_i), formed as the leaving group, is then hydrolyzed. The energy released by pyrophosphate hydrolysis drives the overall process forward.

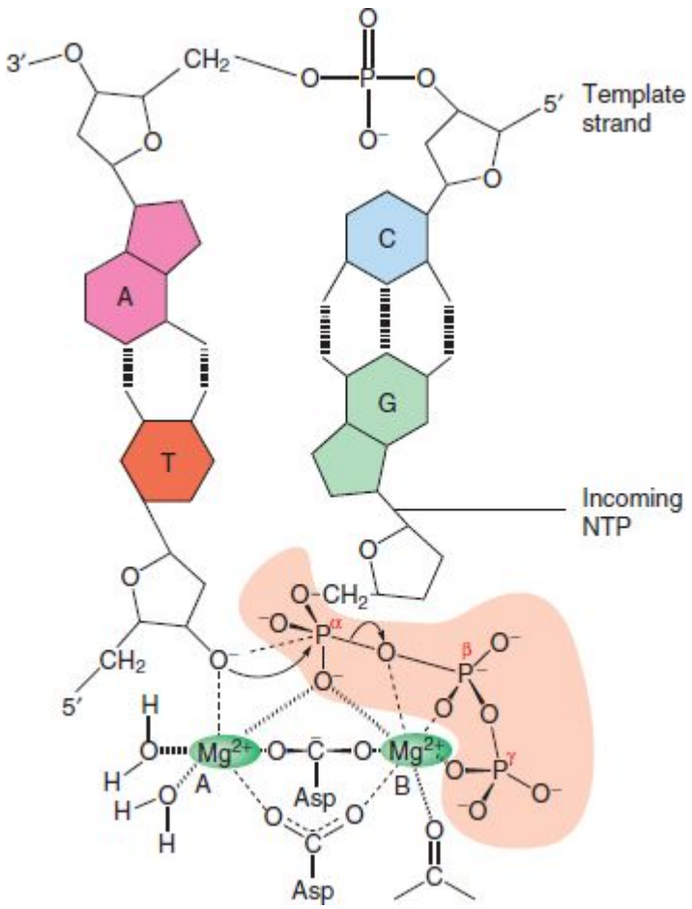


FIGURE 18.4

Mechanism of DNA Polymerases

All DNA (and RNA) polymerases apparently use the same mechanism for template-driven nucleotide polymerization: an inline nucleotidyl transfer. In this example from *E. coli* pol I, two Mg^{2+} ions (labeled A and B) are coordinated with the α-phosphate group of the incoming nucleotide (dNTP) and are themselves bridged by two aspartate side chain carboxylate groups. One ion, Mg^{2+} A, lowers the affinity of the 3′-hydroxyl oxygen for its hydrogen atom. The 3′-oxygen is a nucleophile that attacks the α-phosphate to form a new P—O bond. Both metal ions stabilize the negative charge of the transition state. By stabilizing the negatively charged pyrophosphate, Mg^{2+} B facilitates its departure.

TABLE 18.1 Subunits of DNA Polymerase III

		Subunit	Function
Holoenzyme	Core polymerase	α	$5' \rightarrow 3'$ polymerase
		ε	$3' \rightarrow 5'$ exonuclease
		θ	Assists ε in proofreading
	γ-Complex (Clamp loader)	τ	ATPase, assists in dimerization of core
		γ	ATPase
		δ and δ'	ATPases, stimulates clamp loading
		χ and ψ	Stabilize the clamp loader complex and promote replication initiation
		β	Sliding clamp as β_2

Joining DNA Fragments. Frequently, during DNA replication (as well as DNA repair and recombination processes), DNA strand segments must be joined together. An enzyme called **DNA ligase** catalyzes the formation of a covalent phosphodiester bond between the 3′-OH end of one segment and the 5′-phosphate end of an adjacent segment.

Supercoiling Control. DNA *topoisomerases* prevent tangling of DNA strands. They function ahead of the replication machinery to relieve *torque* (rotary force), which can slow down the replication process. The generation of torque is a very real problem because the double helix unwinds rapidly (as many as 50 revolutions per second during bacterial DNA replication). Topoisomerases are enzymes that change the supercoiling of the DNA (see p. 648) by breaking one or both strands, which is followed by passing the DNA through the break and rejoining the strands. The terms *topoisomerase* and *topoisomers* (DNA molecules that differ only in their degree of supercoiling) are derived from *topology*, a branch of mathematics that examines changes in the shape or position of an object such as a knotted rope that can be achieved without cutting. When appropriately controlled, supercoiling can facilitate the unzipping of DNA molecules. Type I topoisomerases produce transient single-stranded breaks in DNA; type II topoisomerases produce transient double-strand breaks. In prokaryotes, topoisomerase IV (a type II topoisomerase) facilitates the separation (decatanation) of the replication products (the linked circular chromosomes).

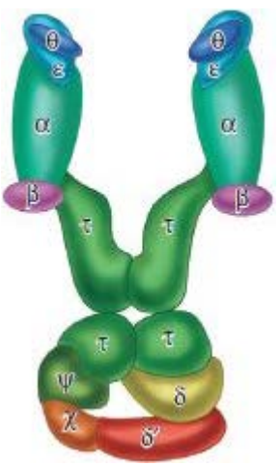


FIGURE 18.5

Diagrammatic View of *E. coli* DNA Polymerase III Holoenzyme

The pol III holoenzyme, which contains 10 different subunits, replicates DNA at high speed (1000 nt per second) and has a low error rate (one error per 10^7 bp per replication).

THE PROKARYOTIC REPLICATION PROCESS The replication of the circular *E. coli* chromosome (**Figure 18.7**) begins when there is a high ATP/ADP ratio and sufficient copies of the DNA-binding protein DnaA (53 kDa). DnaA initiates the replication process by melting the dsDNA and attracting other replication proteins to the site (**Figure 18.8**). The replication initiation site on the *E. coli* chromosome is referred to as oriC. Replication is initiated with helicases unwinding the DNA duplex followed by the assembly of the two replisomes. Replication then proceeds outward in both directions. As the two sites of active DNA synthesis (referred to as **replication forks**) move farther away from each other, a "replication eye" forms. Because an *E. coli* chromosome contains one initiation site, it is considered a single replication unit. A replication unit, or **replicon**, is a DNA molecule (or DNA segment) that contains an initiation site and appropriate regulatory sequences.

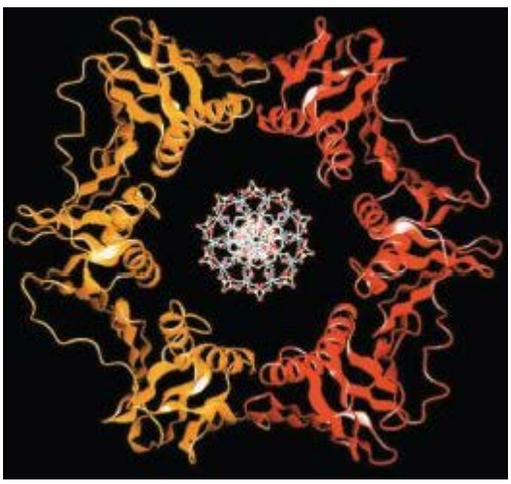


FIGURE 18.6

Cross Section of the β_2-Clamp of DNA Polymerase III

The β-protein is a dimer (shown in red and orange) that encircles the DNA and acts like a clamp.

When DNA replication was first observed experimentally (using electron microscopy and autoradiography), investigators were confronted with a paradox. The bidirectional synthesis of DNA as it appeared in their research seemed to indicate that continuous synthesis occurs in the 5′ → 3′ direction on one strand and in the 3′ → 5′ direction on the other strand. However, all the enzymes that catalyze DNA synthesis do so in the 5′ → 3′ direction only. It is now known that only one strand, referred to as the *leading strand*, is continuously synthesized in the 5′ → 3′ direction. The other strand, the *lagging strand*, is also synthesized in the 5′ → 3′ direction but in small segments (**Figure 18.9**). Reiji Okazaki and his colleagues provided the experimental evidence for discontinuous DNA synthesis. Once the RNA primer is removed from each segment and replaced by DNA, these lagging strand segments, now called **Okazaki fragments**, are covalently linked together by DNA ligase. In prokaryotes such as *E. coli*, Okazaki fragments possess approximately 1000 nucleotides.

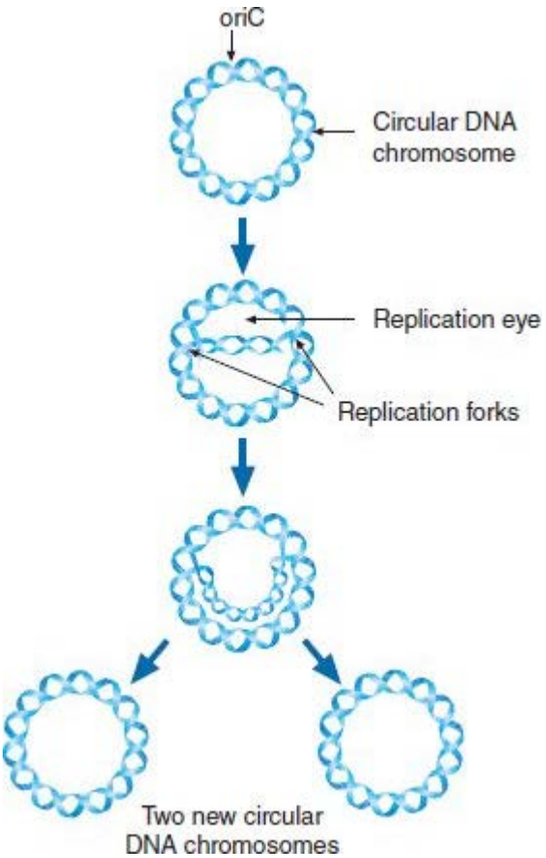


FIGURE 18.7

Replication of Prokaryotic DNA

As DNA replication of a circular chromosome proceeds, two replication forks can be observed using autoradiography. The structure that forms is called a replication eye. At the end of DNA replication, topoisomerase IV (a type of topoisomerase II) separates the two linked chromosomes.

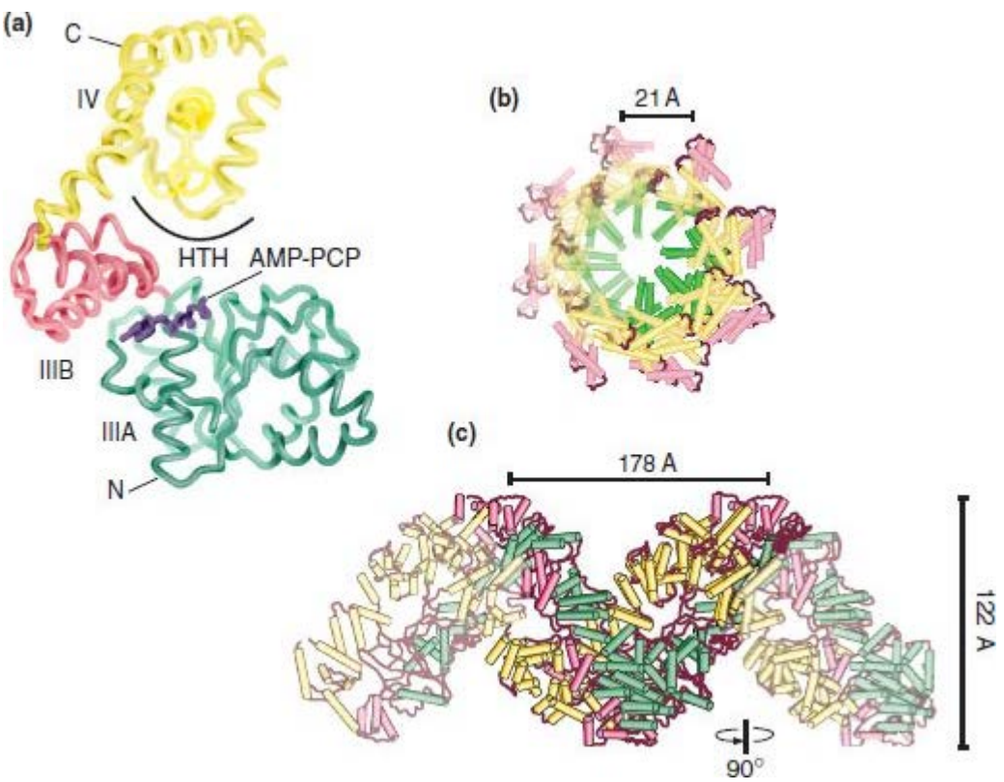


FIGURE 18.8

DnaA Structure

(a) DnaA consists of four domains: III (red and green) and IV (yellow) are illustrated. In this illustration, AMP-PCP, an analogue of ATP (designated in dark blue), is bound in the ATP-binding site of DnaA. When ATP binds to the ATP-binding site within domain IIIA of DnaA monomers, they undergo a conformational change that facilitates the formation of oligomers that bind, via domain IV, to highly conserved 9-bp DNA sequences called DnaA boxes (HTH = helix-turn-helix motif). (b) and (c) Top and side views, respectively, of DnaA oligomers. The architecture of the DnaA oligomer (an open right-handed helix) causes DNA that contains DnaA boxes to wrap around its exterior.

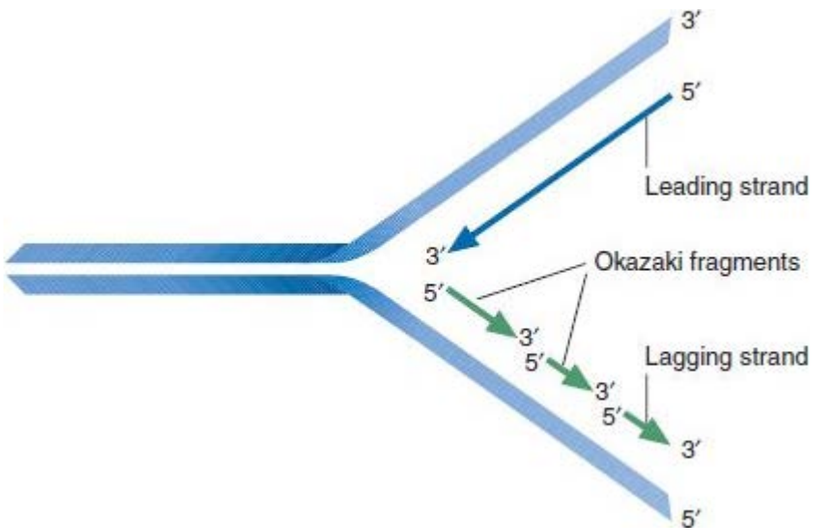


FIGURE 18.9

DNA Replication at a Replication Fork

The 5′ → 3′ synthesis of the leading strand is continuous. The lagging strand is also synthesized in the 5′ → 3′ direction but in small segments (Okazaki fragments).

Replication begins when DnaA binds to five to eight 9-bp sites, referred to as *DnaA boxes*, within the oriC sequence. Prokaryotic organisms vary in the number of DnaA boxes. *E. coli* has five DnaA boxes. The oligomerization of DnaA, which results in a nucleosome-like structure, requires ATP and the histone-like protein HU. As the DnaA-DNA complex forms, localized "melting" of the DNA duplex in a nearby region containing three 13-bp repeats causes a small segment of the double helix to open up (**Figure 18.10**). *DnaB* (a 300-kDa helicase composed of six subunits), complexed with *DnaC* (29 kDa), then enters the open oriC region. When DnaB is loaded onto the DNA, the DnaC is released. The replication fork moves forward as DnaB unwinds the helix. Topoisomerases relieve torque ahead of the replication machinery. As DNA unwinding proceeds, DnaA is displaced. The hydrolysis of bound ATP molecules causes DnaA to revert to an inactive conformation that is incapable of binding DNA. The single strands are kept apart by the binding of numerous copies of *single-stranded DNA-binding protein* (SSB). SSB, a tetramer, may also protect vulnerable ssDNA segments from attack by nucleases.

A model of DNA synthesis at a replication fork is illustrated in **Figure 18.11**. For pol III to initiate DNA synthesis, an RNA primer must be synthesized. On the leading strand, where DNA synthesis is continuous, primer formation occurs only once per replication fork. In contrast, the discontinuous synthesis on the lagging strand requires primer synthesis for each of the Okazaki fragments. The primosome travels along the lagging strand and stops and reverses direction at

intervals to synthesize a short RNA primer. Subsequently, pol III synthesizes DNA beginning at the 3′ end of the primer. As lagging strand synthesis continues, the RNA primers are removed by pol I. Pol I also synthesizes a complementary DNA segment to replace the primer. DNA ligase then joins the Okazaki fragments.

As illustrated in **Figure 18.11**, the synthesis of both the leading and the lagging strands is coupled. The tandem operation of two pol III complexes requires that one strand (the lagging strand) be looped around the replisome. When the lagging strand pol III complex completes an Okazaki fragment, it releases the DNA strand by severing its connection to the sliding clamp. This allows the primosome to move in and synthesize the next RNA primer. The lagging strand then reassociates with a new sliding clamp assembled on the newly synthesized RNA primer by γ-complex that is directly adjacent to the replication fork.

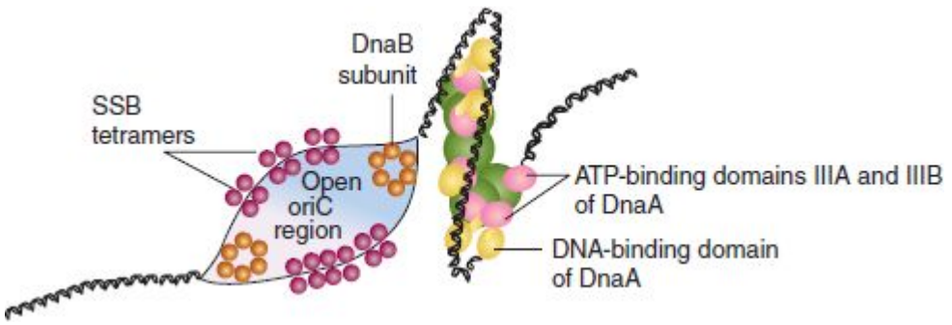


FIGURE 18.10

Replication Fork Formation

After DnaA and DnaB binding, the DnaB helicase separates the duplex DNA strands at two replication forks. Each DNA-binding domain (yellow) of DnaA binds to a 9-bp DnaA box. The binding of SSB to newly formed ssDNA prevents reassociation of the single strands. The helicase loader DnaC (not shown) facilitates the binding of DnaB helicase to DNA.

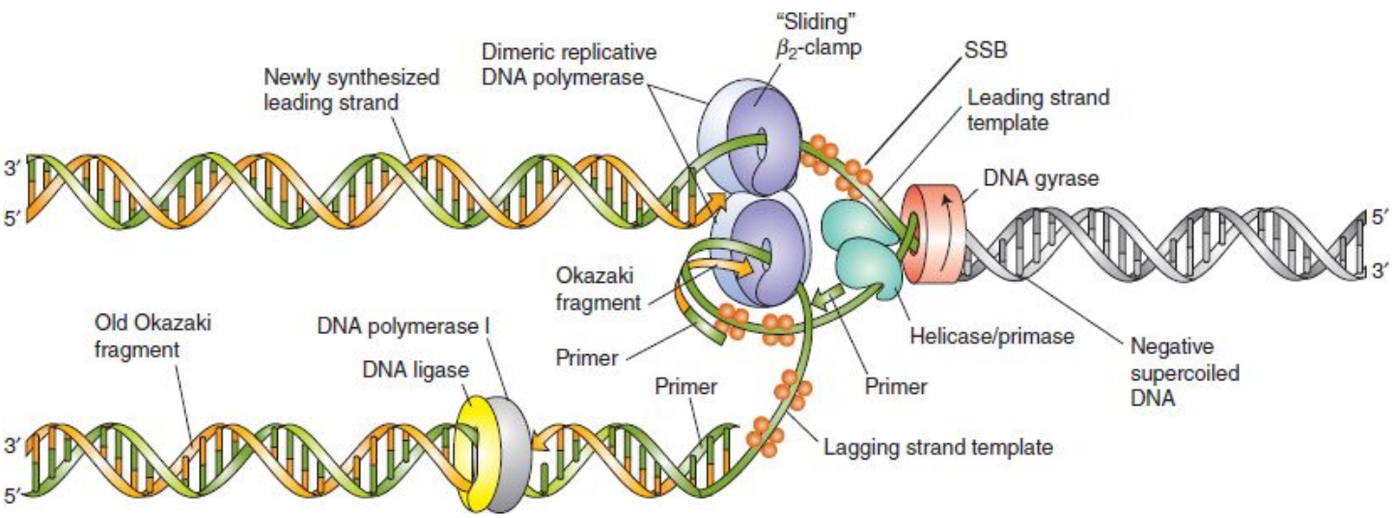


FIGURE 18.11

***E. coli* DNA Replication Model**

The primosome helicase (DnaB) separates the two DNA strands. The complex's primase activity then synthesizes an RNA primer on both the leading and lagging strands. The lagging strand is primed again approximately every 1000 bp throughout the replication process. Since the helicase activity creates positive

supercoiling the DNA duplex at the replication fork must be unwound by DNA gyrase (a type II topoisomerase). DNA gyrase uses ATP hydrolysis to relax the DNA molecule, which relieves torque, and to introduce negative supercoils into DNA just ahead of the replication fork. New strand synthesis is catalyzed by two pol III holoenzymes tethered to each other by the τ-complex (not shown). Each pol III complex is enclosed by a β_2-clamp that tethers the polymerase to the template strand. Both leading and lagging strand synthesis move in the same direction because the lagging strand is looped out before it enters the pol III holoenzyme. When the lagging strand pol III complex completes an Okazaki fragment, it releases the strand. (Because of the alternate lengthening and shortening of the lagging strand, this replication model is referred to as the trombone model.) The primosome then synthesizes a new primer. Working together, pol I and DNA ligase remove the primer and fill and seal the gaps between the Okazaki fragments. Pol III rebinds the lagging strand at a new primer and begins the synthesis of a new Okazaki fragment.

Despite the complexity of DNA replication in *E. coli* and its high processivity rate, this process is amazingly accurate: approximately one error per 10^9 to 10^{10} base pairs per replication. This low error rate is largely a consequence of the precise nature of the copying process itself (i.e., complementary base pairing). Within the active site of pol III and pol I there is a pocket that is precisely shaped to fit a nucleotide base pair in which a purine and a pyrimidine are properly aligned by hydrogen bonds and van der Waals interactions. If the nucleotide bases are mismatched, they do not fit into the pocket, and the incoming nucleotide usually leaves the site before the reaction occurs. Both pol III and pol I also proofread newly synthesized DNA. Most mispaired nucleotides are removed (by the $3' \rightarrow 5'$ exonuclease activities of pol III and pol I) and then replaced. Several postreplication repair mechanisms also contribute to the low error rate in DNA replication.

Replication ends when the replication forks meet on the other side of the circular chromosome at the termination site, the *ter* (τ) region. The ter region in *E. coli* is composed of ten 20-bp terminator sites. When ter binds to *tus*, an antihelicase that halts DNA polymerase movement, replication arrest results. The asymmetric tus-ter complex prevents a replication fork from traveling in one direction but not the other via direction-dependent inhibition of DnaB helicase (**Figure 18.12**). After replication ends, the replisomes disassemble and a type II topoisomerase separates the two daughter molecules.

DNA SYNTHESIS IN EUKARYOTES Although the principles of DNA replication in prokaryotes and eukaryotes have a great deal in common (e.g., semiconservative replication, the DNA polymerase mechanism, and bidirectional replicons), they also have significant differences. Not surprisingly, these differences appear to be related to the size and complexity of eukaryotic genomes.

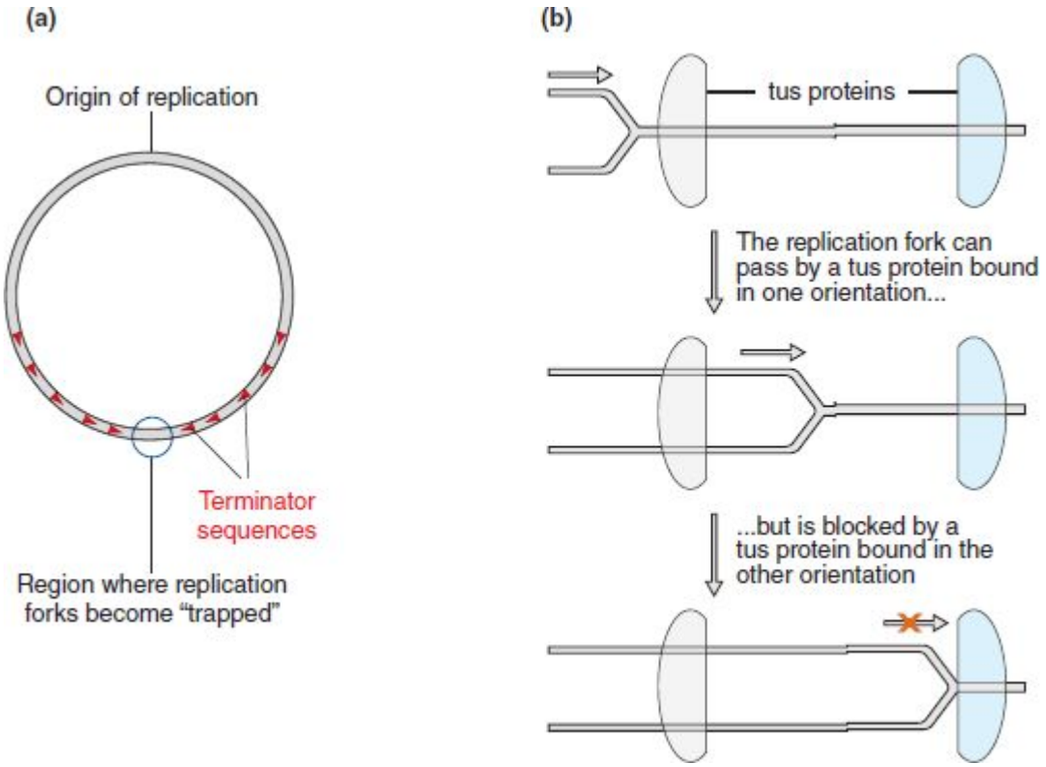


FIGURE 18.12

Role of tus in DNA Replication Termination in *E. coli*

(a) Within the ter region of the *E. coli* chromosome, there are 10 termination sequences. The arrowheads that designate these sequences indicate the direction that each sequence can be passed by a replication fork. (b) A ter sequence binds a pair of tus proteins, oriented in reverse directions. Depending on its binding orientation on a dsDNA ter segment, tus can prevent unwinding. In this diagram, a replication fork is passing by the left-handed tus. The tus bound at this site is disrupted by the DnaB helicase, so the replication fork passes through unobstructed until it encounters the second tus protein. The second tus protein (on the right), oriented on the ter sequence in the opposite direction from the left-handed tus, blocks the replication fork by inhibiting DnaB.

DNA Polymerases. There are 15 eukaryotic DNA polymerases. Of these, 3 (α, δ, and ε) are involved in nuclear DNA replication. DNA polymerase α (pol α) is a primase that initiates DNA replication by synthesizing a short 10-nt RNA segment followed by a 10- to 20-nt DNA segment. After primer synthesis on the leading strand, DNA synthesis is continued by pol ε. Pol δ is the lagging strand polymerase. Both pol δ and pol ε are highly accurate and processive polymerases, with $3' \rightarrow 5'$-exonuclease proofreading activity. Pol δ corrects errors made by pol α. Most of the other DNA polymerases have specific roles in either DNA repair [β, η (eta), ζ (zeta), θ (theta), ι (iota), and κ (kappa)] or recombination [μ (mu)]. Pol γ replicates and repairs mitochondrial DNA. Unlike the prokaryotic DNA polymerases, none of the eukaryotic DNA polymerase enzymes remove RNA primers. Instead, the enzymes Dna2 with nuclease and helicase activities and FEN1, an endonuclease, remove the primers.

Timing of Replication. In contrast to rapidly growing bacterial cells, in which replication occurs throughout most of the cell division cycle, eukaryotic replication is limited to a specific period referred to as the *S phase* (**Figure 18.13**). Eukaryotic cells produce proteins called cyclins (Section 18.3) that regulate phase transitions within the cell cycle.

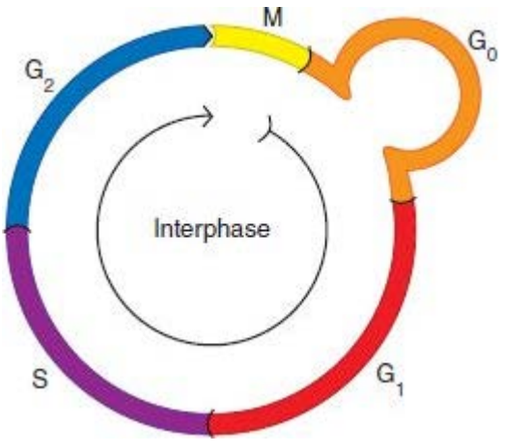


FIGURE 18.13

The Eukaryotic Cell Cycle

Interphase (the period between mitotic divisions) is divided into several phases. DNA replication occurs during the synthesis or S phase. The G_1 (first gap) phase is the time between the previous mitosis and the beginning of the next S phase. During the G_2 phase, protein synthesis increases as the cell readies itself for mitosis (M phase). After mitosis, many cells enter a resting phase (G_0).

Replication Rate. DNA replication is significantly slower in eukaryotes than in prokaryotes. The eukaryotic rate is approximately 50 nucleotides per second per replication fork. (Recall that the rate in prokaryotes is about 20 times higher.) This discrepancy is presumably a result, in part, of the complex structure of chromatin.

Replicons. Despite the relative slowness of eukaryotic DNA synthesis, the replication process is relatively brief, considering the large sizes of eukaryotic genomes. For example, on the basis of the replication rate, the replication of an average eukaryotic chromosome (approximately 150 million bp) should take more than a month to complete. Instead, this process usually is completed in several hours. Eukaryotes use multiple replicons to compress the replication of their large genomes into short periods (**Figure 18.14**). About every 40 kb along eukaryotic chromosomes, there is a site where replication machinery assembles. Humans have about 30,000 origins of replication.

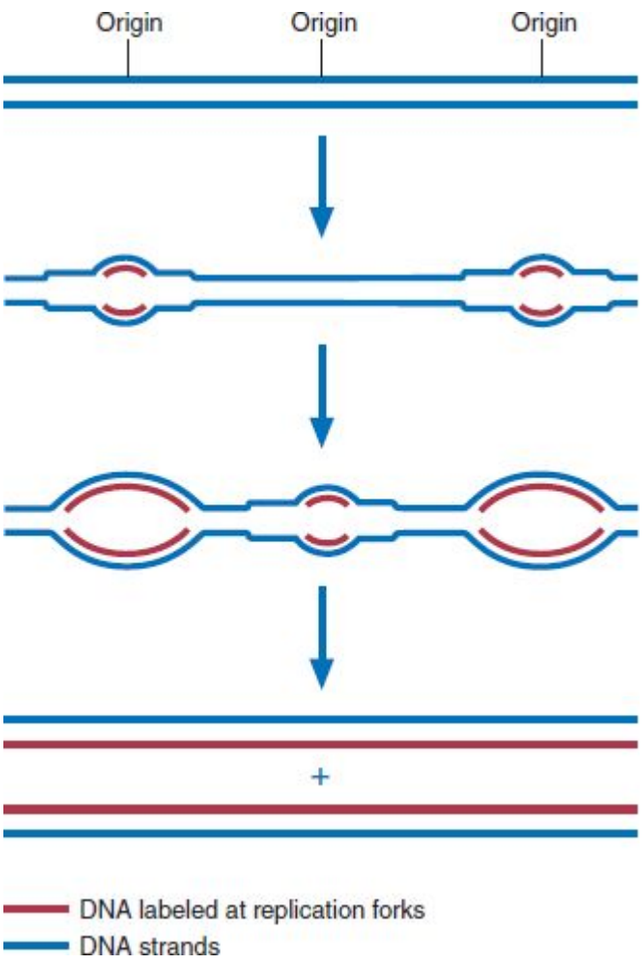


FIGURE 18.14

Multiple-Replicon Model of Eukaryotic Chromosomal DNA Replication

A short segment of a eukaryotic chromosome during replication.

Okazaki Fragments. From 100 to 200 nucleotides long, the Okazaki fragments of eukaryotes are significantly shorter than those in prokaryotes.

THE EUKARYOTIC REPLICATION PROCESS In higher eukaryotes, replication begins with the sequential assembly of the **preinitiation replication complex (preRC)** (**Figure 18.15**). The formation of this complex involves a process that starts in early G_1 phase of the cell cycle when levels of **cyclin-dependent kinase** (Cdk) (p. 740) and **cell division cycle (Cdc) proteins** are low. The process, called licensing, limits DNA replication to once per cell cycle.

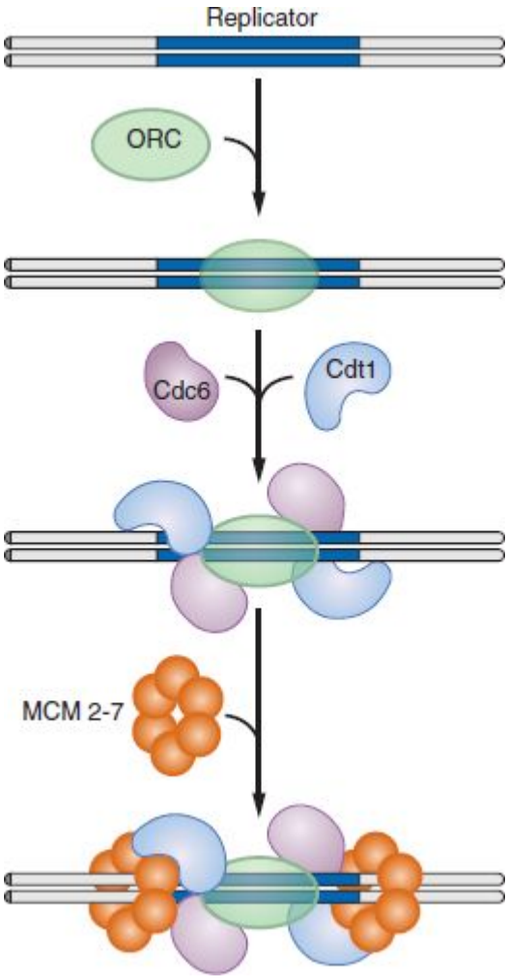


FIGURE 18.15

Formation of a Preinitiation Replication Complex

Assembly of the preRC begins (during G_1 phase of the cell cycle) when the ORC binds to the replication origin sequence, sometimes referred to as the *replicator*. ORC proceeds to recruit Cdc6 and Cdt1 (chromatin licensing and DNA replication factor 1). When the MCM complex subsequently binds, preRC formation is complete.

The components of preRC are assembled beginning when the **origin of replication complex (ORC)**, the subunits of which are analogues of DnaA, binds to the DNA initiation region or origin. Cdc6/Cdc18 and Cdt1 bind to ORC and recruit the **MCM complex** to the site. MCM (*m*ini*c*hromosome *m*aintenance) is the major DNA helicase in eukaryotes. Licensing, which occurs during preRC formation ensures that DNA synthesis at each origin of replication is initiated only once during each cell cycle. The conversion of licensed preRC to an active initiation complex requires the addition of pol α/primase, pol ε, and a number of accessory proteins, which occurs only at the onset of S phase. The cell cycle regulating kinases then phosphorylate and activate the components of the preRC. The proteins that bind to the ORC and complete the structure of the preRC are referred to as **replication licensing factors**. Once licensing has occurred, the recruitment of additional proteins of the cell cycle control machinery (e.g., cdk2-cyclin E and cdc45) allows for the DNA replication proteins to be loaded onto the replication fork and DNA synthesis to begin (**Figure 18.16**).

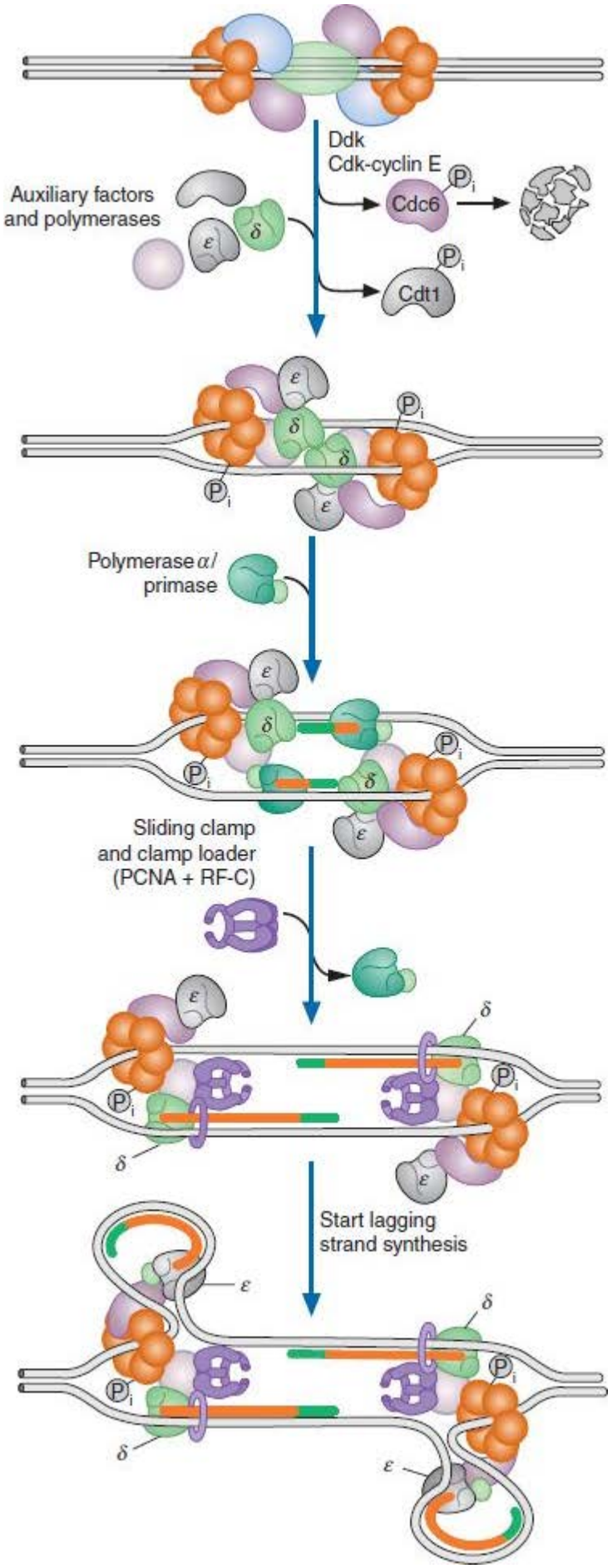

Ddk
Cdk-cyclin E
Auxiliary factors
and polymerases
Cdc6
Cdt1
Polymerase α/
primase
Sliding clamp
and clamp loader
(PCNA + RF-C)
Start lagging
strand synthesis

FIGURE 18.16

Eukaryotic Replication Fork Formation

Ddk and Cdk-cyclin E trigger replication initiation by phosphorylating several proteins. Among the results is the release of Cdc6 and Cdt1 from ORC. Upon recruitment of DNA polymerase δ and ε, the initiation complex is complete. Polymerase α/primase is then recruited. Once the RNA primers have been synthesized on the leading strand and the primer sequence is briefly extended by polymerase α, the clamp loader (RFC) binds to the sliding clamp (PCNA). The sliding clamp/clamp loader complex then transforms polymerase δ into a processive enzyme. After DNA has unwound sufficiently, lagging strand synthesis is initiated. In this illustration, DNA polymerase ε is bound to the lagging strand. The sliding clamp/clamp loader complex then binds to both the template and the newly synthesizing strands, which transforms pol δ into a processive enzyme.

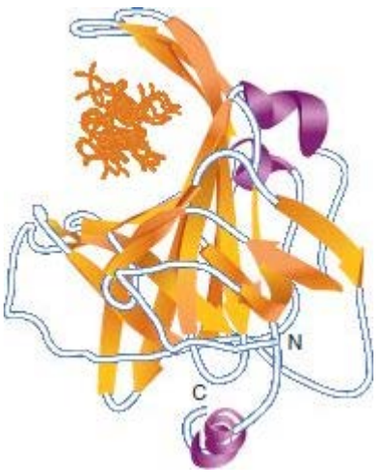


FIGURE 18.17

Replication Protein A Structure

Eukaryotes use RPA, a single-stranded DNA-binding protein, to prevent DNA strands from reannealing or being degraded by nucleases. The β-sheet in RPA forms a channel in which DNA (dark orange) binds.

When the initiation complex is active, newly phosphorylated MCM separates the DNA strands, each of which is then stabilized by **replication protein A (RPA)** (**Figure 18.17**). Replication is begun by primase, which synthesizes the RNA primers of the leading strand and each Okazaki fragment on the lagging strand. Pol α/primase extends each primer by a short DNA strand about 20 nt long. Pol α is then displaced, and pols δ and ε continue the replication process. The attachment of pol ε to the leading strand and pol δ to the lagging strand is controlled by **replication factor C (RFC)**, which is a clamp loader protein. After binding ATP, RFC then binds to PCNA, a processivity factor. The RFC/PCNA complex, which converts pols δ and ε into processive enzymes, then loads either polymerase onto the DNA, triggering ATP hydrolysis.

DNA replication of each chromosome continues until the replicons meet and fuse. When the replication machinery reaches the 3′ end of the lagging strand, there is insufficient space to synthesize a new RNA primer. Incomplete lagging strand synthesis leaves the template strand without its complementary base pairs at the end of the chromosome. Chromosomes with 3′-ssDNA overhangs are very susceptible to nuclease digestion and tend to fuse with each other, leading to chromosomal breakage during mitosis. Eukaryotic cells can compensate for this problem with **telomerase**, a ribonucleoprotein with reverse transcriptase activity, and an RNA molecule with a base sequence that is complementary to the TG-rich sequence of telomeres. Recall that telomeres (p. 659) are minisatellite sequences that occur at the ends of linear chromosomes. Telomerase uses the RNA base sequence to synthesize a single-stranded DNA to extend the 3′ strand of the telomere (**Figure 18.18**). Afterward, the normal replication machinery synthesizes a primer and a

new Okazaki fragment. The chromosome ends are then sequestered and stabilized by **telomere end-binding proteins (TEBPs)** that bind to GT-rich telomere sequences and **telomere repeat-binding factors (TRFs)** that secure the 3′ overhang (now further away from critical coding sequences) into a knot-like T-loop.

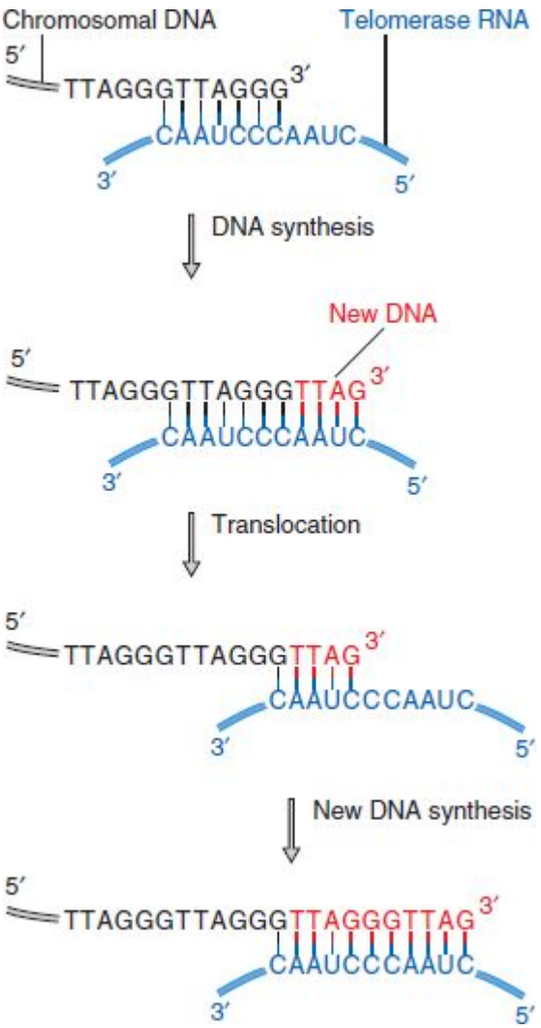


FIGURE 18.18

Telomerase-Catalyzed Extension of a Chromosome

The GT-rich, 3′-ssDNA end of a chromosome (the telomere): a portion of the RNA component of telomerase base-pairs with this region, and the reverse transcriptase activity further extends the template strand in the 5′ to 3′ direction. The telomerase translocates to the end of the new segment, and the process is repeated until the ssDNA is long enough to accommodate the replication machinery and a primer. A new Okazaki fragment is then synthesized.

In most multicellular eukaryotes, telomerase is active only in germ cells (cells that give rise to eggs and sperm). In the human body during normal aging, the telomeres of somatic cells (differentiated cells not including eggs and sperm) shorten over time. Once telomeres are reduced to a critical length, chromosomes can no longer replicate. As a result, somatic cells eventually die. It is noteworthy that the fibroblasts (connective tissue cells) of patients with Hutchinson–Guilford progeria syndrome have abnormally short telomeres. These patients age rapidly, with death occurring in the preteen years. It is also known that telomerase is overexpressed in approximately 90% of all cancers.

QUESTION 18.1

Compare the replication processes of prokaryotes and eukaryotes.

DNA Repair

Cells continuously monitor for DNA damage caused by normal metabolic activities and environmental exposures (pp. 643–45). Humans, for example, are estimated to have an average of 10,000 lesions per cell per day. And yet, only a small fraction of these lesions are passed on to daughter cells or offspring as mutations. Direct DNA sequencing analysis of humans has yielded a mutation rate of between 100 and 200 mutations per generation. Typically, low mutation rates in both prokaryotic and eukaryotic organisms are due to the evolution of several mechanisms that detect DNA lesions and then repair them. Since DNA damage is continuously occurring, cells must be able to identify it and mobilize repair processes. Some damage can be easily reversed. For example, a break in a 3′,5′-phosphodiester linkage can be readily repaired by DNA ligase. Similarly, most organisms (excluding mammals such as humans) repair pyrimidine dimers (**Figure 18.19**) with **photoreactivation repair**, a process in which the photoreactivating enzyme DNA photolyase uses light (300–500 nm) to convert the dimer into two monomers. Light energy captured by the enzyme's flavin (FAD) and pterin chromophores breaks the dimer's cyclobutane ring. (The pterin chromophore is a heterocyclic pteridine component of folic acid [p. 545]). Other repairs, however, are more complex, often requiring multienzyme complexes.

KEY CONCEPTS

- DNA is replicated by a semiconservative mechanism involving several enzymes.
- The leading strand is synthesized continuously in the 5′ → 3′ direction.
- The lagging strand is synthesized in pieces in the 5′ → 3′ direction; the pieces are then covalently linked.

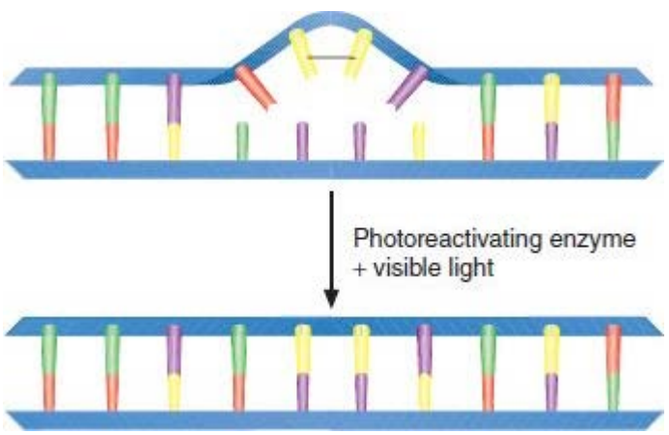


FIGURE 18.19

Photoreactivation Repair of Thymine Dimers

Light provides the energy for converting the dimer to two thymine monomers. The enzyme's light-harvesting cofactor, the pteridine ring-containing N^5,N^{10}-methenyl THF, absorbs a blue light photon and then excites a two-electron–reduced $FADH^-$ by energy transfer. The subsequent transfer of the electron by the excited $FADH^-$ breaks the dimer, after which the electron is then transferred back to the enzyme. No nucleotides are removed in this repair mechanism. Photoreactivation repair occurs widely but unevenly in bacteria, archaea,

protozoa, fungi, plants, and animals (but not humans).

There are several classes of DNA repairs. Direct repairs eliminate chemical damage to DNA (e.g., pyrimidine dimers) by reversing it. When damage is localized to one of the two DNA strands, several forms of excision repair (base excision repair, nucleotide excision repair, and mismatch repair) can be used. Double-stranded DNA breaks are repaired by either nonhomologous end joining or homologous recombination repair.

In humans and other mammals, DNA damage responses are regulated by ATM and ATR, members of a superfamily of serine-threonine kinases that also includes mTOR (p. 612). Of these, ATM (ataxia telangiectasia mutated) and ATR (*a*taxia *t*elangiectasia and *R*ad3-related protein) initiate global responses to DNA damage that activate large numbers of DNA repair and cell cycle regulatory proteins. (*Ataxia telangiectasia* is a rare human disease characterized by radiation sensitivity, a genomic instability that predisposes affected individuals to neurodegeneration and cancer.)

KEY CONCEPTS

- DNA is constantly exposed to chemical and physical processes that alter its structure.
- Each organism's survival depends on its capacity to repair this structural damage.
- Examples include photoreactivation repair, base excision repair, nucleotide excision repair, mismatch repair, nonhomologous end joining, and homologous recombination repair.

DIRECT REPAIRS Several repair mechanisms repair damage limited to a single-strand DNA segment using the complementary, undamaged strand as a template. **Base excision repair** is a mechanism that removes and then replaces individual nucleotides whose bases have undergone damage (e.g., alkylation, deamination, or oxidation). One of several enzymes called **DNA glycosylases** cleaves the N-glycosidic linkage between the damaged base and the deoxyribose component of the nucleotide (**Figure 18.20**). The resulting **apurinic** or **apyrimidinic** (AP) **sites** are resolved through the action of nucleases that remove the deoxyribose residue and several additional nucleotides. The gap is repaired by a DNA polymerase (pol I in bacteria and DNA polymerase β in mammals) and DNA ligase.

In **nucleotide excision repair** (NER), a large variety of bulky, helix-destabilizing (2–30 nt) lesions (e.g., pyrimidine dimers and benzo[*a*]pyrene adducts) are removed, and the resulting gap is filled. There are two forms of nucleotide excision repair: *global genomic repair* (GG-NER) and *transcription coupled repair* (TC-NER). They differ in the mechanisms by which DNA damage is recognized. In both forms, the excision enzymes appear to recognize the physical distortion rather than specific base sequences. In *E. coli* GG-NER (**Figure 18.21**), the excision nuclease (exinuclease), composed of Uvr A, B, and C, cuts the damaged DNA and removes a 12- to 13-nt ssDNA sequence containing the lesion. An $UvrA_2UvrB$ (A_2B) complex scans DNA for damage (e.g., a thymine dimer). Once UvrA senses a helix distortion, it partially unwinds the affected segment. UvrB further destabilizes the segment by inserting a β-hairpin domain. Next, A_2B bends the DNA, and UvrA dissociates from UvrB–DNA. UvrC then binds to UvrB, cutting the damaged DNA strand 4 or 5 nucleotides to the 3′ side of the thymine dimer. Then UvrC cuts the strand 8 nucleotides to the 5′ side. UvrD, a helicase, releases UvrC and the thymine dimer–containing oligonucleotide. The excision gap is repaired by pol I and DNA ligase. **Transcription coupled repair** occurs only on a strand actively being transcribed. The damage is recognized when the transcribing enzyme, RNA polymerase, now tightly bound to DNA, is stalled. In *E.coli*, Mfd (mutation frequency decline), a transcription–repair coupling factor, then proceeds to displace the

polymerase and recruit the $UvrA_2B$ complex that will initiate damage removal.

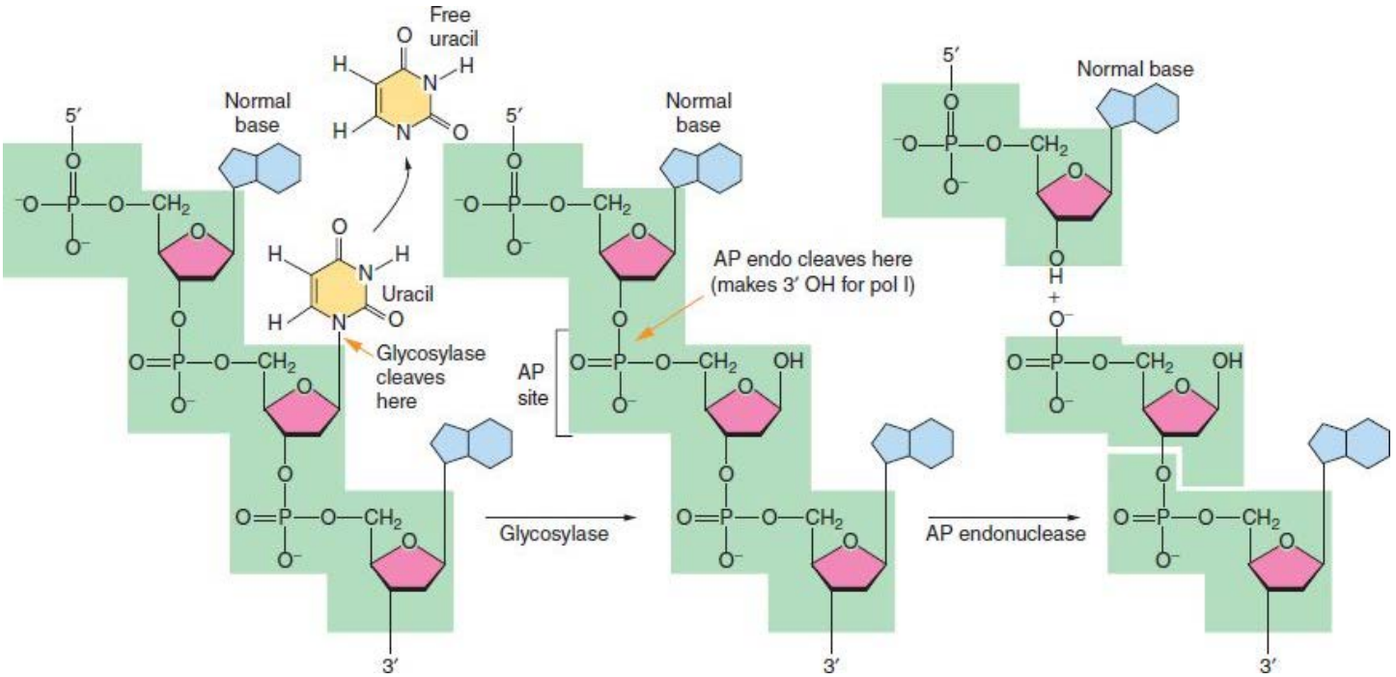


FIGURE 18.20

Base Excision Repair

DNA glycosylase hydrolyzes the N-glycosidic linkage to release the base (in this case, uracil). AP endonuclease cleaves the DNA backbone at the 5′ position of the AP (apyrimidinic) site. An endonuclease removes the AP residue and several additional nucleotides (not shown) in the 5′ → 3′ direction. A DNA polymerase then fills in the gap and DNA ligase repairs the nick.

Nucleotide excision repair of pyrimidine dimers and other helix-destabilizing lesions in humans, like that of other eukaryotes, is more complicated than the prokaryotic process, involving numerous proteins (e.g., 30 in mammals). Most of these proteins are named for their association with two diseases caused by their deficiency: *xeroderma pigmentosum* (an inherited disorder characterized by extreme sensitivity to light) and *Cockayne syndrome* (premature aging and hearing and eye abnormalities). In GG-NER, XPA, XPC, and XPE are involved in damage recognition and the subsequent recruitment of other proteins. Examples include helicases (XPB and XPD), which unwind the DNA segment containing the damage, and repair proteins such as XPF and XPG, which excise the oligonucleotide strand. A DNA polymerase uses the undamaged strand as a template to synthesize the complementary strand. GG-NER is complete when DNA ligase repairs the single-strand breaks.

In eukaryotic TC-NER, the stalled RNA polymerase serves as the damage recognition signal. The polymerase binds to CSB (an ATP-dependent chromatin remodeling enzyme) and CSA (a component of an E3 ubiquitin ligase complex). Following this DNA damage detection step, CSB triggers the recruitment of the same NER proteins used in GG-NER.

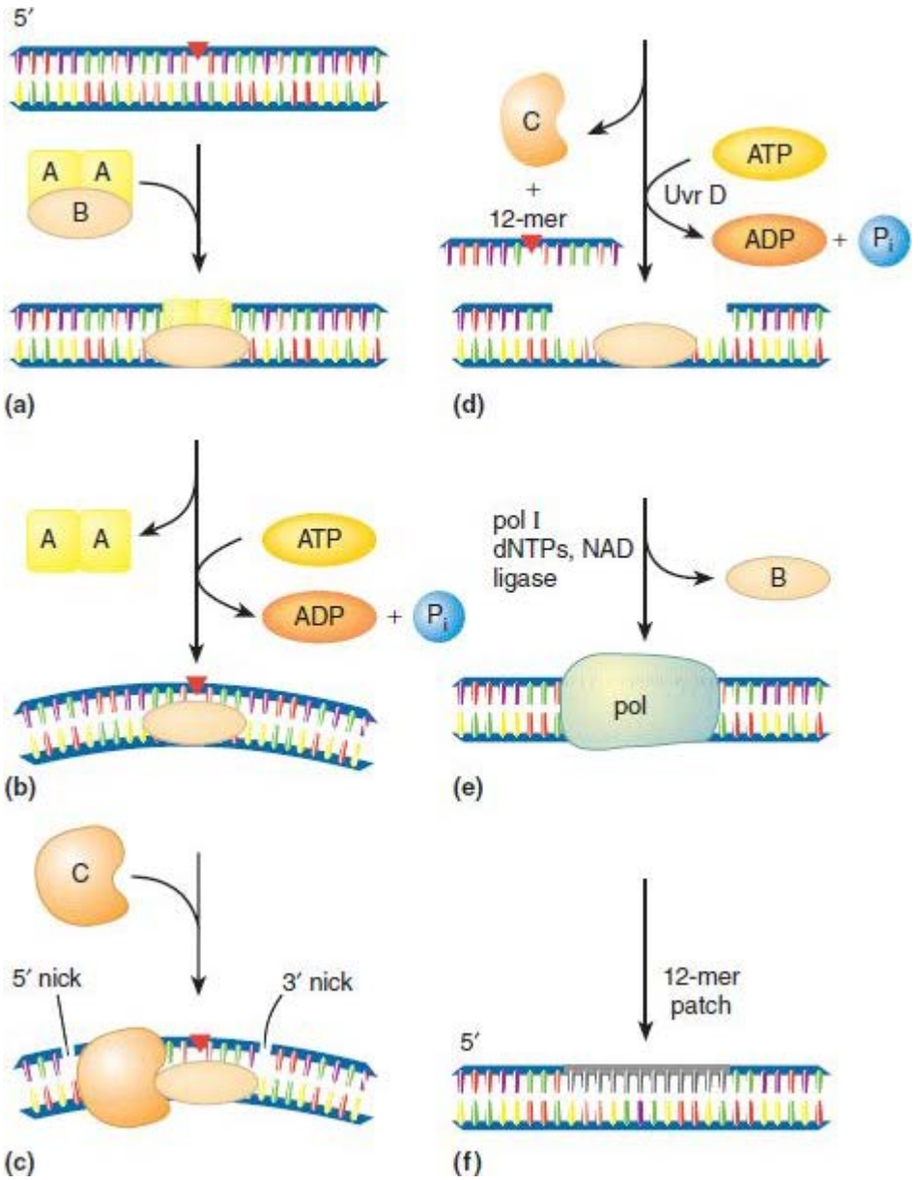


FIGURE 18.21

Excision Repair of a Thymine Dimer in *E. coli*

Uvr A, a damage recognition protein, detects helical distortion caused by DNA adducts such as thymine dimers (a). It then associates with Uvr B to form the A_2B complex. After binding to the damaged segments, A_2B forces DNA to bend. Uvr A then dissociates (b). The binding of the nuclease Uvr C to Uvr B (c) and the action of the helicase Uvr D (d) result in the excision of a 12-nucleotide DNA strand (12-mer). Then Uvr B is released (e), and the excision gap is repaired by pol I (f).

Mismatch repair (MMR) is a single-strand repair mechanism that corrects helix-distorting base mispairings that are the result of replication proofreading errors or of replication slippage. (*Replication slippage* is a type of error that occurs when repeat sequences are either skipped or copied twice, causing the formation of bubbles that require repair.) The replication process often results in deletions (skipped sequences) or insertions (recopied sequences). Some recombination errors and several forms of chemical damage (e.g., 8-oxoguanine and carcinogen adducts) can also be repaired. MMR has been estimated to cause a 100-fold increase in replication fidelity. A key feature of MMR is the capacity to distinguish between old and newly synthesized strands. In *E. coli*, this is accomplished by the methylation of both strands of parental DNA. DNA methyl transferase (Dam) methylates N-6 of adenine residues in 5′-GATC-3′ sequences, and DNA

cytosine methylase (Dcm) converts the cytosines in 5′CCTGG-3′ and 5-CCAGG-3′ sequences to 5-methycytosine. Because these sequences form palindromes, both strands are methylated equally. For a finite amount of time after replication, each daughter DNA is hemimethylated (i.e., it consists of one methylated strand and one nonmethylated strand). It is during this brief period that DNA is scanned for mismatched base pairs. The MMR system consists of three proteins: MutS, MutL, and MutH. A MutS homodimer recognizes and binds to a site on the newly synthesized strand containing a mispaired base. MutH then locates and binds the nearest GATC site. After ATP-dependent activation by MutS and MutL, MutH proceeds to nick the unmethylated strand, thus marking it for repair. A DNA helicase separates the two strands, and several exonucleases proceed to degrade the unmethylated strand from the nicked site to a few nucleotide residues past the mismatch site. Pol III then proceeds to resynthesize the missing segment using the methylated strand as a template.

The human homologues of MutS and MutL, MSH2 and MLH1, respectively, cause microsatellite instability if mutated. Most cases of hereditary nonpolyposis colorectal cancers (HNPCCs) are linked to mutations in the genes for these two proteins. The risk for other cancers (e.g., endometrial, ovarian, stomach, and small intestine) is also increased by MMR protein mutations.

DOUBLE-STRAND BREAKS DNA double-strand breaks (DSBs) are especially dangerous for cells because they can result in genome rearrangements or a lethal breakdown of chromosomes. Caused by radiation, ROS, or DNA-damaging chemicals (e.g., asbestos or cisplatin, an anticancer drug) or as a result of errors in DNA replication, DSBs are repaired by two mechanisms. One mechanism ligates two DNA ends together (nonhomologous end joining), while the other takes advantage of the base sequence information on a homologous chromosome (homologous recombination). In eukaryotes, the choice between nonhomologous end joining (NHEJ) and homologous recombination is controlled by the cell cycle (**Figure 18.13**). Although NHEJ DSB repair occurs during the entire cycle, it is most often used in the G1 phase. Homologous recombination repair is limited to late S and G2 phases.

In humans and other mammals, NHEJ is the preferred pathway for DSB repair. Both processes begin with sensing that a DSB has occurred. In mammals DSB sensing involves several molecules. Among these are ATM and ATR (p. 700), which initiate global responses to DNA damage that activate large numbers of DNA repair and cell cycle regulatory proteins, and DNA-PK (DNA-dependent protein kinase), which is involved in DSB repair.

NHEJ begins with the binding of DNA-PK to the two DNA ends. DNA-PK is a heterotrimer of DNA-PK_{CS} (protein kinase DNA-activated catalytic polypeptide) and a Ku dimer. DNA-PK_{CS} undergoes autophosphorylation and then recruits and phosphorylates a protein called artemis. *Artemis* is a nuclease that converts the two ends into ligation substrates. A trimeric complex containing DNA ligase IV and two accessory proteins finalizes the repair by ligating the broken DNA ends. Because there is no requirement for sequence homology, NHEJ is an error-prone pathway. For example, the occurrence of several DSBs in a cell can lead to the inadvertent joining of DNA ends from different chromosomes, and the loss of nucleotides at the break site can result in deletions. NHEJ proteins are used by immune system cells to generate antibody diversity (estimated at 10 billion different antibodies in humans) by recombination of V, D, and J gene segments. Homologous recombination repair, which closely resembles general recombination, is discussed in the next section.

THE SOS RESPONSE The **SOS response** is a global bacterial repair mechanism that is triggered when cells are exposed to high levels of UV light or mutagenic chemicals. SOS response enzymes

and other proteins prevent cell death when high levels of DNA damage prevent replication. Rec A is a multifunctional protein with roles in DSB repair, stalled replication fork restarts, and homologous recombination. When activated by DNA damage, RecA, a multifunctional ATPase, interacts with LexA, a repressor protein bound to promoters for SOS genes. As a result, LexA undergoes autocatalytic cleavage, thereby allowing SOS response gene expression. Early in the SOS response, cell division is halted and pol II and pol IV gene expression is upregulated, followed by expression of the pol V gene. These repair enzymes, referred to as the *translesion polymerases* because they repair damaged DNA, displace pol III at stalled replication forks. Once that damage has been repaired, a translesion polymerase is displaced and replaced with pol III. Replication continues until the next lesion is encountered. Since pol IV and pol V are error prone, the SOS response results in multiple mutations. In effect, damaged bacterial cells are exchanging multiple mutations for what otherwise would be certain death.

QUESTION 18.2

List six types of DNA repair. Explain the basic features of each.

DNA Recombination

Recombination can be defined as the rearrangement of DNA sequences by exchanging segments from different molecules. The process of recombination, which produces new combinations of genes and gene fragments, is primarily responsible for biological diversity among living organisms. More important, the large number of variations made possible by recombination gives species opportunities to adapt to changing environments. In other words, recombination is a principal source of the variations that make evolution possible. Recombination is also used to repair dsDNA breaks.

There are two forms of recombination: general and site-specific. In eukaryotes, **general** (or **homologous**) **recombination-mediated sequence exchanges** between homologous DNA molecules occur most notably during meiosis. (Meiosis is the form of eukaryotic cell division in which haploid gametes are produced.) In **site-specific recombination**, the exchange of sequences from different molecules requires only short regions of DNA homology. These regions are flanked by extensive nonhomologous sequences. Site-specific recombinations, which depend more on protein-DNA interactions than on sequence homology, occur throughout nature. For example, this mechanism is used by bacteriophage to integrate its genome into the *E. coli* chromosome. In eukaryotes, site-specific recombination is responsible for a wide variety of developmentally controlled gene rearrangements that are partially responsible for cell differentiation in complex multicellular organisms. One of the most interesting examples of gene rearrangement is the generation of antibody diversity in mammals. In the variation of site-specific recombination referred to as **transposition**, DNA sequences called **transposable elements** or **transposons** (p. 659) move from one chromosome or chromosomal region to another.

GENERAL RECOMBINATION General recombination occurs in all living organisms, but has been investigated primarily in *E. coli* and fungi such as *S. cerevisiae* and *Aspergillus nidulans*. In addition to generating genetic diversity, general recombination is also an important DNA damage repair process. Several models explaining recombination have been proposed. Examples include the Holliday, Meselson–Radding, double-strand break repair, and synthesis-dependent strand annealing models.

The first model explaining general recombination was based on Robin Holliday's work with fungi. The Holliday model (**Figure 18.22**) involves the following steps:

1. Two homologous DNA molecules become paired.
2. Two of the DNA strands, one in each molecule, are nicked (cleaved) at identical locations.
3. The two nicked strand segments cross over, thus forming a Holliday intermediate.
4. DNA ligase seals the cut ends.
5. Branch migration caused by base-pairing exchange leads to the transfer of a segment of DNA from one homologue to the other.
6. A second series of DNA strand cuts occurs.
7. DNA polymerase fills any gaps, and DNA ligase seals the cut strands.

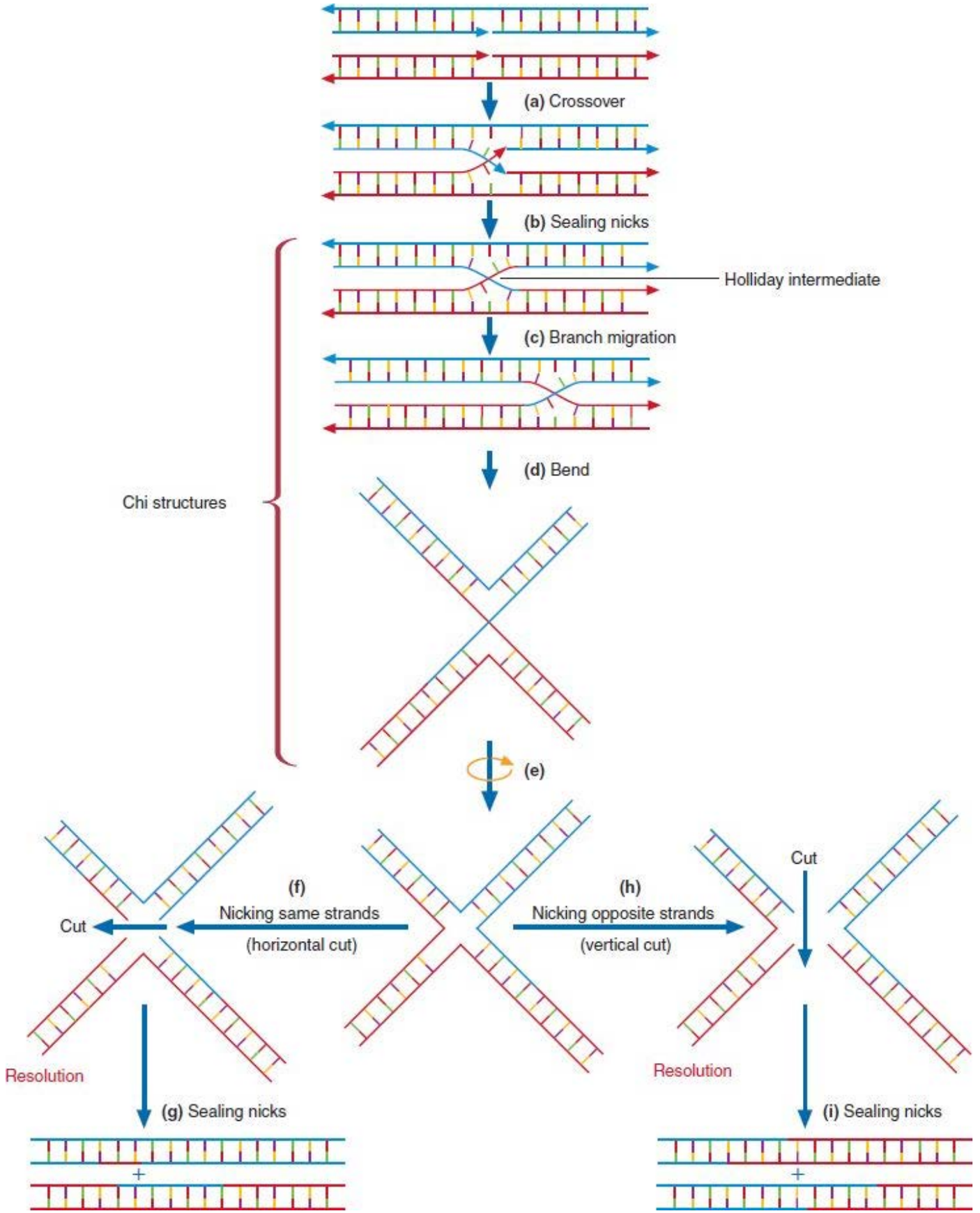


FIGURE 18.22

General Recombination: The Holliday Model

Once one strand in each duplex is nicked, each broken strand invades the other duplex (a). Covalent bonds are formed (b), cross-linking the two duplexes. Branch migration then occurs (c). The bending of the chi structure in (d) and (e) makes later events easier to understand. In (f) the same original strands are nicked. The resulting

hetero-duplex (g) contains a patch. Nicking the opposite strands (h) results in the formation (i) of a spliced heteroduplex.

The **Meselson–Radding model** is the result of efforts to account for several laboratory observations not explained by the Holliday model. Among these is the fact that recombination sometimes results in only one of the homologous DNA molecules having a recombinant strand. According to the Meselson–Radding model (**Figure 18.23**) recombination occurs as follows:

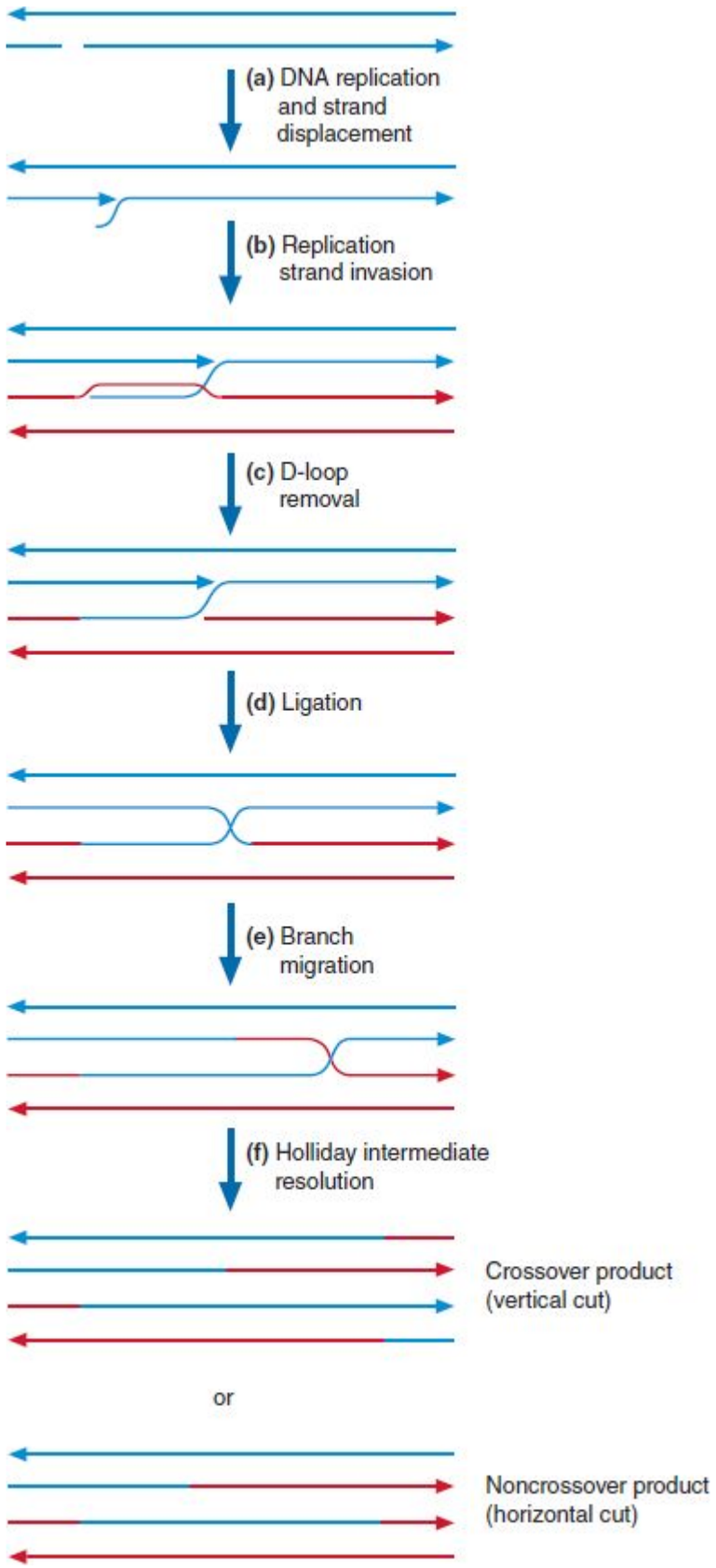


FIGURE 18.23

General Recombination: The Meselson–Radding Model

A single-strand nick on a DNA molecule initiates the recombination process. As DNA polymerase proceeds to synthesize a strand extension at the 3′-end, the DNA strand on the other side of the nick (i.e., the 5′-end) is displaced (a). It is this 5′-tail that invades into a homologous DNA molecule (b) to form a D-loop. The D-loop is subsequently removed (c) by a nuclease. A Holliday intermediate forms as the two free ends are ligated (d).

Branch migration may occur (e). As in the Holliday model (**Figure 18.22 e–i**) the Holliday intermediate resolves (f) into either crossover or noncrossover products.

1. One strand of one of the two homologous DNA molecules is nicked.
2. The extension of the newly created 3′-end by DNA polymerase causes the displacement of the strand on the other side of the nick. As the growing strand becomes longer, the displaced strand invades the double helix of a homologous segment of the second chromosome to form a D-loop structure.
3. The D-loop is cleaved, and the invading strand is ligated to the newly created 3′-end of the homologous strand.
4. The 3′-end of the newly synthesized strand and the 5′-end of a homologous strand are ligated to form a Holliday intermediate.
5. Branch migration may occur.
6. Strand nicks and resolution of the Holliday junction (refer to **Figure 18.22f–i**) result in either a crossover product (vertical cut) or a noncrossover product (horizontal cut).

In eukaryotes, general recombination has a featured role in *meiosis*, a type of cell division necessary for sexual reproduction. In a mechanism that accurately aligns homologous maternal and paternal chromosomes, the general recombination machinery facilitates the exchange of chromosomal segments that creates genetic diversity. The **double-strand break repair model** (DSBR) (**Figure 18.24a–d**) explains many of the features of meiotic recombination not accounted for by the Holliday and Meselson–Radding models. DSBR mechanisms also protect prokaryotic and eukaryotic cells from DSBs (p. 703) caused by mutagenic factors such as ionizing radiation and ROS. The major steps in DSBR are as follows:

1. An endonuclease induces a DSB in one of a pair of homologous DNA molecules.
2. Exonucleases proceed to degrade the 5′-ends, leaving 3′-tails.
3. One of the 3′-tails invades a homologous DNA molecule and after a successful homology search forms a D-loop. Note that Rad51, the eukaryotic homologue of RecA (p. 703), is involved in the homologous sequence search. It also catalyzes homologous pairing and strand exchange.
4. DNA polymerase extends both the invading and the noninvading 3′-tails.
5. Branch migration combined with the ligation between the invading 3′-tail by the 5′-end on the other side of the break results in the formation of a double Holliday junction.
6. Resolution of the two Holliday structures can give crossover or noncrossover products.

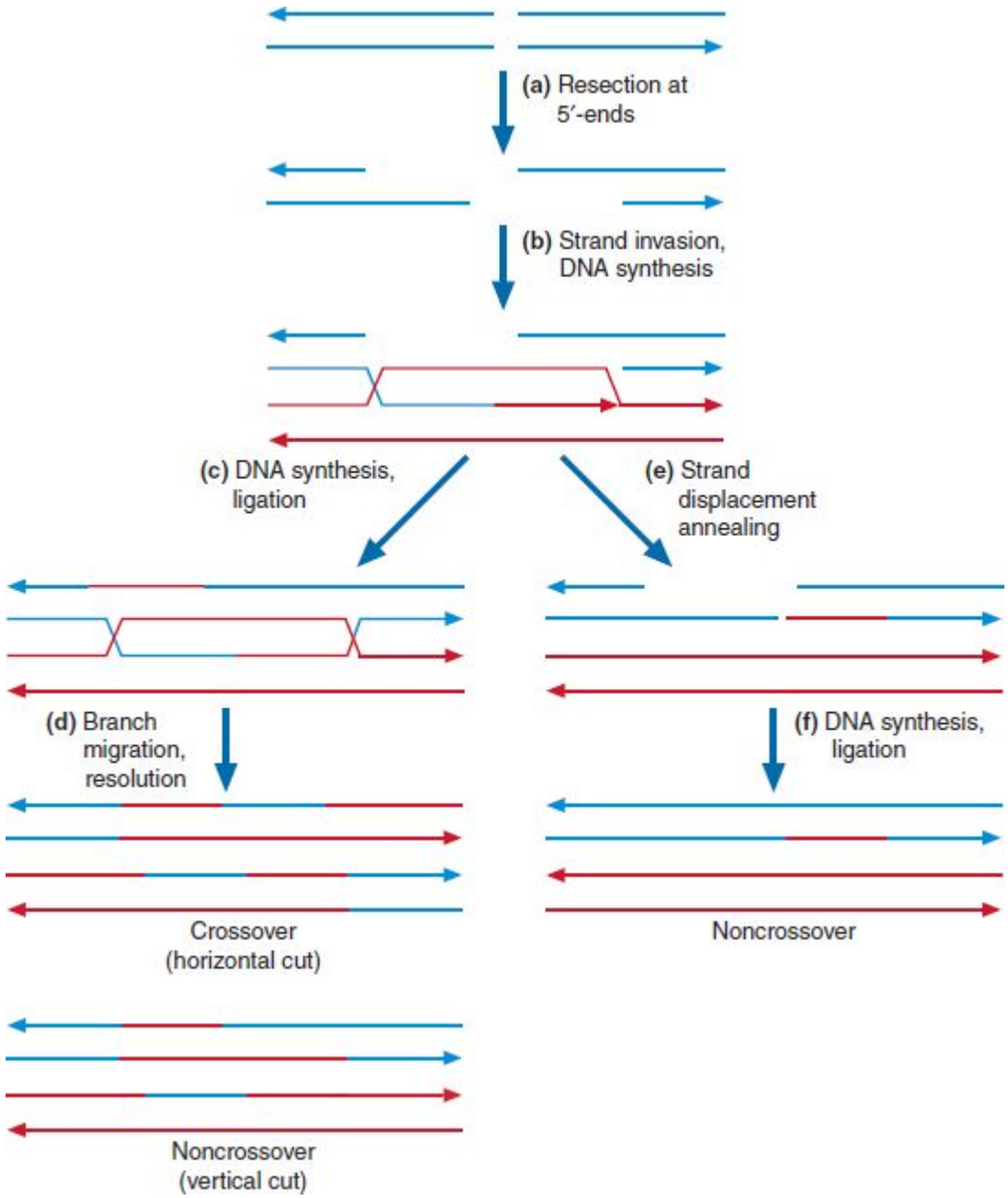


FIGURE 18.24

Double-Strand Break Repair (DSBR) and Synthesis-Dependent Strand Annealing (SDSA) Models of General Recombination

DSBR begins with a double-strand cut of one of the homologous DNA molecules. An exonuclease then resects (trims) the two 5′-ends (a) to yield two 3′-tails. (b) One of the 3′-tails invades an equivalent segment of the second homologous DNA molecule to form a D-loop. As the 3′-end of this tail is extended by DNA synthesis, using the homologous DNA strand as a template, a Holliday intermediate forms. (c) The other 3′-tail is extended by DNA synthesis using the complementary strand segment in the D-loop as a template. The DNA synthesis phase of the process ends with the action of DNA ligase. (d) The double Holliday junction undergoes branch migration and is then resolved by an endonuclease into crossover or noncrossover products. In SDSA after 5′-end resection (a), strand invasion, and DNA synthesis (b), a second Holliday intermediate does not form. Instead, (e) the invading strand is displaced, thus allowing a reannealing process that provides a template for the DNA synthesis repair (f) of the other broken strand. SDSA ends with ligation of the single-stranded gaps.

DSBR also occurs in mitosis, usually without crossover products. The **synthesis-dependent strand-annealing model** (SDSA) explains how this phenomenon occurs. SDSA diverges from DSBR when the strand in the D-loop is displaced (**Figure 18.24e** and **f**). The invading strand

proceeds to anneal with the complementary single strand of the other DSB end. The process ends with further DNA synthesis, followed by ligation.

PROKARYOTIC RECOMBINATION Recombination in *E. coli* is initiated when RecBCD, an enzyme complex that possesses exonuclease and helicase activities, encounters a Chi (crossover hotspot instigator) sequence site. After binding to a DNA molecule, RecBCD cleaves one of the strands and proceeds to unwind the double helix until it reaches 5′-GCTGGTGG-3′ (the Chi site), a sequence that occurs frequently in *E. coli* DNA. Strand exchange begins when a nucleoprotein filament formed when monomers of RecA, an ATPase with recombinase activity, coats one of the strands. Powered by ATP hydrolysis, the RecA-coated strand segment then searches for homology in nearby dsDNA. Once a homologous segment is located, DNA synthesis causes strand displacement, which is followed by D-loop cleavage, strand capture, and Holliday junction formation. Subsequent branch migration (**Figure 18.25**) is initiated as RuvA recognizes and binds to the Holliday junction. Two copies of RuvB, a hexamer with ATPase and helicase activities, then form a ring on either side of the junction. The RuvAB complex catalyzes branch migration. This molecular machine separates, rotates, and pulls the strands in the two sets of helices, even after RecA dissociates. The migration ends when a specific sequence [5′-(A or T)TT(G or C)-3′] is reached. As RuvAB detaches, two RuvC proteins bind to the junction. Recombination ends as RuvC cleaves the crossover strands and the Holliday structure resolves itself to form two separate double-helical DNA molecules.

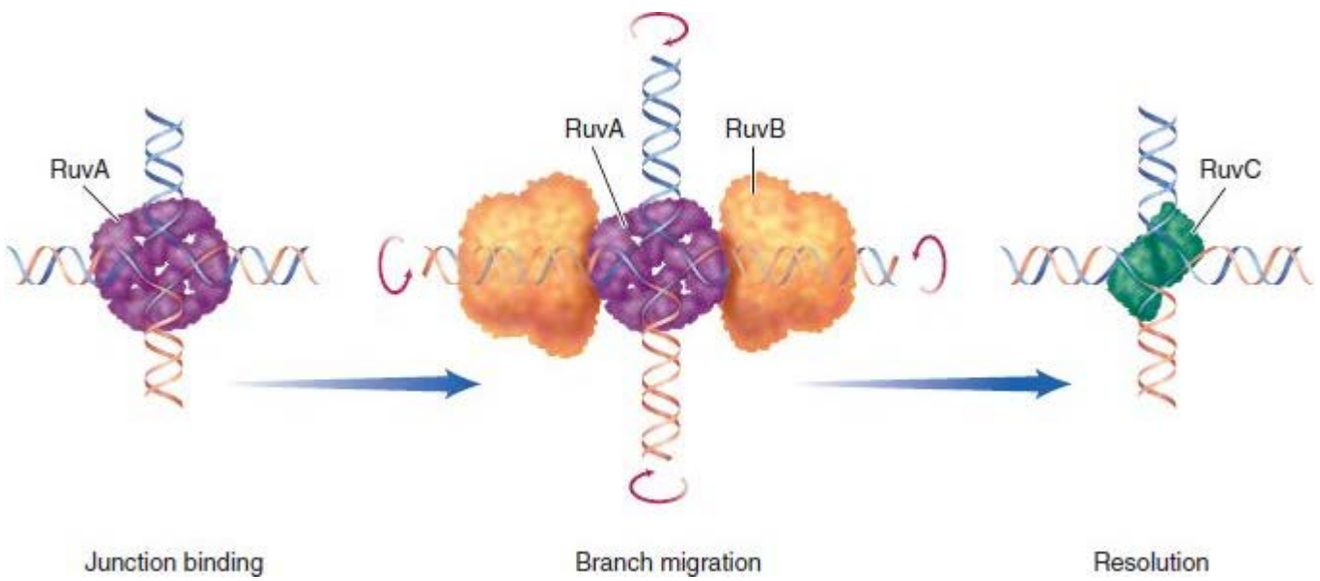


FIGURE 18.25

Model of the Association of Ruv Proteins with a Holliday Junction

RuvA, a tetramer, first binds to the Holliday junction point. Two hexameric RuvB rings then form on both sides of the DNA/RuvA complex, with the DNA passing through the rings. Branch migration occurs as ATP hydrolysis drives the two RuvB rings to rotate the DNA helices in opposite directions. After branch migration, RuvA and RuvB detach and two RuvC proteins bind to the junction. RuvC, a nuclease, proceeds to cut the crossover strands, thus resolving the Holliday structure.

In bacteria, general recombination appears to be involved in several forms of intermicrobial DNA transfer: transformation, transduction, and conjugation. In **transformation**, naked DNA fragments enter a bacterial cell through a small opening in the cell wall and are introduced into the

bacterial genome. (See the online Biochemistry in Perspective essay A Short History of DNA Research: The Early Years, cited in Chapter 17, for a description of Fred Griffith's groundbreaking experiment.) **Transduction** occurs when bacteriophage inadvertently carry bacterial DNA to a recipient cell. After a suitable recombination, the cell uses the transduced DNA. **Conjugation** is an unconventional sexual mating that involves a donor cell and a recipient cell. The donor cell possesses a specialized plasmid (p. 44) that allows it to synthesize a sex pilus, a filamentous appendage that functions in a DNA exchange process. After the pilus has attached to the surface of the recipient cell, a fragment of the donor's genetic material is transferred. The transferred DNA segment can be integrated into the recipient's chromosome by recombination, or it may exist outside it in plasmid form.

EUKARYOTIC RECOMBINATION In eukaryotes, general recombination occurs during the first phase of meiosis to ensure accurate homologous chromosome pairing and *crossing over*, the mechanism for introducing genetic variation. DNA damage repair of DSBs using homologous DNA molecules occurs during the S and G2 phases of the cell cycle (p. 696) when newly replicated DNA is available. (In the other phases of the cycle, DSBs are repaired by NHEJ.)

The basic mechanism of eukaryotic homologous recombination repair of DSBs is believed to be similar to the process in prokaryotes. However, because of vastly more complex genomes, the number of proteins involved in eukaryotic recombination is substantially larger. In mammals, the MRN complex (Mre11, Rad50, and Nbs1) creates a scaffold that stabilizes the DNA ends at DSBs caused by exogenous damage. MRN also catalyzes 5′-end resection that produces the 3′-ssDNA ends required for strand invasion. When exogenous agents cause DSBs, MRN recruits and activates ATM (p. 700), which in turn activates several DNA repair proteins and cell cycle regulatory proteins. Examples of ATM-activated DNA repair proteins include Rad51, BRCA1, and BRCA2. After 5′-end resection, RPA (p. 698) binds to and stabilizes the now exposed single 3′ DNA strand ends. With the aid of BRCA2 (breast cancer susceptibility protein type 2), the recombinase RAD51 displaces RPA, binds to the DNA strand ends to form a nucleoprotein filament, and proceeds to catalyze strand invasion. Once a homologous strand has been identified in an undamaged chromatid, the RAD51/ssDNA complex invades the homologous dsDNA sequence (the donor duplex) to form a *heteroduplex* structure that generates the D-loop (p. 706). (A *chromatid* is a newly copied chromosome linked to the damaged chromosome by a centromere and cohesin complex rings.) Several accessory proteins facilitate RAD51 removal. The resolution of the heteroduplex (**Figure 18.24**), DNA synthesis by pol δ, and annealing by RAD52 yield two undamaged chromatids. BRCA1 functions in a number of repair-associated pathways, including cell cycle regulation and chromatin remodeling. BRCA1, linked to BRCA2, interacts with Rad51 during DSB repair. Note that BRCA1 also promotes genome stability by repressing satellite DNA transcription.

SITE-SPECIFIC RECOMBINATION AND TRANSPOSITION Site-specific recombination relies on short segments of homologous DNA called **attachment (att) sites** or **insertional (IS) elements**. Recombination at these sites can lead to insertions, inversions, deletions, and translocations that may benefit or harm the cell. Examples of site-specific recombination include insertion of viral DNA into a host cell genome, acquisition of antibiotic resistance by bacteria, phenotypic variation in plants, and antibody maturation in mammals. A simple case of insertion is illustrated by the integration of bacteriophage λ DNA into the *E. coli* chromosome (**Figure 18.26**). The process requires homologous att sites in the phage and bacterial genomes, a viral recombinase called λ integrase, and a bacterial gene product, the integration host factor. The mechanism with Holliday junction resolution results in the insertion of the λ genome into the bacterial

chromosome.

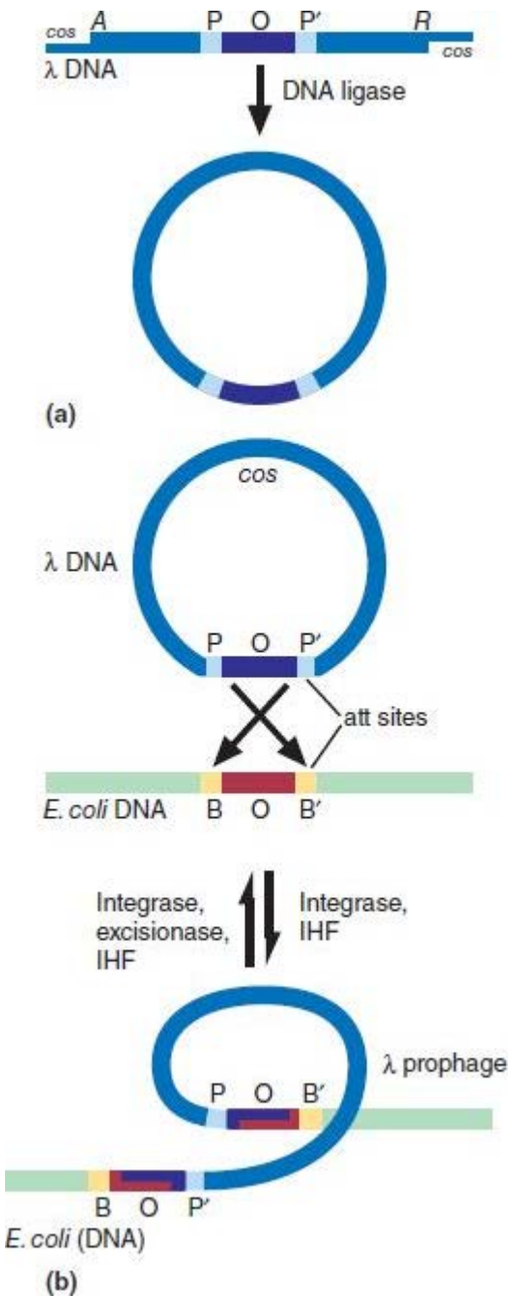


FIGURE 18.26

Insertion of the Bacteriophage λ Genome into the *E. coli* Chromosome

(a) The *λ* DNA circularizes as the single-stranded *cos* sequences anneal. The viral att site sequence consists of POP′ where O is a 15-bp sequence at the center. The bacterial att sequence BOB′ contains an identical O sequence. (b) Insertion occurs through site-specific recombination between short homologous O sequences within the phage and bacterial att sites.

QUESTION 18.3

Bacterial conjugation has medical consequences. For example, certain plasmids contain genes that code for toxins. The causative agent of a deadly form of food poisoning, *E. coli* 0157, synthesizes a toxin that causes massive bloody diarrhea and kidney failure. This toxin is now believed to have originated in *Shigella*, another bacterium that causes dysentery. Similarly, the

growing problem of antibiotic resistance is partly attributable to the spread of antibiotic-resistant genes among bacterial populations. Antibiotic resistance develops because antibiotics are overused in medical practice and in livestock feeds. Suggest a mechanism by which this extensive use promotes antibiotic resistance. [*Hint*: The high-level use of antibiotics acts as a selection pressure.]

Bacterial Conjugation

The IS elements of simple prokaryotic transposons consist of the gene for the transposition enzyme transposase, flanked by short inverted terminal repeats that define the boundary of the transposon (**Figure 18.27**). (An **inverted repeat** is a sequence that is the reverse of another downstream sequence. Inverted repeats without intervening sequences constitute a **palindrome**, a DNA sequence whose complement reads the same in reverse. See p. 648 for an example of an inverted repeat.) More complicated bacterial transposons, called **composite transposons**, are made up of two separate transposons that are linked by the DNA between them. The two transposons (IS elements) become linked when inactivating mutations (e.g., in the inner inverted repeats or one of the IS transposase genes) prevent independent jumping of the two IS elements. The presence of DNA sequences between the two IS elements that have useful properties (e.g., antibiotic resistance genes) also promotes retention of a composite transposon. Two transposition mechanisms have been observed, nonreplicative and replicative.

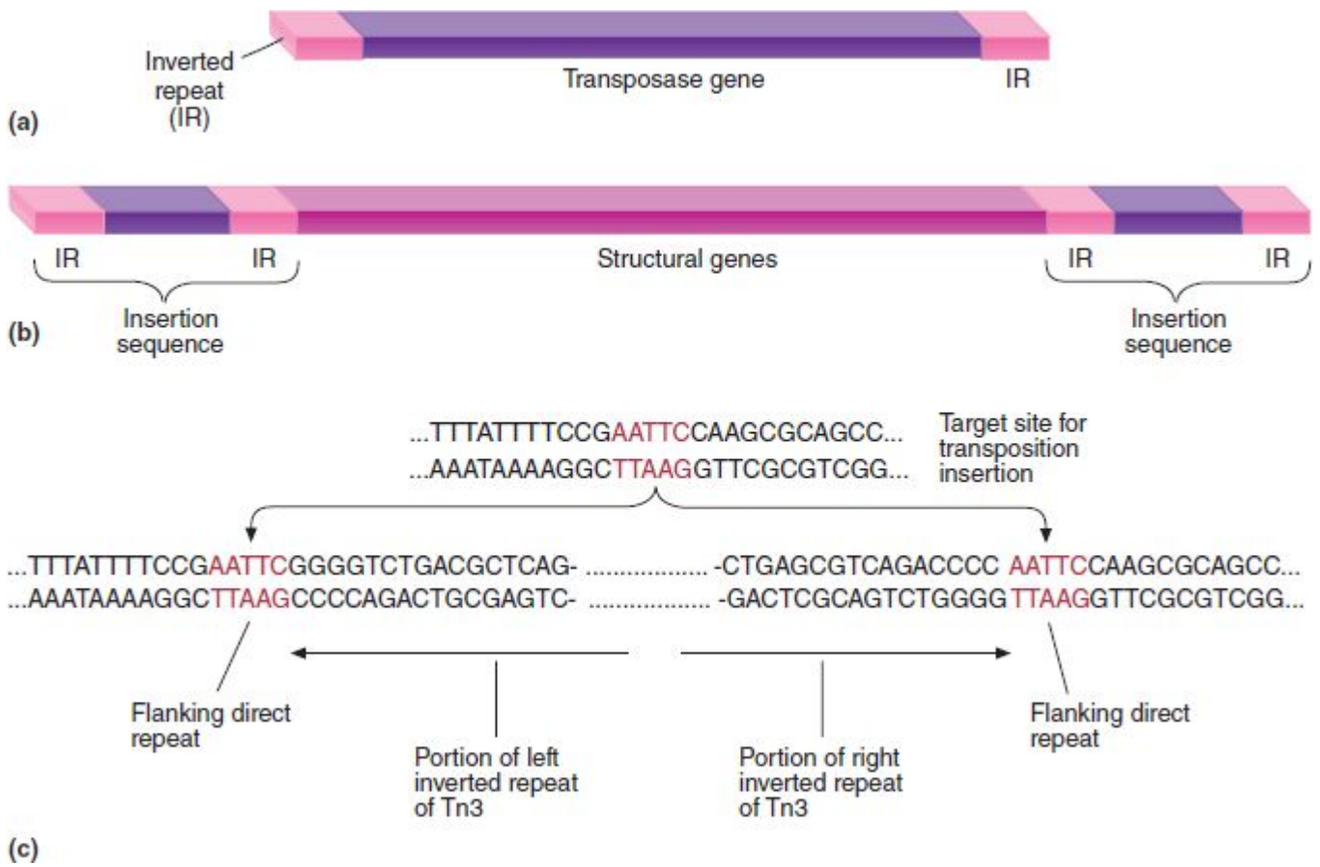


FIGURE 18.27

Bacterial Insertion Elements

(a) An insertion sequence. (b) A composite transposon. (c) Insertion of a transposon (Tn3) into bacterial DNA. The insertion process involves the duplication of the target site.

KEY CONCEPTS

- In general recombination, the exchange of DNA sequences takes place between homologous sequences.
- In site-specific recombination, DNA–protein interactions are principally responsible for the exchange of nonhomologous sequences.
- The DNA sequences called transposons can move from one site in a genome to another.

Nonreplicative Transposition. In this process the transposase makes a double-stranded cut in the donor DNA and splices it into the staggered ssDNA cut ends of the target site (a "cut-and-paste" mechanism). The cell's DNA repair system fills the gaps in the target DNA, resulting in a short, direct repeat on either side of the transposon insertion (**Figure 18.28**). An unrepaired gap can be lethal to the cell.

Replicative Transposition. In replicative transposition, only one strand of the donor DNA is transferred to the target position, and replication followed by site-specific recombination results in the duplication of the transposon rather than insertion at the new site (**Figure 18.29**). An intermediate (the cointegrate) then forms that consists of the donor segment covalently linked to the target DNA. An additional enzyme, called resolvase, catalyzes a site-specific recombination that allows the resolution of the cointegrate into two separate molecules.

Some transposons found in eukaryotes resemble those found in bacteria. For example, the Ac element, a maize transposon, is composed of a transposase gene flanked by short inverted repeats. Its discoverer, Barbara McClintock (Nobel Prize in Physiology or Medicine, 1983), called the Ac transposon a controlling element because it appeared to control the synthesis of the pigment anthocyanin in corn kernels (i.e., there was a change in gene expression). Many other eukaryotic transposons, however, have somewhat different structures than those observed in bacteria. As described previously (p. 660), retrotransposons are a significant feature of eukaryotic genomes. Depending on the changes and their location, the effects of transposons can be viewed either as disruptive and damaging or as providing opportunities for genetic diversity. Some effects of transposition are observed as changes in gene expression, a topic that is discussed in Section 18.3.

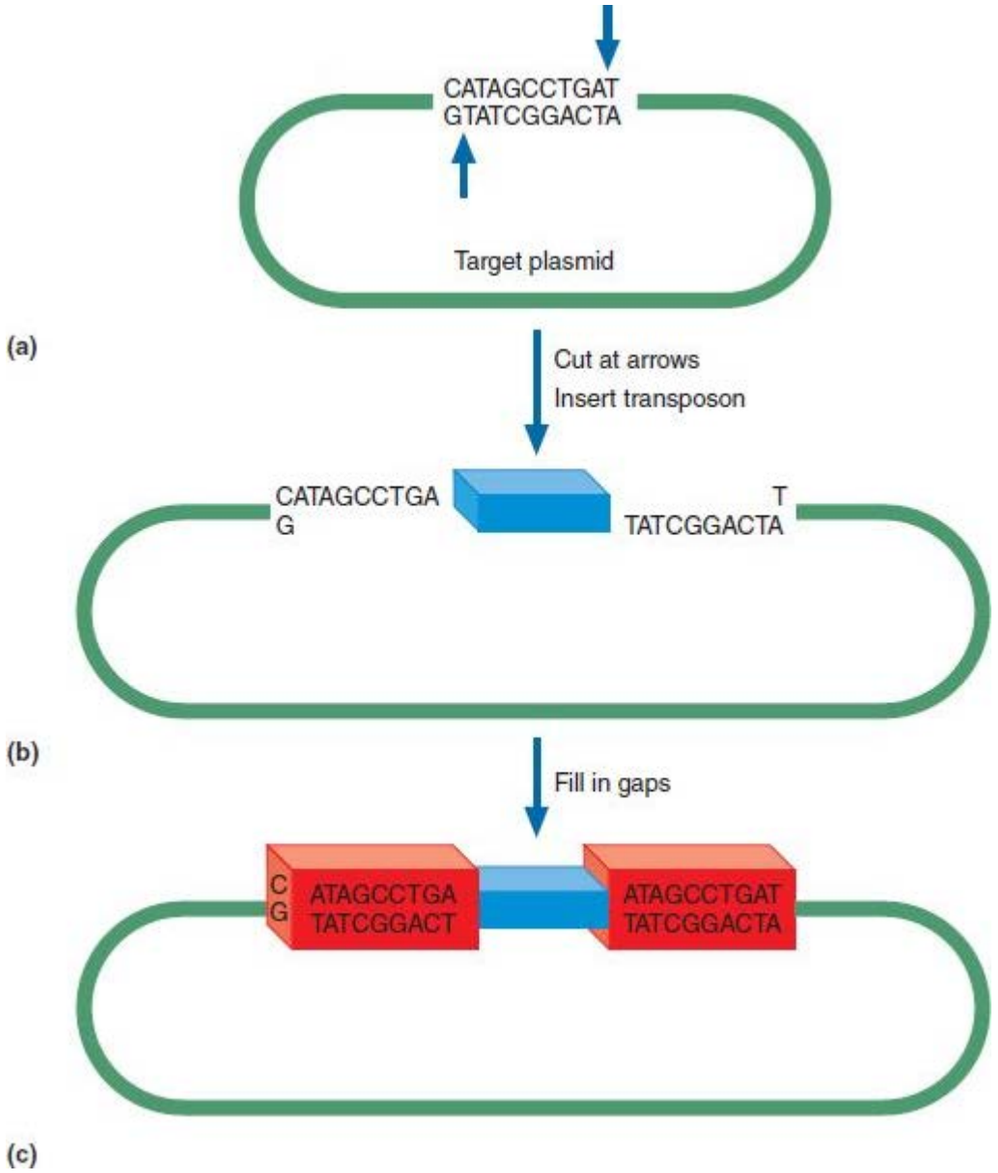


FIGURE 18.28

Nonreplicative Transposition

(a) Host DNA is cut (arrows) in a staggered fashion. (b) Each end of the transposon (blue) is covalently attached to an overhanging end of host DNA. (c) After the gaps have been filled by a DNA polymerase, there are nine base pair repeats of host DNA (red) flanking the transposon.

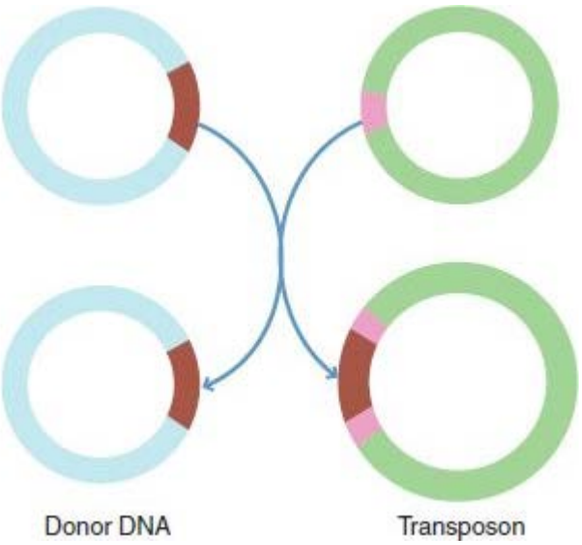


FIGURE 18.29

Replicative Transposition

The transposition of a transposable element (red) to a new location (pink) does not involve its loss from the original location. Note that in this type of transposition process, which occurs via replication of the entire transposon followed by site-specific recombination, the donor DNA is undamaged.

QUESTION 18.4

One fascinating aspect of complex organisms such as mammals is the existence of gene families, groups of genes that code for the synthesis of a series of closely related proteins. For example, several different types of collagen are required for the proper structure and function of connective tissues. Similarly, there are several types of globin gene. It is currently believed that gene families originate from a rare event in which a DNA sequence is duplicated. Some gene duplications provide a selective advantage by providing larger quantities of important gene products. Alternatively, the two duplicate genes evolve independently. One copy continues to serve the same function, while the other eventually evolves to serve another function. Can you speculate about how gene duplications occur? Once a gene has been duplicated, what mechanisms introduce variations?

Biochemistry IN THE LAB

Genomics

Genomics and functional genomics have accelerated research efforts in all of the life sciences. **Genomics** is the large-scale analysis of entire genomes. **Functional genomics** is a methodology used to analyze the functional properties of genes and proteins and how these molecules interact within living organisms. As a result of major breakthroughs in DNA technology, genomics has yielded genome sequence information for more than several thousand organisms and a variety of functional genome technologies. For example, with DNA chips (glass

or plastic wafers to which thousands of different DNA sequence probes are attached), researchers can simultaneously monitor the expression of thousands of genes in cultured cells. The following discussion describes the basic tools used in genomic technology.

Isolation, characterization, and manipulation of DNA sequences, now considered commonplace, are made possible by a series of techniques referred to as **recombinant DNA technology**. The essential feature of this technology is that DNA molecules obtained from various sources can be cut and spliced together. These techniques have made genomes more accessible to investigation than ever before because the large number of DNA copies required in DNA sequencing methods can be obtained through molecular cloning and PCR (polymerase chain reaction). Commercial applications of recombinant DNA techniques have revolutionized medical practice. For example, human gene products such as insulin and growth hormone, as well as certain vaccines (e.g., hepatitis B vaccine) and diagnostic tests (e.g., HIV diagnosis), are now produced in large quantities by bacterial cells into which recombinant genes have been inserted. Currently, several research groups are investigating the use of recombinant techniques in human gene therapy, a process in which (it is hoped) defective genes can be replaced by their functional counterparts.

Figure 18A illustrates the basic features of recombinant DNA construction. The process begins by using a restriction enzyme (see Biochemistry in the Lab: Nucleic Acid Methods, pp. 666–70, in Chapter 17) that generates sticky ends to cleave DNA from two different sources. The DNA fragments are then mixed under conditions that allow annealing (base pairing) between the sticky ends. Once base pairing has occurred, the fragments are covalently bonded together by DNA ligase. After recombinant DNA molecules have been isolated and purified, it is usually necessary to reproduce them so that sufficient quantities are available for further investigation. Molecular cloning, a commonly used method for increasing the number of copies of DNA, is discussed next.

Molecular Cloning

The term *molecular cloning* refers to the experimental methods used to create recombinant DNA molecules within host organisms. In molecular cloning, a piece of DNA isolated from a donor cell (e.g., any animal or plant cell) is spliced into a vector. A **vector** is a DNA molecule capable of replication that is used to transport a foreign DNA sequence, often a gene, into a host cell.

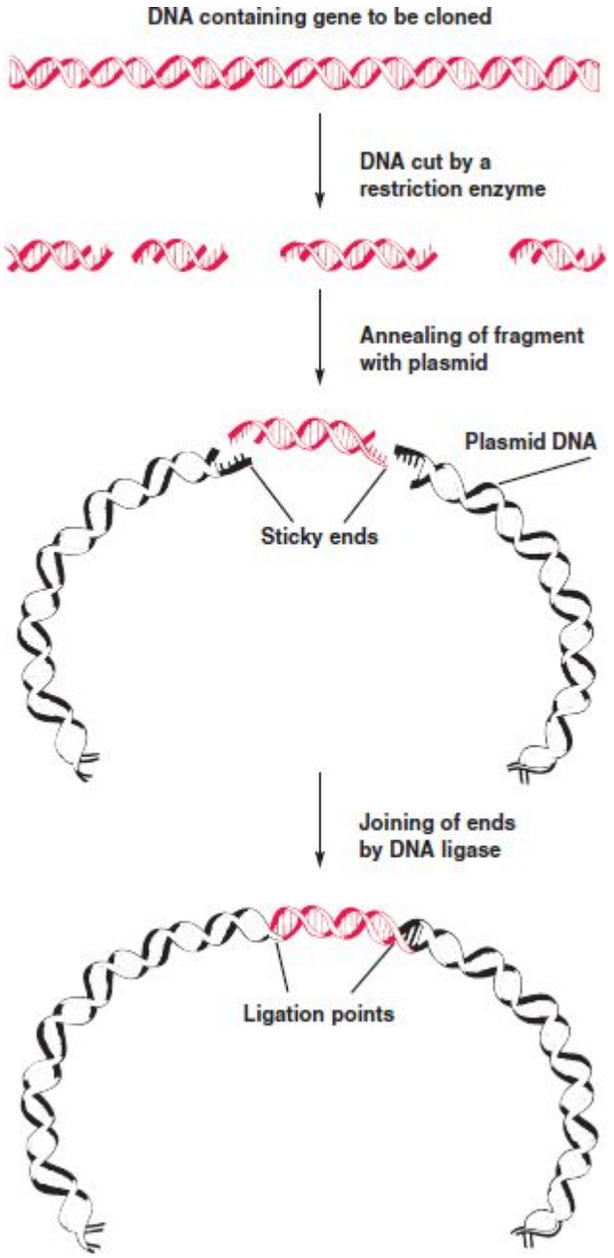


FIGURE 18A

Recombinant DNA Construction

Creation of recombinant DNA molecules begins by treating DNA from two sources with restriction enzymes. Under hybridizing conditions, DNA fragments with sticky ends anneal together. Once base pairing has occurred, DNA ligase covalently joins the fragments together.

The choice of vector depends on the size of donor DNA. For example, bacterial plasmids are often used to clone small pieces (15 kb) of DNA. Somewhat larger pieces (24 kb) are incorporated into bacteriophage λ vectors, whereas cosmid vectors are used for DNA fragments as large as 50 kb. Bacteriophage λ can be used as a vector because a substantial portion of its genome does not code for phage production and can therefore be removed. The removed viral DNA can then be replaced by foreign DNA. **Cosmids** are cloning vehicles that contain λ bacteriophage *cos* sites incorporated into plasmid DNA sequences with one or more selectable markers. The *cos* sites allow delivery to the host cell in a phage head, the plasmid DNA facilitates independent replication of the recombined unit, and the selectable markers permit detection of successful recombinants. Still larger pieces can be inserted into bacterial artificial chromosomes and yeast artificial

chromosomes. Bacterial artificial chromosomes (BACs), derived from a large *E. coli* plasmid called the F-factor, are used to clone DNA sequences as long as 350 kb. **Yeast artificial chromosomes** (YACs), which can accommodate up to 3000 kb, are constructed by using yeast DNA sequences that are autonomously replicating (i.e., they contain a eukaryotic DNA replication origin).

As noted, forming recombinant DNA requires a restriction enzyme, which cuts the vector DNA (e.g., a plasmid) open (**Figure 18B**). After the sticky ends of the plasmid have annealed with those of the donor DNA, a DNA ligase activity joins the two molecules covalently. Then the recombinant vector is inserted into host cells.

In some circumstances, the introduction of a cloning vector into a host cell is trivial. For example, phage vectors are designed to introduce recombinant DNA in an infective process called **transfection**, and some bacteria take up plasmids unaided. However, most host cells must be induced to take up foreign DNA. Several methods are used. In some prokaryotic and eukaryotic cells, the addition of Ca^{2+} to the medium promotes uptake. In others, a process called *electroporation*, in which cells are treated with an electric current, is used. One of the most effective methods for transforming animal and plant cells is the direct microinjection of genetic material. Transgenic animals, for example, are created by the microinjection of recombinant DNA into fertilized ova. Once the recombinant DNA is introduced, each type of cell in the transgenic animal replicates it along with its own genome. Note that recombinant vectors must contain regulatory regions recognized by host cell enzymes.

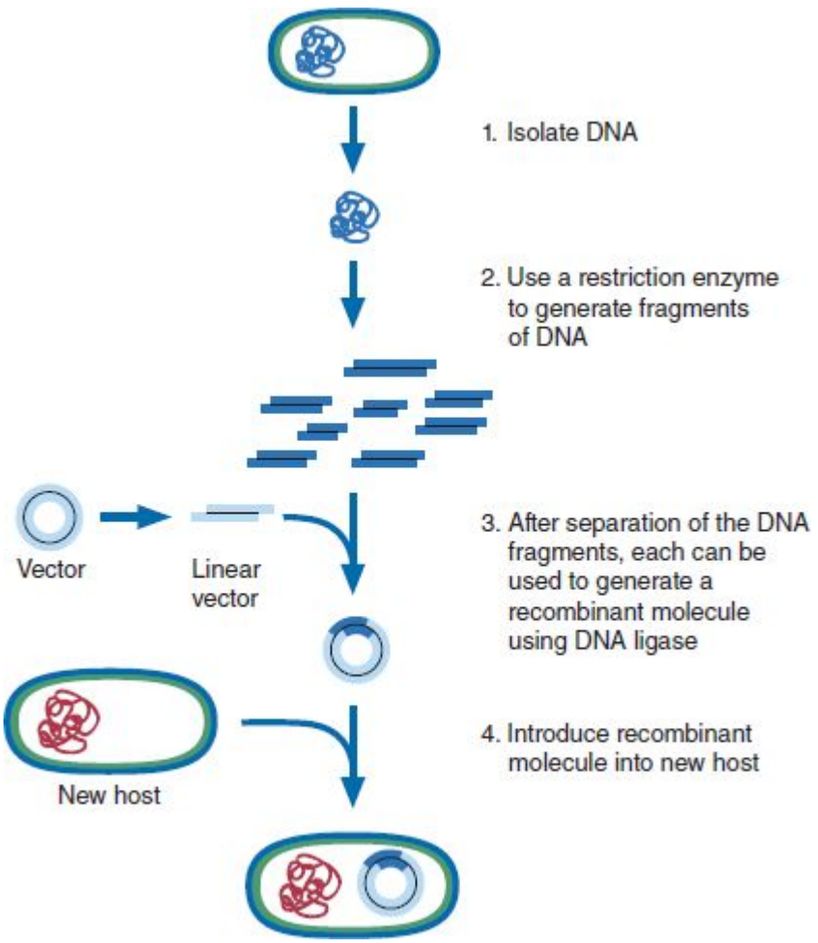


FIGURE 18B

DNA Cloning

In the cloning process, each clone is produced by introducing a recombinant molecule into a host cell, which then replicates the vector along with its own genome. The same restriction enzyme is used to create the linear vector and the DNA fragments.

As host cells that have been successfully transformed proliferate, they rapidly amplify the recombinant DNA. For example, under favorable conditions of nutrient availability and temperature, a single recombinant plasmid introduced into an *E. coli* cell can be replicated a billion times in about 11 hours. However, transformed and untransformed cells usually look exactly alike. Consequently, researchers often design cloning protocols that use vectors with selectable *marker genes* (genes whose presence can be detected) to facilitate the identification of transformed cells. Antibiotic resistance genes or color selection markers (**Figure 18C**), for example, are usually incorporated into the plasmid vectors introduced into bacteria. When bacteria exposed to plasmids with antibiotic resistance genes are plated out on a medium containing the antibiotic, only the transformed cells will grow. With eukaryotic organisms such as yeast, cells that lack an enzyme required to synthesize a nutrient may be used. For example, vectors containing the *LEU2* gene are used to transform mutant yeast cells that lack a specific enzyme in the leucine biosynthetic pathway. Only cells that have successfully transformed are able to grow in a leucine-deficient medium.

In another approach, the *colony hybridization technique*, bacteria are screened by using a radioactively labeled nucleic acid probe, an RNA molecule or a single-stranded DNA molecule with a sequence complementary to that of a specific sequence within the recombinant DNA. Bacterial cells are plated out onto solid media in petri dishes and allowed to grow into colonies. Each plate is then blotted with a nitrocellulose filter. (Some cells from each of the original colonies remain on the petri dishes.) The cells on the nitrocellulose filter are lysed, and the released DNA is treated so that hybridization with the probe can occur. Once nonhybridized probe molecules have been washed away, autoradiography (see Biochemistry in the Lab: Cell Technology, pp. 68–70, in Chapter 2) is used to identify the colonies on the master plate that possess the recombinant DNA of interest.

Polymerase Chain Reaction

Although cloning has been immensely useful in molecular biology, **PCR** is a more convenient method for obtaining large numbers of DNA copies. Using a heat-stable DNA polymerase from *Thermus aquaticus* (Taq polymerase), PCR can amplify any DNA sequence, provided the flanking sequences are known (**Figure 18D**). Flanking sequences must be known because PCR amplification requires primers. Priming sequences are produced by automated DNA-synthesizing machines.

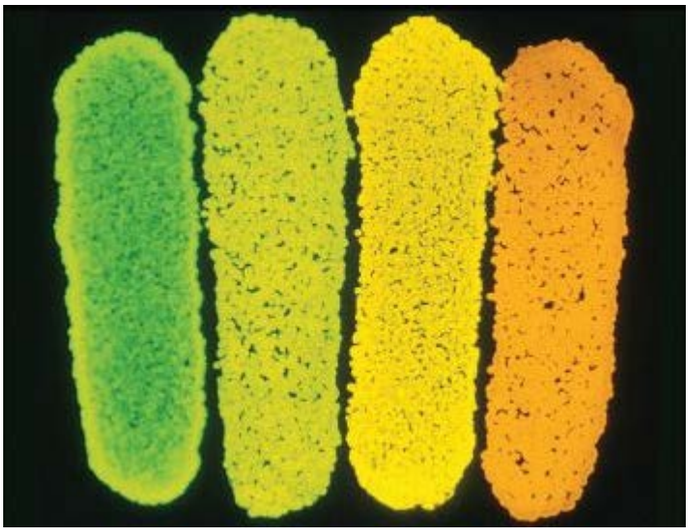


FIGURE 18C

Four Recombinant *E. coli* Cells, Each with a Different Variant of the Gene for Luciferase

Luciferase is an enzyme found in fireflies, mollusks, and several types of deep-sea fish. When in the presence of ATP and luciferin, luciferase catalyzes a light-emitting reaction. The luciferins are a group of bioluminescent compounds that emit light when they are oxidized by O_2 in a reaction catalyzed by a luciferase.

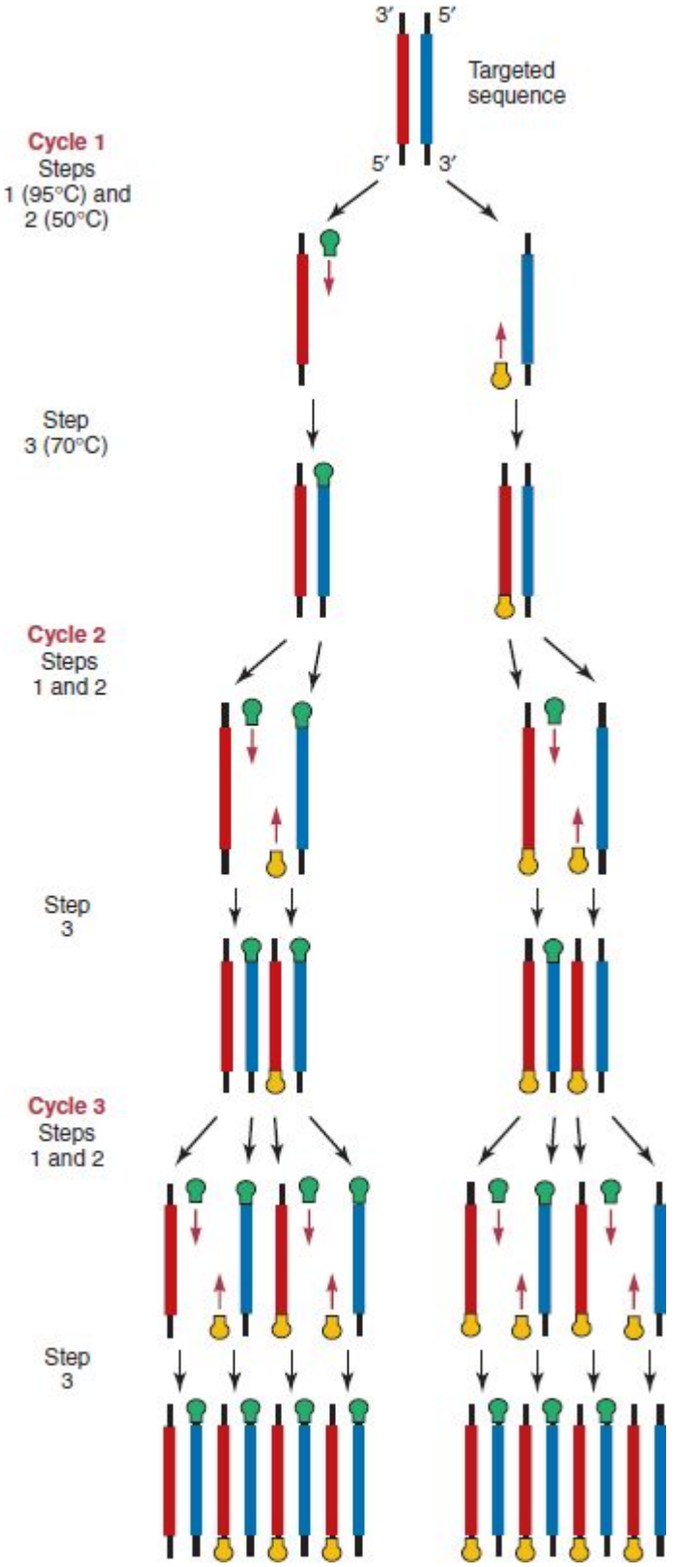


FIGURE 18D

Polymerase Chain Reaction

A single DNA molecule can be amplified millions of times by replicating a three-step cycle. In the first step, the dsDNA sample is denatured by heating to 95°C. In step 2, the temperature is quickly lowered to 50°C and an oligonucleotide primer is added. The primer hybridizes to complementary sequences on the ends of the two strands. During step 3, DNA synthesis occurs as the temperature is raised to 70°C, the optimal temperature of *Taq* polymerase. The cycle is then repeated, with both old and new strands serving as templates.

PCR begins by adding *Taq* polymerase, the primers, and the ingredients for DNA replication to a heated sample of the target DNA. (Recall that heating DNA separates its strands.) As the mixture cools, the primers attach to their complementary sequences on either side of the target sequence. Each strand then serves as a template for DNA replication. At the end of this process, referred to as a *cycle*, the copies of the target sequence have been doubled. The process can be repeated indefinitely, synthesizing an extraordinary number of copies. By the end of 30 cycles, for example, a single DNA fragment has been amplified 1 billion times.

Genomic Libraries

Genomic libraries, also called clones or gene banks, are collections of clones derived from fragments of entire chromosomes or genomes. They are used for a variety of purposes, the most important of which are the isolation of specific genes whose chromosomal location is unknown and in genome-wide sequencing efforts (gene mapping). Genomic libraries are produced in a process, referred to as **shotgun cloning**, in which a genome is randomly digested (**Figure 18E**). The range of fragment sizes, which is determined by the type of restriction enzyme and the experimental conditions chosen, must be compatible with the vector. To ensure that all sequences of interest are represented in the library, DNA samples are often only partially digested. The location of any gene can be identified if an appropriate probe is available.

When a DNA sequence (a clone) in a genomic library is too large to sequence, the cloned DNA is fragmented and subcloned in a process referred to as **chromosome walking**. One of the subclones is picked and sequenced and a small fragment at one end is used as a probe to select one of the remaining subclones that contains that sequence (**Figure 18F**). The new fragment is sequenced, and a portion is used as a probe to select other overlapping clones. In this way, contiguous sequences can be mapped. A set of overlapping sequences is referred to as a **contig**. When eukaryotic genomes are analyzed, their large size often requires the use of large cloning vectors such as YACs (p. 713) and a technique called chromosomal jumping. In **chromosomal jumping**, a type of chromosome walking, the overlapping clones contain DNA sequences of several hundred kb that are generated using restriction enzymes that cut infrequently. Used to bypass difficult regions (e.g., repetitive sequences), this technique allows rapid movement through genome mapping or target gene location.

In a variation of genomic libraries, collections of complementary DNA molecules called **cDNA libraries** are produced from mRNA or miRNA molecules by reverse transcription. This technique can be used to evaluate the transcriptome of certain cell types under specified circumstances. In other words, it is a method for determining which genes are expressed in a particular cell type. For example, with the use of DNA chip technology (DNA microarrays), gene expression in normal and diseased cells can be investigated and compared. cDNA libraries are especially useful when eukaryotic DNA is cloned because mRNA molecules lack noncoding or intron sequences. Consequently, gene products can be more easily identified, and large amounts of gene product can be generated in bacteria, which cannot process introns.

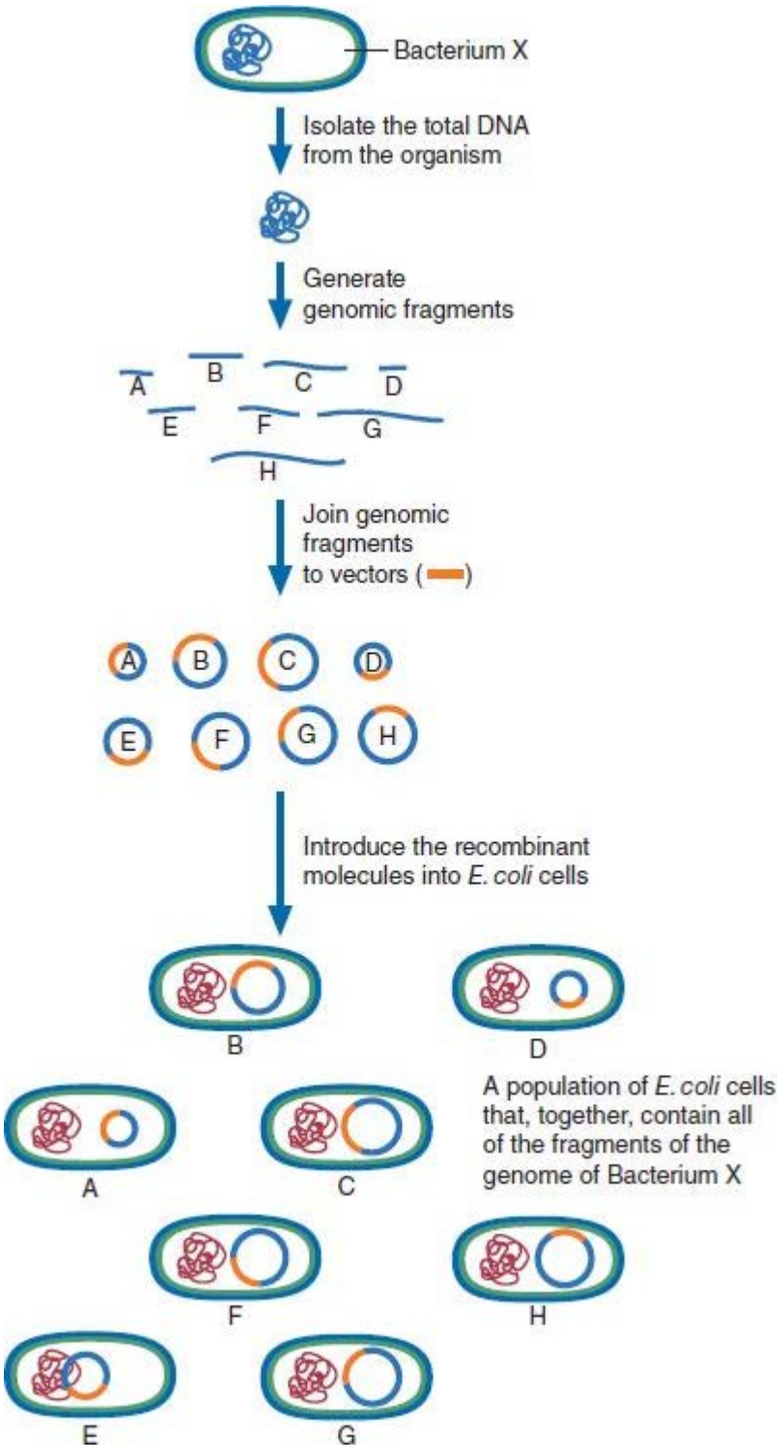


FIGURE 18E

Creation of a DNA Library Using the Shotgun Method

After the organism's DNA is isolated and purified, it is cleaved into fragments with a restriction enzyme. The fragments are then randomly incorporated into vectors, and the recombinant molecules are introduced into cells such as *E. coli*. The collection of these cells is called a genomic library.

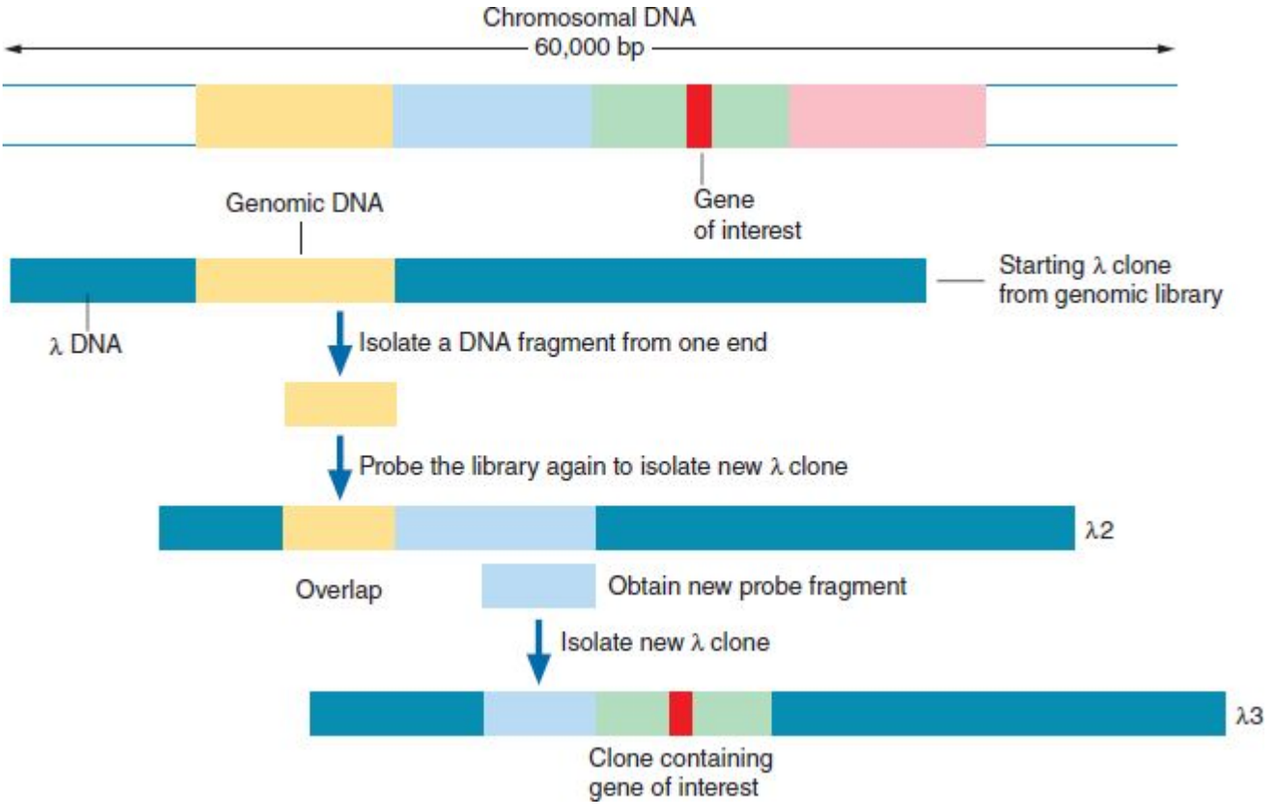


FIGURE 18F

Chromosome Walking

In chromosome walking, DNA clones, which contain overlapping sequences, are systematically identified. They may then be mapped and sequenced. Unknown genes may also be searched for. The process begins when DNA is cleaved into pieces and cloned. (In this example, bacteriophage λ vectors are used.) One end of the starting clone is labeled and used as a probe to identify the clone in the λ library that contains both that sequence and an adjacent sequence. Repeating this step results in the labeling of the end of the second clone, which is used as a probe to identify yet another overlapping clone. The process continues until a collection of clones that together contain all the sequences in the original DNA fragment has been sequenced and mapped.

Microarrays

Microarrays, or "chips," are used to analyze the expression of genes simultaneously (**Figure 18G**). Often no larger than a postage stamp, a **DNA microarray** consists of thousands or hundreds of thousands of oligonucleotides such as cDNAs or antisense RNAs (RNAs that are complementary in sequence to mRNAs or miRNAs) that are attached to the minute pores within a glass, plastic, or silicon support. The molecule in each position in a microarray is designed to hybridize with a target molecule (a specific DNA or RNA sequence). Researchers also utilize protein and antibody chips that are used to track protein–protein interactions or target antigens (molecules that induce antibody production), respectively.

In investigations of gene expression, an entire set of mRNA molecules, for example, from the cells of interest (i.e., the transcriptome) are reverse-transcribed into cDNA. The cDNA molecules, labeled with a fluorescent dye, are then incubated with a microarray under hybridizing conditions. After incubation, the microarray is washed to remove unhybridized molecules. Researchers then determine which genes are being expressed by identifying the positions on the microarray that are fluorescing. Scientists have used microscopes, photomultiplier tubes, and computer software to observe changes in gene expression in a variety of circumstances. Examples include comparisons

of normal and cancerous cells and cells exposed to different nutrients or signal molecules.

Genome Projects

Each genome project determines the entire set of DNA base sequences of a particular organism. The process entails taking a large number of sequence fragments obtained by fracturing the genome and then determining their base sequences with an automatic sequencing method. The sequence data for each of the fragments are then assembled using computer methods to yield the entire genome sequence.

The Human Genome Project (completed in 2003) was an intensive international effort to determine the nucleotide sequence of the human genome. As this goal was reached, the attention of researchers shifted to the **annotation** (i.e., functional identification) of approximately 20,000 human genes. Just as scientists have historically used structural and functional comparisons of other organisms in anatomy, biochemistry, physiology, and medicine to better understand human biology, the current effort to interpret human genome data is being aided immensely by comparisons with the information obtained in other genome projects. The genomes of well-researched organisms as diverse as bacteria (e.g., *E. coli*), yeast (e.g., *S. cerevisiae*), the worm *Caenorhabditis elegans*, the fruit fly *Drosophila*, and various mammals (e.g., the mouse) have been used in genome structure analysis and in the assignment of recently discovered genes in other organisms. Other interesting genome projects include those of Neanderthals and the human microbiome.

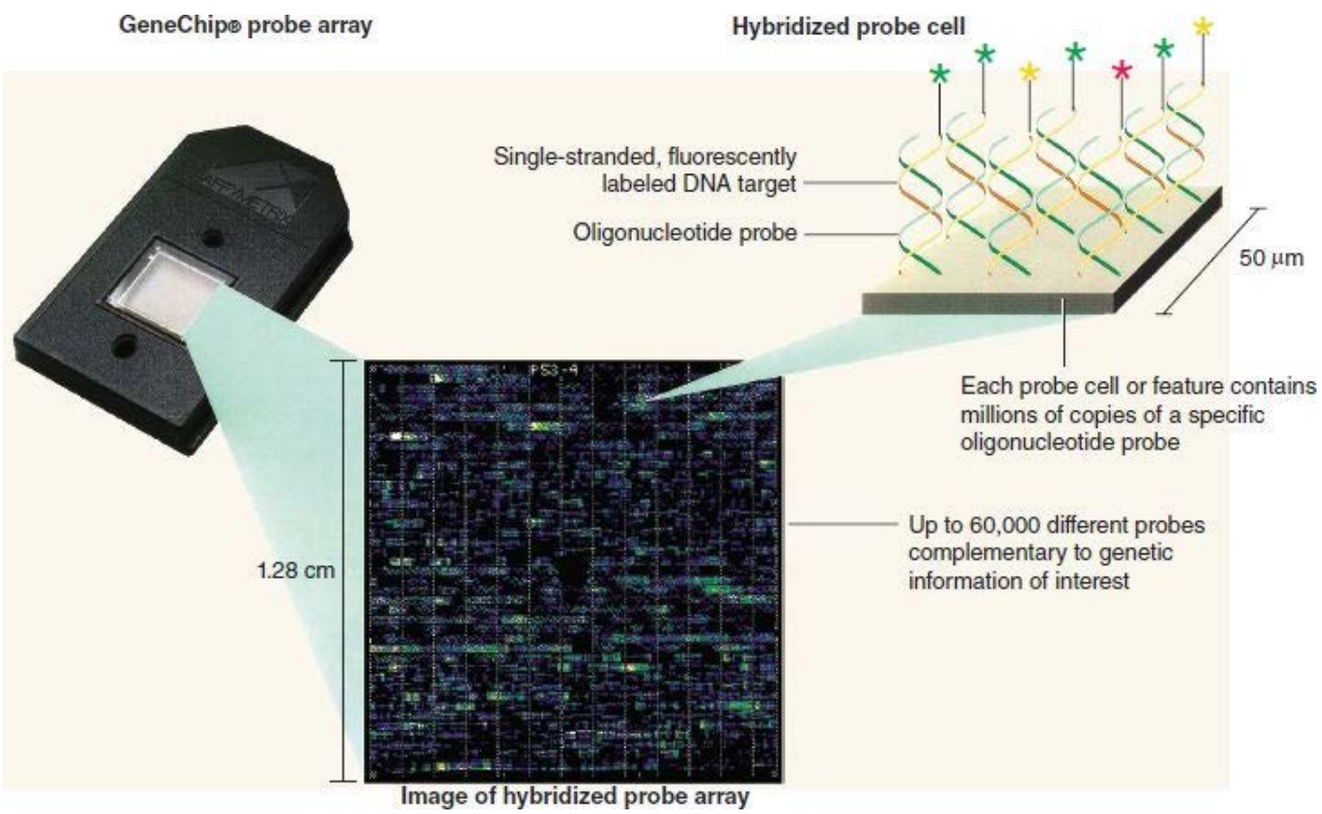


FIGURE 18G

DNA Microarray Technology

DNA microarrays can be used to determine which genes are expressed in a specific cell type because each "chip" can accommodate from thousands to millions of DNA probes. (Oligonucleotide probes are

synthesized on the chip surface by means of photolithographic techniques similar to those used in the manufacture of computer chips.) The microarray is incubated under hybridizing conditions with fluorescently labeled cDNA. The cDNA molecules are derived from mRNA extracted from the cells of interest.

ENCODE (the Encyclopedia of DNA Elements) is a research project, funded by the U.S. National Human Genome Research Institute (a division of the U.S. National Institutes of Health), that involves a worldwide group of research laboratories. ENCODE's goal, made possible by ever more powerful DNA sequencing technologies, is twofold: (1) to determine the function of the 98.5% of the human genome that does not code for proteins, and (2) to stimulate the development of DNA-based therapies to prevent and treat human diseases. Examples of ENCODE's preliminary results include the surprising fact that the human genome is pervasively transcribed and that the vast majority (about 75%) of human DNA sequences are in close proximity of ENCODE-identified regulatory sequences. Recently reported additional information in ENCODE includes more than 8 million identified DNA sequences that are recognition motifs for DNA-binding proteins, as well as genome-wide binding patterns of over a hundred major transcription factors.

Bioinformatics

The emergence of *high-throughput* (i.e., rapid, high-volume, automated) technologies to analyze living organisms has created a vast amount of data on nucleic acid and polypeptide sequences. The information, which is collected from genome and proteome sequencing projects and from microarray analysis of cell processes such as transcription, is placed in databases that are available to the scientific community. How do scientists analyze such enormous volumes of raw data? As a result of technological advances in computer science, applied mathematics, and statistics, **bioinformatics** has provided researchers with a powerful investigational tool. The use of computer algorithms has made the investigation of a wide variety of previously intractable problems feasible, as the following examples illustrate.

1. Genes can be located by a process referred to as sequence inspection. Gene prediction programs utilize several clues to locate sequences that can potentially code for polypeptides called open reading frames (ORFs) (p. 673). ORFs are DNA base sequences (a series of codons starting with a start codon and ending with a stop codon) that, when converted to an mRNA sequence, has the capacity to be translated to yield a polypeptide.
2. Alignment of DNA sequences allows researchers to search the genomes of hundreds of organisms for similarities among gene or regulatory sequences and has provided invaluable insight into the relatedness of living organisms and the mechanisms used to sustain living processes.
3. Protein structure prediction has been facilitated by a method called homology modeling. Once a new protein-coding gene has been discovered, bioinformatic analysis is used to search among homologous or near-homologous molecules whose structure is already known.
4. Bioinformatic analysis of the massive data derived from protein microarrays and MS-derived cell proteome data provides an invaluable means of analyzing cell protein synthesis patterns. For example, this type of data analysis allows medical scientists to compare how normal cell proteins are altered in disease states.
5. In the field of evolutionary biology, bioinformatic programs have been used to trace

the lineages of organisms based on rare events such as gene duplications and lateral gene transfer. (Lateral gene transfer is the transfer of genes between species.)

6. High-throughput gene expression analysis is now used to identify the genes involved in medical disorders (e.g., in comparisons of the transcription products of normal and cancerous cells).
7. Several bioinformatic approaches have been developed to investigate cellular organization. Examples include HiC, a type of chromosome conformation capture analysis (cross-linked DNA protein fragment analysis) and ChIA-Pet (*ch*romatin *i*nteraction *a*nalysis by *p*aired-*e*nd *t*ag sequencing), which have been devised to investigate the three-dimensional structure and nuclear organization of chromatin.
8. Systems biologists' use of complex mathematical modeling combined with an ever-increasing source of biological data promises to significantly improve our understanding of life's operating systems.

18.2 TRANSCRIPTION

Transcription, the creation of RNA copies of DNA sequences, is a highly regulated process that transforms environmental cues (e.g., nutrient availability in bacteria and developmental signals in multicellular eukaryotes) into the products of gene expression changes. As with all aspects of nucleic acid function, the synthesis of RNA molecules is a complex process involving a variety of enzymes and associated proteins. RNA polymerase, sometimes referred to as DNA-dependent RNA polymerase, or RNAP, catalyzes RNA synthesis. RNA synthesis occurs in the 5′ to 3′ direction. The reaction catalyzed by all RNA polymerases is

$$\text{NTP} + (\text{NMP})_n \rightarrow (\text{NMP})_{n+1} + \text{PP}_i$$

where each new nucleotide is added to the 3′ end of the molecule. Because the nontemplate or plus (+) strand of DNA has the same base sequence as the RNA transcription product (except for the substitution of U for T), it is also called the **coding strand** (**Figure 18.30**). By convention, the direction of the gene, a segment of double-stranded DNA, is the same as the direction of the coding strand. Polymerization proceeds from the 5′ end to the 3′ end of the gene because the template DNA strand, also called the minus (−) strand, and the newly made RNA molecule are antiparallel. As noted, transcription generates several types of RNA, of which rRNA, tRNA, and mRNA are directly involved in protein synthesis (Chapter 19).

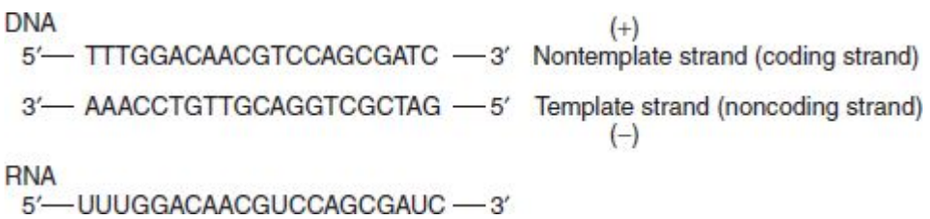


FIGURE 18.30

DNA Coding Strand

One of the two DNA strands, referred to as the template (−) strand, is transcribed. The RNA transcript is identical in sequence to the nontemplate (+) or coding strand, except for the substitution of U for T.

Transcription in Prokaryotes

The RNA polymerase (**Figure 18.31**) in *E. coli* catalyzes the synthesis of all RNA classes. The core RNAP enzyme (α_2, β, and β') has a molecular weight of 370 kDa. Another protein, the ω subunit, promotes the assembly of the RNAP holoenzyme. The transient binding of the σ (sigma) factor to the core enzyme allows it to bind both the correct template strand and the proper site to initiate transcription. A variety of σ factors have been identified. For example, in *E. coli*, σ^{70} is involved in the transcription of most genes, whereas σ^{32} and σ^{28} promote the transcription of heat shock genes and the flagellin gene, respectively. (As its name suggests, flagellin is a protein component of bacterial flagella.) The superscript indicates the protein's molecular weight in kilodaltons.

Transcription of an *E. coli* gene is illustrated in **Figure 18.32**. The entire process consists of three stages: initiation, elongation, and termination. Each is discussed briefly. Transcription initiation involves the binding of RNA polymerase to a **promoter**, a regulatory DNA sequence that is located upstream (i.e., toward the 5′ end of a polynucleotide) of a gene. Although prokaryotic promoters are variable in size (from 20 to 200 bp), two short sequences at positions about 10 and 35 bp away from the transcription initiation site are remarkably similar among various bacterial species. There is a set of base sequences, called consensus sequences, at each of these sites. A **consensus sequence** represents an average of a number of closely related but nonidentical sequences. The sequences shown in **Figure 18.33** are named in relation to the transcription starting point, the −35 region and the −10 region. (The −10 region is also called the *Pribnow box* [consensus sequence TATAAT], after its discoverer.) RNA polymerase slides along the DNA until it reaches a promoter sequence. Promoters vary widely in the efficiency with which they productively bind RNA polymerase. Transcription initiation rates between "strong" promoters (close to the consensus sequence) and "weak" promoters (far from the consensus sequence) may vary by as much as a thousandfold. Mutations within a promoter sequence usually weaken the promoter but can also convert a weak promoter into a stronger one. Neither possibility is favorable.

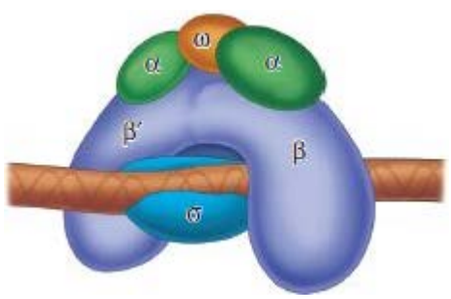


FIGURE 18.31

***E. coli* RNA Polymerase**

The *E. coli* RNA polymerase consists of two α subunits and one each of β, and β' subunits. The ω subunit facilitates the assembly and stability of RNAP. The transient binding of a σ subunit (not shown) allows binding of the core enzyme to appropriate DNA sequences.

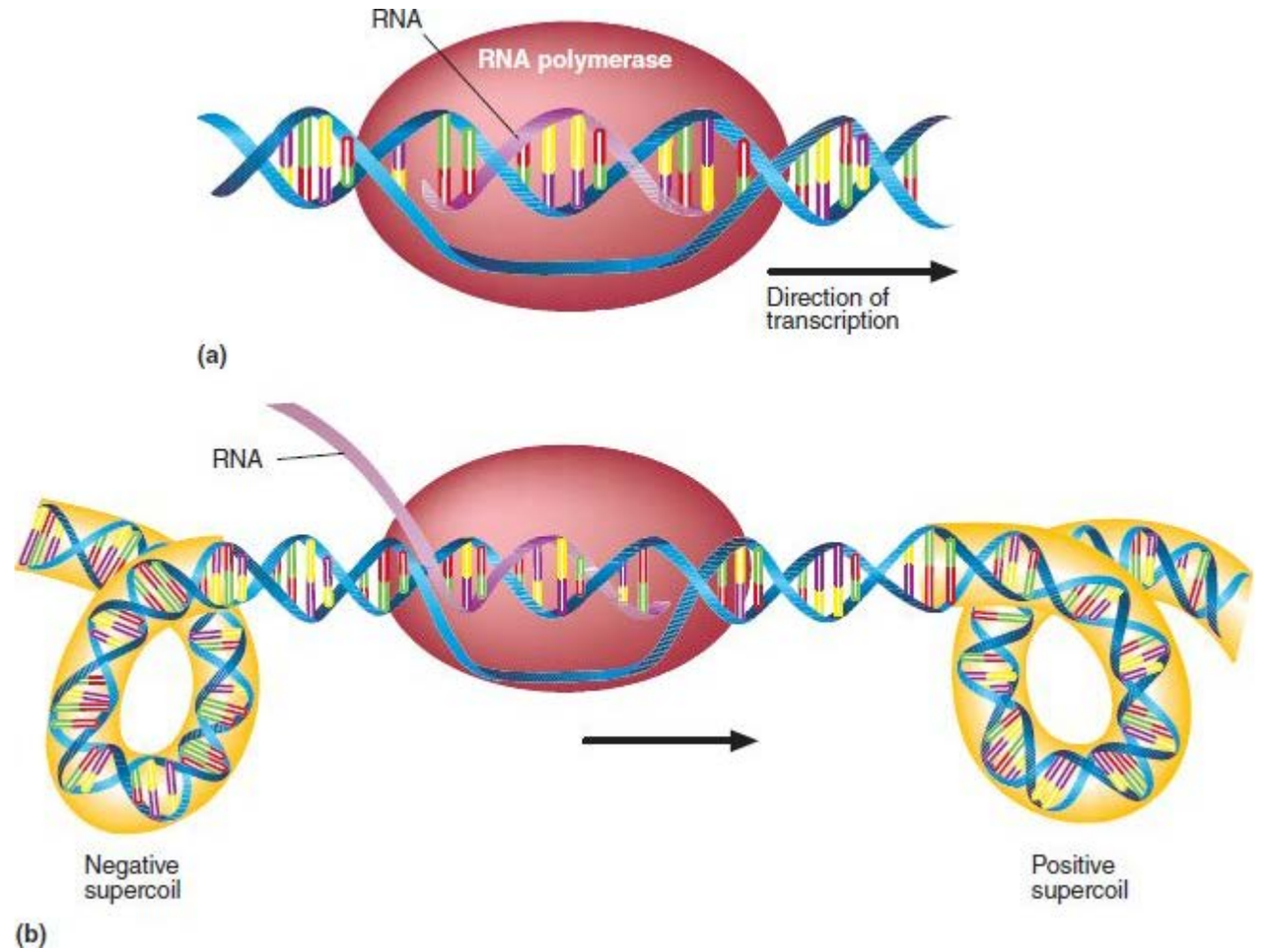


FIGURE 18.32

An Idealized Image of Transcription Initiation in *E. coli*

(a) A transcription bubble forms as a short DNA segment unwinds. An RNA-DNA hybrid forms as transcription progresses. The bubble moves to keep up with transcription as DNA unwinds ahead of it and rewinds behind it. (b) Transcription induces coiling. Positive supercoils form ahead of the bubble, while negative supercoils form behind it. These supercoils are relaxed by topoisimerases.

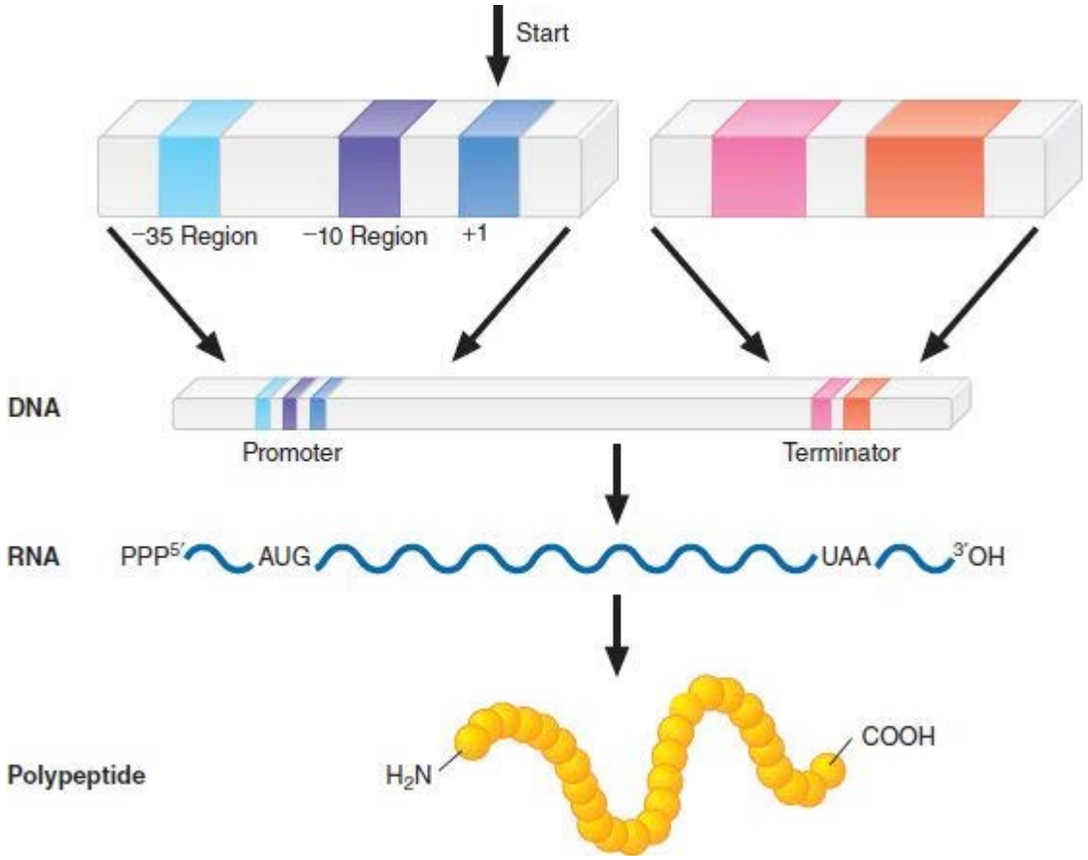


FIGURE 18.33

Typical *E. coli* Transcription Unit

If RNA polymerase can bind to the promoter, DNA transcription begins at the start site +1. Translation of mRNA begins as soon as the ribosome-binding site on the mRNA transcript is available.

RNAP and the Prokaryotic Transcription Process

The RNAP holoenzyme (i.e., the core enzyme and its associated σ factor) binds to the promoter region. A short DNA segment then unwinds as the β' and σ subunits break the hydrogen bonds between 13 bp in the Pribnow box. Because the DNA strands are now separated, the enzyme-promoter complex is referred to as "open."

Transcription begins with the binding of the first nucleoside triphosphate (usually ATP or GTP) to the RNA polymerase complex. A nucleophilic attack by the 3′-OH group of the first nucleoside triphosphate on the α-phosphate of a second nucleoside triphosphate (also positioned by base pairing in the adjacent site) causes the first phosphodiester bond to form. (Because the phosphate groups of the first nucleotide molecule are not involved in this reaction, the 5′ end of prokaryotic transcripts possesses a triphosphate group.) RNAP proceeds to catalyze the formation of additional phosphodiester bonds between ribonucleotides that are base-paired to the DNA template strand within the promoter. For the initiation phase to end successfully (i.e., for the RNAP holoenzyme to move away from the promoter), the growing RNA chain must reach a length of about 10 nt. Several attempts at "promoter clearance" usually fail, and the truncated transcripts are released and then degraded. When the transcribed sequence reaches a length of 10 nucleotides, the conformation of the RNA polymerase complex changes, the σ factor is released, and the initiation phase ends. As soon as one RNA polymerase has moved beyond the promoter site, another RNA polymerase can move in, bind to the site, and start another round of RNA synthesis.

The elongation phase begins once the σ factor has detached and the affinity of the RNAP

complex for the promoter site has decreased. The core RNA polymerase converts to an active transcription complex as it binds several accessory proteins. As RNA synthesis proceeds in the 5′ → 3′ direction, the DNA unwinds ahead of the *transcription bubble* (the transiently unwound DNA segment, composed of 12 to 14 bp, in which an RNA-DNA hybrid has formed). At any one time, there are the equivalent of about 30 bp of DNA within RNAP. The active site of the enzyme complex lies between the β and β' subunits (**Figure 18.34**). As the dsDNA enters the enzyme and is separated into two strands, the template strand enters the active site through a channel. The nontemplate strand is looped away from the active site and travels in its own channel. When the template and nontemplate strands emerge from their separate channels, they re-form a double helix. Meanwhile, the growing RNA **transcript** exits through its own channel formed in the β and β' subunits. The unwinding action of RNA polymerase creates positive supercoils ahead of the transcription bubble and negative supercoils behind the bubble, which are resolved by topoisomerases. As the bubble moves down the gene, it is said to move "downstream." The incorporation of ribonucleotides continues until a termination signal is reached.

There are two types of transcription termination in bacteria: intrinsic termination and rho-dependent termination. In **intrinsic termination** (also referred to as **rho-independent termination**), RNA synthesis is terminated as the result of the transcription of the termination sequence that consists of an inverted repeat sequence followed by from 6 to 8 As (**Figure 18.35**). When the termination sequence is transcribed, the inverted repeat forms a stable hairpin that causes the RNA polymerase to slow or stop. The RNA transcript is then released because the RNA-DNA hybrid dissociates owing to weak base pair interactions between the As in a short polyadenylate [poly(A)] sequence that follows the inverted repeat and the complementary Us in the transcript. In **rho-dependent termination**, strong hairpins do not form, and termination requires the aid of a protein known as the rho factor, an ATP-dependent helicase (**Figure 18.36**). Rho factor binds to a specific recognition sequence on the nascent mRNA strand upstream from the termination site. It then proceeds to unwind the RNA-DNA helix to release the transcript and dislodge the polymerase.

KEY CONCEPTS

- During transcription, an RNA molecule is synthesized from a DNA template.
- In prokaryotes, this process involves a single RNA polymerase activity.
- Transcription is initiated when the RNA polymerase complex binds to a promoter sequence.

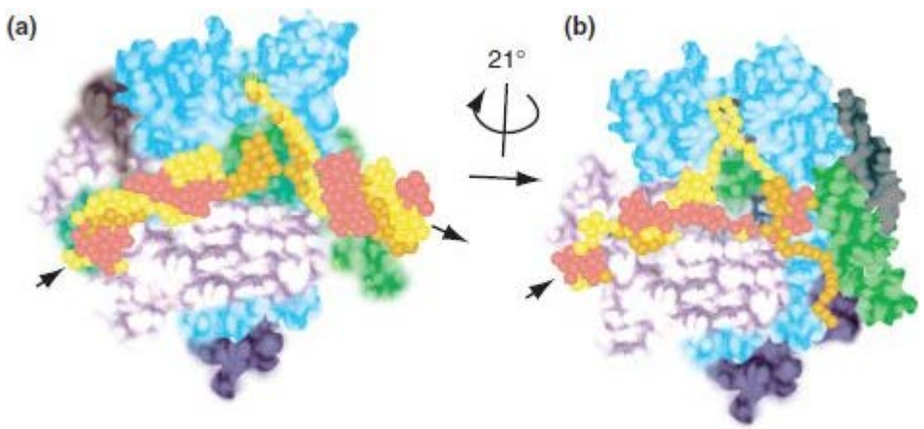


FIGURE 18.34

Bacterial RNA Polymerase Model

In this model, the components of RNAP are color-coded as follows: β (blue), β' (pink). The DNA template strand is red and the nontemplate strand is yellow. In (a) and (b), the double helix lies within a horizontal trough between the β and β' subunits. The black arrows indicate the direction of DNA movement within the enzyme. Rotation of the model by 21° reveals the growing RNA transcript (gold) as it exits the RNAP.

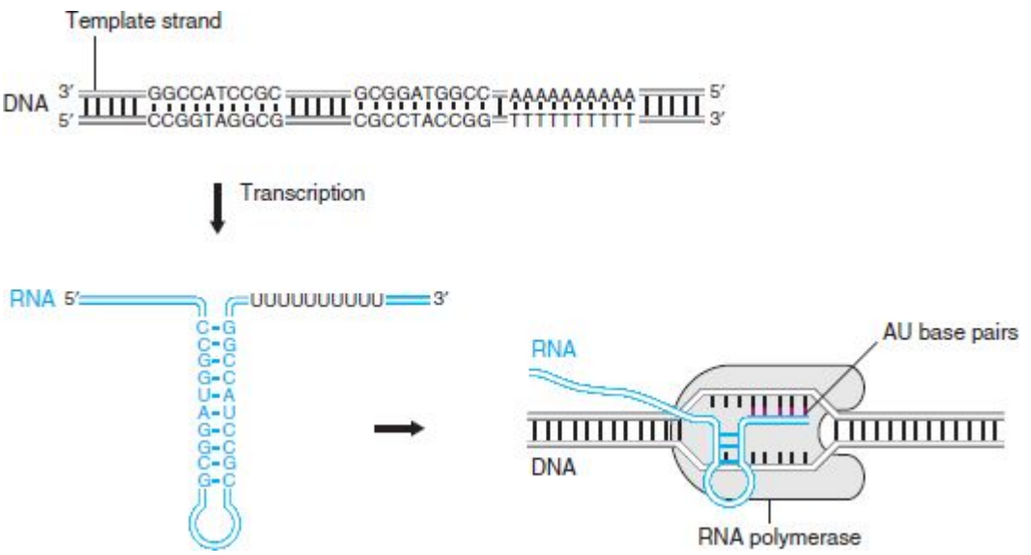


FIGURE 18.35

Intrinsic Termination

When the termination sequence (an inverted repeat followed by a series of As) in the template strand has been transcribed, the resulting RNA sequence forms into a stable hairpin (stem-loop) structure. After the disruption of DNA-RNA interactions by the hairpin, the RNA transcript is held on the template strand only by a short sequence of AU base pairs. The RNA molecule quickly dissociates because AU interactions are very weak.

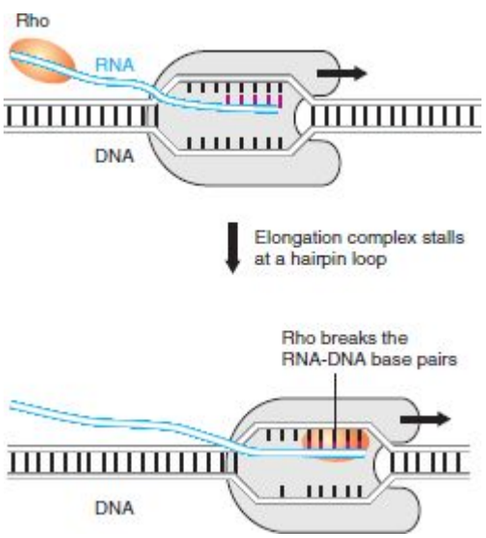


FIGURE 18.36

Rho-Dependent Termination

Rho factor is an ATP-dependent helicase that binds to a C-rich sequence in the RNA transcript. Once bound, rho moves along the transcript in the 5′ → 3′ direction until it reaches the termination site now inside the transcription bubble. Rho disrupts DNA-RNA interactions by using ATP-derived energy to unwind the DNA-RNA hybrid, and the RNA transcript is released.

In prokaryotes, mRNA is translated as soon as the transcript's ribosome-binding site is exposed (cotranscriptional translation). Mature rRNA and tRNA molecules are produced from larger transcripts by posttranscriptional processing. The RNA processing reactions for *E. coli* rRNA are outlined in **Figure 18.37**. The *E. coli* genome contains several sets of the rRNA genes 16S, 23S, and 5S. In the primary processing step, the polycistronic 30S transcript is methylated and then cleaved by several RNases into a number of smaller segments. Further cleavage by different RNases produces mature rRNAs. A few tRNAs are also produced. The other tRNAs are produced from primary transcripts in a series of processing reactions in which they are trimmed down by several ribonucleases (RNases). In the last step of tRNA processing, a large number of bases are altered by several modification reactions (e.g., deamination, methylation, and reduction).

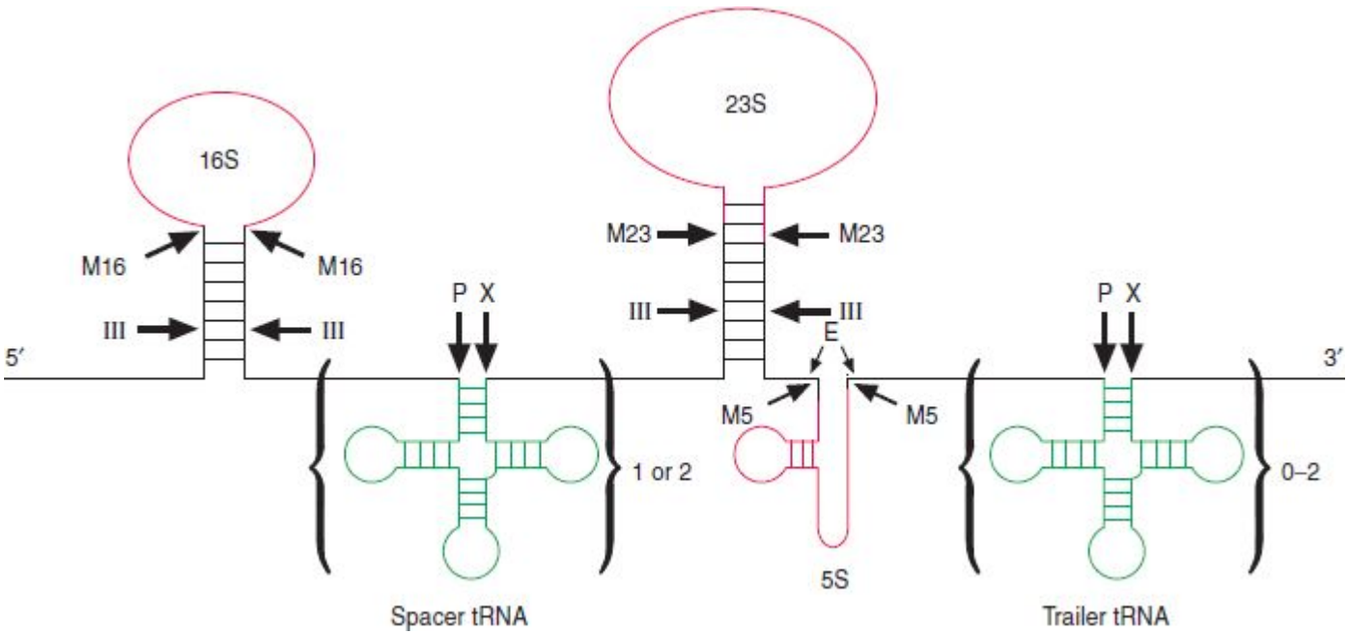


FIGURE 18.37

Ribosomal RNA Processing in *E. coli*

Each rRNA operon encodes a primary transcript that contains one copy each of 16S, 23S, and 5S rRNAs. Each transcript also encodes one or two spacer tRNAs and as many as two trailer tRNAs. Posttranscriptional processing involves numerous cleavage reactions catalyzed by various RNases and splicing reactions. (Individual RNases are identified by letters and/or numbers, e.g., M5, X, and III.) RNase P is a ribozyme.

Transcription in Eukaryotes

DNA transcription in prokaryotes and eukaryotes is similar in several respects. For example, the bacterial RNA polymerases and their eukaryotic counterparts are structurally similar and use a common transcription mechanism (e.g., promoter recognition, truncated transcripts, and promoter clearance). Also, although the initiation factors in bacteria and eukaryotes are only distantly related, they perform similar functions. These two types of organisms differ significantly, however, in the regulatory mechanisms that control gene expression. One of the most prominent examples of these differences is the limited access of the eukaryotic transcription machinery to DNA. Chromatin is usually at least partially condensed. Yet for transcription to occur, DNA must be sufficiently exposed and accessible for RNA polymerase activity. For DNA to be permissive to transcription, the histone tails must be acetylated by histone acetyltransferases, and chromatin-remodeling complexes must weaken the histone–DNA contacts. There are two classes of chromatin-remodeling complex. The SWI and SNF proteins (as a SWI/SNF complex) possess a

DNA-stimulated ATPase activity that destabilizes histone–DNA interactions that facilitate the release of the histone particle. The NURF proteins complex contains an ATPase that catalyzes nucleosome sliding that releases the histone–DNA contacts only enough to allow the histone particle to slide out of the way (**Figure 18.38**).

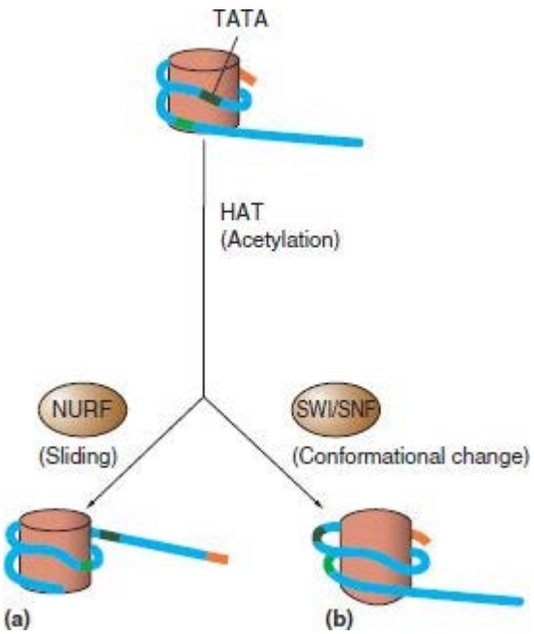


FIGURE 18.38

Chromatin Remodeling

The acetylation by histone acetyltransferase (HAT) of the histone tails breaks their contacts with DNA. The core histones are then released from DNA contacts by sliding out of the way through the action of NURF proteins (a) to expose the promoter region on the DNA to the transcriptional machinery or by a localized conformational change in chromatin (b) promoted by the SWI/SNF chromatin-remodeling complex.

Principally because of the complexity of eukaryotic genomes, eukaryotic DNA transcription is not understood as completely as the prokaryotic process. However, the eukaryotic process is known to possess several unique features.

RNA POLYMERASE ACTIVITY Eukaryotes possess three nuclear RNA polymerases, each of which differs in the type of RNA synthesized, subunit structure, and relative amounts. RNA polymerase I (RNAP I), localized within the nucleolus, transcribes the larger rRNAs (28S, 18S, and 5.8S). The precursors of mRNA and most miRNAs and snRNAs are transcribed by RNA polymerase II (RNAP II), and RNA polymerase III (RNAP III) is responsible for transcribing the precursors of the tRNAs, 5S rRNA, U6 snRNA, snoRNAs and lncRNAs. Approximately 80% of a cell's RNA is synthesized by RNAPs I and III.

Each RNA polymerase possesses two large subunits and several (8–12) smaller subunits. For example, the two large subunits of RNAP II have molecular weights of 215 and 139 kDa. The number of smaller subunits varies among species; for example, plants possess eight, whereas vertebrates have six. Some of the smaller subunits are also present in the other two RNA polymerases. In contrast to the prokaryotic RNA polymerase, the eukaryotic enzymes cannot initiate transcription themselves. Various transcription factors must be bound at the promoter before transcription can begin.

RNA polymerases function within discrete clusters where active genes come together via chromatin looping to be transcribed. These clusters, called *transcription factories*, are located in

interchromatin channels (p. 652) throughout the nucleus. The number of transcription factories varies with species and cell type. The nuclei of HeLa cells (epithelial cells from an immortal cell line used in research), for example, contain about 10,000 factories, of which 8000 are RNAPII factories and 2000 are RNAPIII factories. The nucleolus is considered a large RNAPI factory. RNAPII transcription factories, each of which contains an estimated eight polymerase molecules, have diameters ranging from 40 to 180 nm (depending on the level of transcriptional activity). Each RNAPII-associated factory has a protein-rich core, with active RNA polymerase enzyme complexes attached to its surface.

EUKARYOTIC PROMOTERS Eukaryotic promoters are more complex and variable than those of prokaryotes. Each consists of a *core promoter*, the minimum DNA sequence required for RNA polymerase binding and transcription initiation, and additional proximal (nearby) and distal (distant) DNA sequences that contribute to transcription regulation.

There are two classes of eukaryotic core promoters: focused and dispersed. Focused promoters contain a *transcription start site* (TSS) that consists of a single nucleotide or a very narrow range of nucleotides and a variable set of sequence motifs, referred to as *core promoter elements* (CPEs), which allow binding of the transcription machinery. The TATA box (typically 25–100 bp upstream of a gene's TSS), originally identified in yeast, is the best-researched eukaryotic CPE. Found only in about 10% of mammalian core promoters, the AT-rich TATA box is recognized and bound by TATA-binding protein (TBP), a subunit of the transcription factor TFIID. Other examples of CPEs found in focused core promoters include Inr (initiator), BRE (B recognition element), MTE (motif ten element), and DPE (downstream promoter element) (**Figure 18.39**). Dispersed core promoters are so named because they possess many TSSs (usually one strong and several weak start sites), which are distributed seemingly at random over a broad region of 50 to 100 bp. Commonly found in vertebrates and especially in mammals, dispersed core promoters typically occur within CpG islands (p. 662) and are mainly associated with housekeeping genes (p. 654). Unmethylated CpGs are now believed to facilitate transcription by destabilizing nucleosomes.

Proximal promoter elements are transcription factor-binding sites up to about 250 bp upstream of the TSS. Examples of proximal promoter elements include the GC and CCAAT boxes, both of which enhance transcription when bound to certain transcription factors. For example, transcription rates increase when the transcription factor Sp1 binds to GC box elements upstream of genes that code for proteins such as the insulin receptor (p. 611) and HMG-CoA reductase (p. 476).

Distal regulatory sequences are distance-independent DNA elements that can be located upstream or downstream or within the introns of their target genes. Distal elements can increase (*enhancers*) or decrease (*silencers*) transcription. Transcription is activated when an activator protein binds to an enhancer element, and it is inhibited when a repressor protein binds to a silencer element.

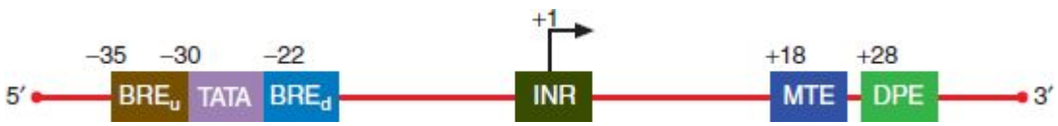


FIGURE 18.39

The Eukaryotic RNAPII Core Promoter

The core promoter, the minimum DNA sequence required for RNA polymerase binding and transcription initiation for a specific gene, often contains sequence motifs that researchers have identified as core promoter

elements (CPEs). The TATA box is a well-researched CPE. BRE_u and BRE_d are CPEs that, when present, are located immediately upstream and downstream of the TATA box, respectively. Together or separately, BRE_u and BRE_d modulate transcription activation via their interactions with TFIIB. The majority of genes lack a TATA box. Instead, they may contain Inr (initiator element) or DPE (downstream core promoter element). MTE (motif 10 element), acting in conjunction with Inr, promotes the binding of TFIID to the core promoter. No CPE is present in all eukaryote core promoters, and some core promoters have no known CPEs. Note that when it is present in a core promoter, Inr contains the transcription start sequence.

When enhancer or silencer elements come into contact with an RNA polymerase complex at a promoter, the intervening DNA sequences (as long as 100 kbp) bend into loops stabilized by cohesin (p. 727). Recall that the interactions of enhancers and promoters can be blocked by boundary DNA elements called *insulators*. In vertebrates, CTCF interacts with cohesin at insulator elements.

RNA POLYMERASE II AND THE EUKARYOTIC TRANSCRIPTION PROCESS DNA-directed RNA polymerase II (RNAP II) is the most investigated type of RNA polymerase in eukaryotes, largely because of its role in mRNA synthesis. Roger Kornberg (Nobel Prize in Chemistry, 2006) determined the structural and functional properties of yeast RNAP II.

The core RNAP II enzyme contains 12 subunits in humans. RBP1, the largest subunit, forms part of the enzyme's active site, the groove that binds DNA. RBP1 also contains a C-terminal domain (CTD) with between 25 and 52 repeats of the heptad YSPTSPS (52 in humans). The CTD's capacity to act as a docking site for proteins involved in different phases of transcription is regulated by reversible covalent modifications of various heptad residues (primarily phosphorylation of Ser2 and Ser5). RNAP II can bind to promoters when the heptads are unphosphorylated. As transcription progresses, a pattern of phosphorylated Ser2 and Ser5 residues created by kinases and phosphatases allows the recruitment of specific RNA processing enzymes. During transcription, the pattern of serine phosphorylation changes from high Ser5P and low Ser2P levels near the beginning to the reverse further downstream. Examples of processes affected by CTD serine phosphorylation patterns include the conversion of RNAP II into a processive enzyme, histone remodeling, mRNA 5′-cap formation, and splicing initiation (p. 728). By the time RNAP II begins synthesizing the poly(A) tail (p. 727), all Ser5 residues are unphosphorylated. Complete CTD dephosphorylation allows RNAP II disengagement and its subsequent recycling in another round of transcription.

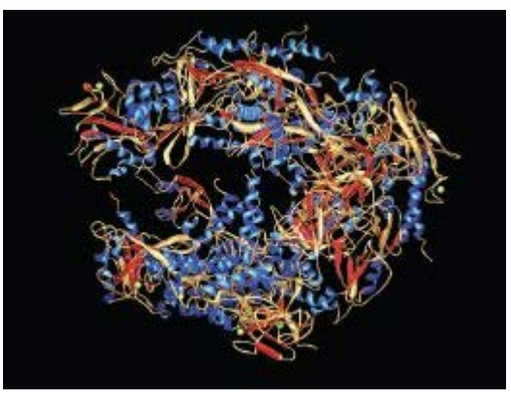


FIGURE 18.40

RNA Polymerase II Core Enzyme

The 12 individual subunits are distinguished by colors.

In addition to the core enzyme (**Figure 18.40**), the RNAP II transcription machinery (3000 kDa) consists of a set of six transcription factors referred to as *general transcription factors* and a multisubunit protein complex called Mediator. The general transcription factors (GTFs) TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH are the minimum number of additional proteins that are necessary for accurate transcription. They facilitate construction of the transcription **preinitiation complex** (PIC), recognition of the promoter, and the ATP- dependent unwinding of DNA.

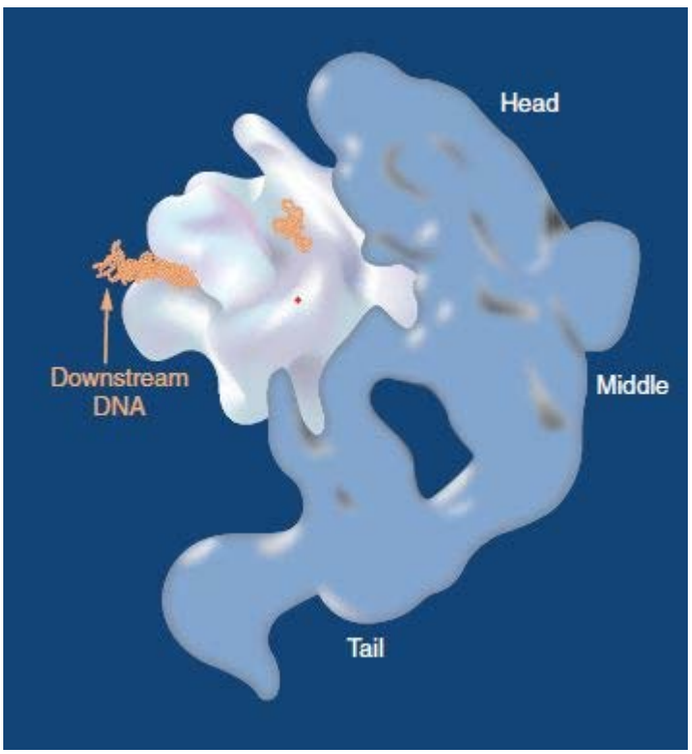


FIGURE 18.41

The Yeast Mediator–RNA Polymerase II Holoenzyme Complex

The RNAP II preinitiation complex (RNAP II with associated transcription factors), shown in white, interacts with the crescent-shape Mediator (shown in blue) via its head and middle domains. The tail domain of Mediator interacts with a regulatory protein that is bound to DNA.

Mediator is a protein complex required for the transcription of almost all RNAP II promoters. Human mediator is a 26-subunit complex with a molecular mass of 1.2 MDa and a structure that consists of three domains: head, middle, and tail (**Figure 18.41**). Mediator is essentially a signal integration platform that acts as an adaptor between RNAP II and the transcription factors that are bound at positive (enhancers) and negative (silencers) regulatory DNA sequences, which may be some distance away from the gene(s) they modulate (**Figure 18.42**). **Figure 18.43** illustrates the PIC bound to mediator in the context of a DNA loop.

Eukaryotic transcription occurs in several phases: PIC assembly, initiation, elongation, and termination. PIC assembly (**Figure 18.44**) begins with the binding of the TATA-binding protein (TBP) subunit of TFIID to the TATA box, a process that is facilitated by TFIIA. TBP is a saddle-shape protein that causes strand separation by distorting DNA structure. The insertion of four phenylalanine side chains between base pairs in the minor groove causes the formation of a bend in the DNA segment. As the DNA bends its interactions with TBP increase, thus enhancing the shape of the bend to approximately 90°. The subsequent binding of the other GTFs followed by Mediator yields the transcriptionally active PIC. Transcription bubble formation is catalyzed by

the ATPase and helicase activities of TFIIH subunits, which create negative supercoil tension. The bubble is maintained in the open position by TFIIF, which binds to the coding strand. The noncoding strand then enters the active site of RNAP II. Once the first bond is synthesized, the transcription process continues until the mRNA reaches about 23 nt, where it often enters a paused state. The binding of cohesin at the promoter is believed to stimulate the transition to the elongation phase of transcription. This phase begins with promoter clearance and the release of the GTFs TFIIE and TFIIH.

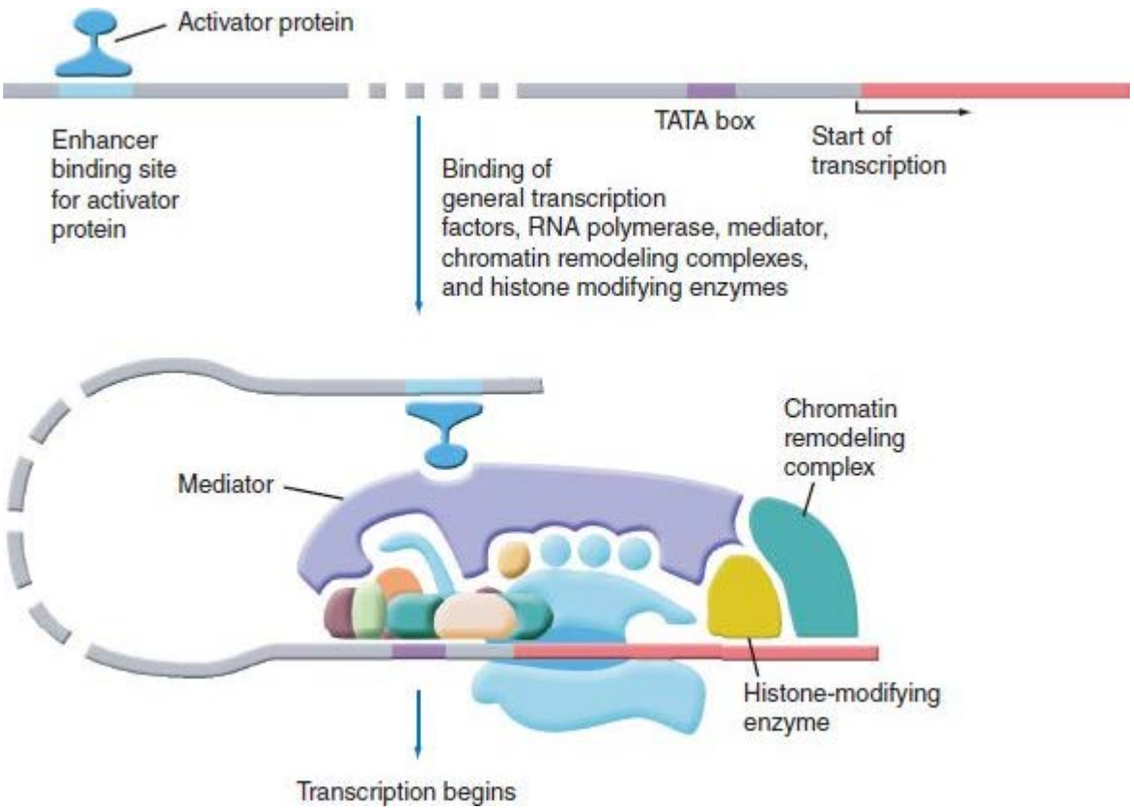


FIGURE 18.42

Transcription Activation by Activator Proteins

The interaction between an activating transcription factor bound to an enhancer element distal from the core promoter (indicated by the dashed lines) and Mediator facilitates the recruitment of RNAP II and the GTFs to the core promoter. Mediator binds to the CTD of RNAP II. Chromatin remodeling complexes and histone-modifying enzymes are also recruited to the PIC. Although only one activator–mediator interaction is illustrated, the transcription of most genes involves several activator proteins that together specify the rate of the process. Mediator can also act to repress transcription when it interacts with repressor proteins (not shown) bound to silencers.

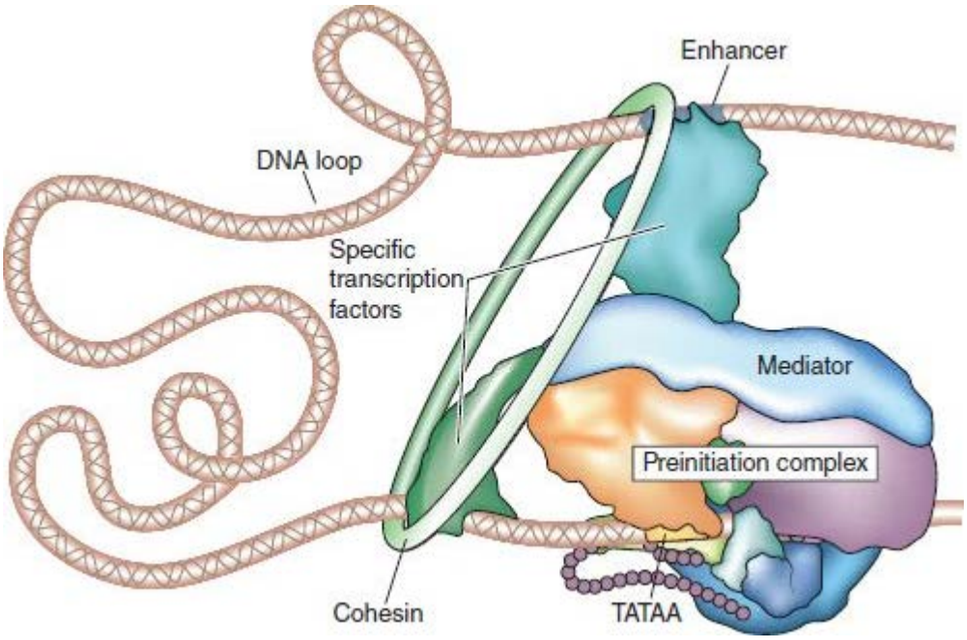


FIGURE 18.43

DNA Looping and Gene Activation

A DNA loop forms between an enhancer element bound to an activator transcription factor and a distant core promoter site via linkage to a PIC/ Mediator complex. Cohesin is a ring-like complex that stabilizes the loop. In some cases of enhancer-promoter DNA loops, the insulator protein CTCF dimer (not shown) binds in close proximity to cohesin to block the interaction of enhancers in one DNA domain with promoters in another. The DNA loops created by Mediator and cohesin are different in different cells because they co-occupy different enhancers and promoters as a result of the unique gene expression pattern of each cell type.

Once promoter clearance has been achieved, the RNAP II holoenzyme dissociates from Mediator and the elongation phase of transcription begins. The RNAP II complex continues the transcription process well past the functional end of the nascent transcript until a signal sequence (5′-AAUAAA-3′) called a *poly(A) signal* is reached. As soon as the poly(A) signal is transcribed, several proteins linked to the RNAP II CTD transfer to the poly(A) signal and cause termination by recruiting an endonuclease that cleaves the transcript about 10–30 nt downstream of the poly(A) signal.

RNA PROCESSING In contrast to prokaryotic mRNA, which usually requires little or no processing, eukaryotic mRNAs and many lncRNAs are the products of extensive posttranscriptional processing. Throughout processing, which begins shortly after transcription initiation, the nascent mRNA transcripts, called pre-mRNAs, become associated with about 20 different types of nuclear protein in ribonucleoprotein particles (hnRNP). Shortly after the transcription of the primary transcript begins, a modification of the 5′ end called *capping* occurs. The cap structure (**Figure 18.45**), which consists of 7-methylguanosine (m^7G) linked to the mRNA through a triphosphate linkage, is synthesized when the pre-mRNA is about 30 nt long. This linkage protects the 5′ end from exonucleases, facilitates transfer into the cytoplasm, and promotes mRNA translation by ribosomes.

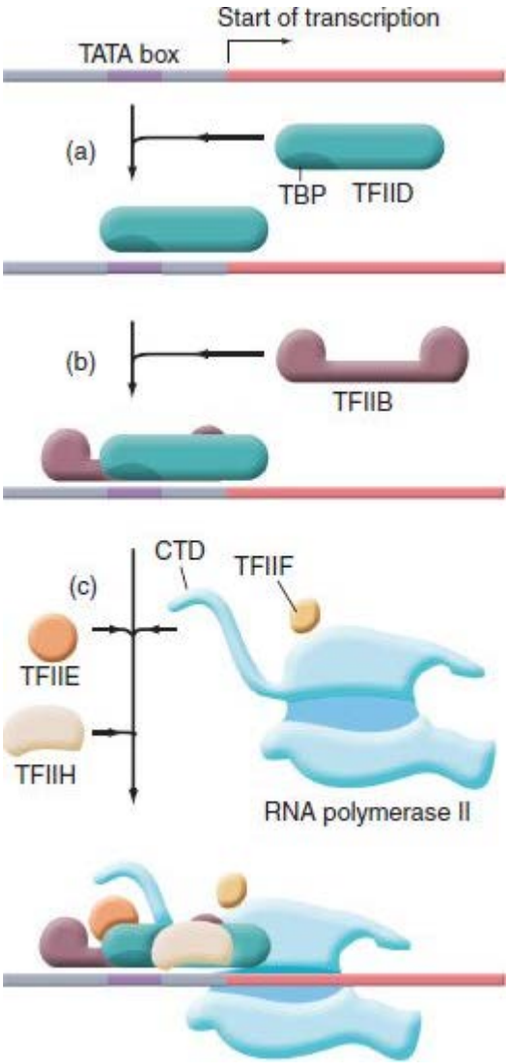


FIGURE 18.44

Preinitiation Complex Formation at a TATA Box

(a) The DNA-distorting TBP subunit of the GTF TFIID recognizes and binds to the TATA box. TFIIA (not shown) stabilizes the interaction between TFIID and the TATA box. (b) The binding of TFIID enables the subsequent binding of TFIIB. (c) The other GTFs (TFIIE, TFIIF, and TFIIH) and RNAP II assemble at the promoter to form an active preinitiation complex. Note that the mediator complex (not shown) controls PIC formation.

With the exception of the replication-dependent histone mRNAs of multicellular animals, all eukaryotic mRNAs are linked to a polyadenylate strand (100–250 adenylate residues) called a *poly(A) tail*. Transcript termination by RNAP II and poly(A) addition are independent events, although the enzyme complex involved in the process is located at the polymerase's CTD tail. A protein complex containing CPSF (cleavage polyadenylation specificity factor) cleaves a nucleotide sequence at the mRNA's 3′-end. The poly(A) tail is then synthesized by a poly(A) tail polymerase and then covalently linked to the 3′-terminus of the mRNA transcript. The poly(A) tail facilitates transport of mRNAs out of the nucleus and translation on ribosomes. In most cells, the poly(A) tails of mRNAs in the cytoplasm gradually get shorter, with the result that they are translated less and become vulnerable to degradation. The mRNAs of many protein-coding genes have more than one site at which the poly(A) tail is added. Consequences of *alternative polyadenylation* (different 3′-UTR sequences) include the synthesis of different proteins (if the coding region is altered) and different binding sites for miRNAs.

One of the more remarkable features of eukaryotic RNA processing is the removal of introns from RNA transcripts. In this process, called **RNA splicing**, introns are cut out of the primary transcript, and the exons are linked together to form a functional product. Most research efforts have been concerned with the splicing of pre-mRNAs, which is now described. The number of introns in the protein-coding eukaryotic genes varies widely, from one in the intron-containing genes in lower eukaryotes such as yeast to dozens or even hundreds in some mammalian genes. The human dystrophin gene, the largest in humans, is 2.4 million base pairs (Mb) long and contains 79 exons. (Synthesis of the dystrophin primary transcript takes 16 h!) The dystrophin-spliced mRNA (14 kb) codes for a structural protein (427 kDa), with 3600 amino acid residues that connects the muscle cell cytoskeleton with the plasma membrane and the extracellular matrix. Several mutations in the dystrophin gene cause the X-linked recessive disorder, Duchenne muscular dystrophy, which presents as progressive muscle degeneration.

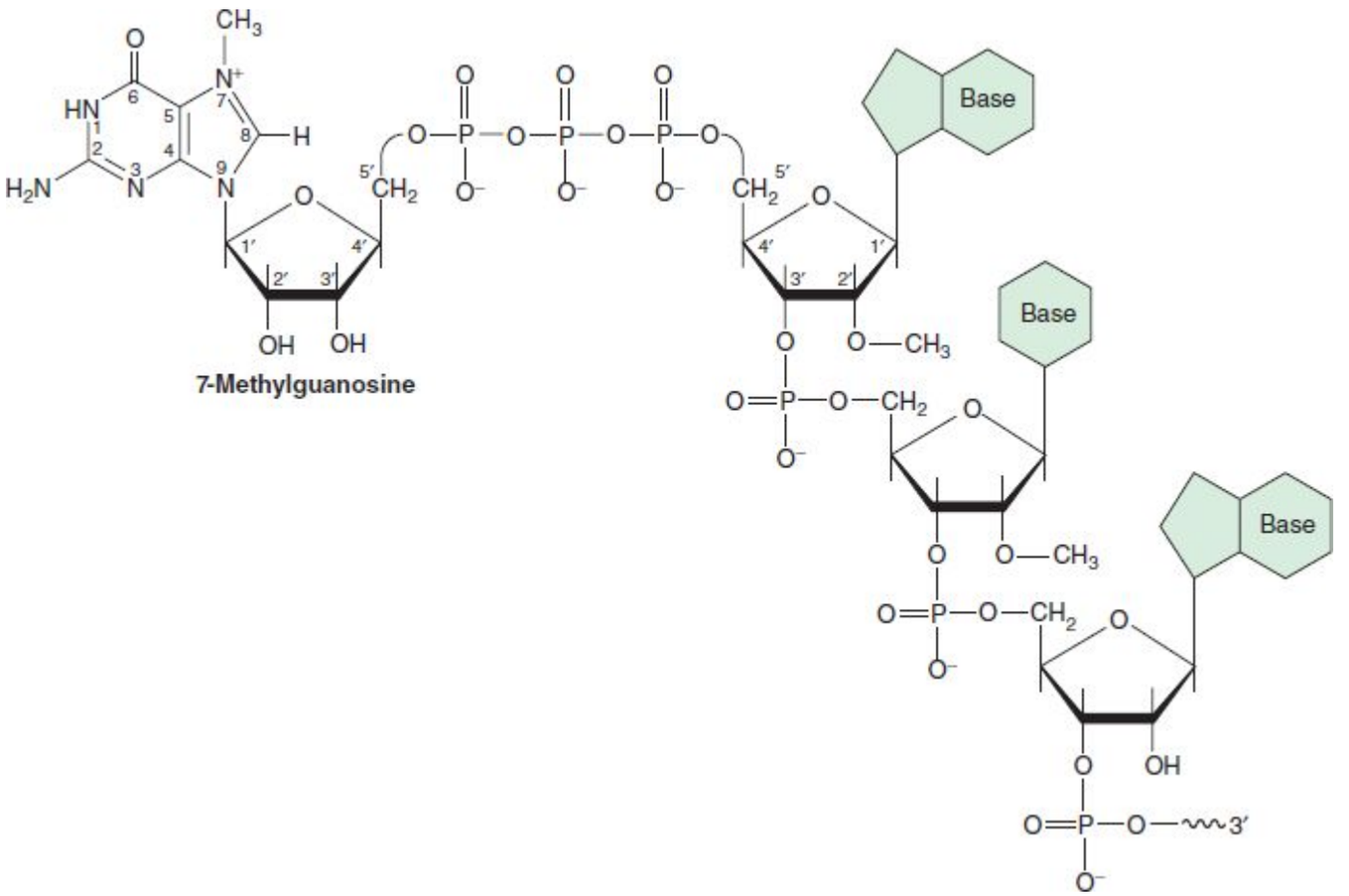


FIGURE 18.45

The Methylated Cap of Eukaryotic mRNA

The cap structure consists of a 7-methylguanosine attached to the 5′ end of an RNA molecule through a unique 5′ → 5′ linkage. The 2′-OH of the first two nucleotides of the transcript are methylated. The 5′ → 5′ bond formation is catalyzed by a guanyl transferase; N-7 methylation is catalyzed by guanine methyl transferase.

RNA splicing occurs in close proximity to transcription in a multi-megadalton (MDa) RNA-protein complex called a **spliceosome** formed from the snRNAs U1, U2, U4, U5, and U6 (p. 674). Recent research has revealed that yeast and human spliceosomes, within nuclear speckles in close proximity to RNA polymerases, undergo seven major rearrangements during splicing. As the process unfolds more than 100 proteins and the five snRNP complexes come and go. Despite this

complexity, the essential splicing features may be described as follows. In eukaryotic nuclear pre-mRNA transcript processing most splicing events occur at GU-AG sequences. In GU-AG introns, 5′-GU-3′ and 5′-AG-3′ are the first and last dinucleotide sequences, respectively, of the intron. The splicing event is initiated by complementary binding between the intron's 5′-splice site sequence and the 5′ terminal sequence of the snRNA U1. This is followed by the binding of a U2 sequence in the U2 snRNP to the "branch site," a 2′-OH of an adenosine nucleotide within the intron. The other spliceosome components (U4/U6 and U5 RNPs) then bind to form the entire spliceosome with the U5 sequence binding upstream of the 5′ splice site. Splicing (**Figures 18.46** and **18.47**) is composed of two reactions:

Duchenne Muscular Dystrophy

1. A 2′-OH of the "branch site" attacks a phosphate in the 5′ splice site in an S_N2 transesterification reaction. This reaction, which is accompanied by the displacement of the U1 and U6 RNPs and the release of the U1 and U4 RNP, results in the formation of a loop called a *lariat* because of the newly created 5′ → 2′ phosphodiester bond.

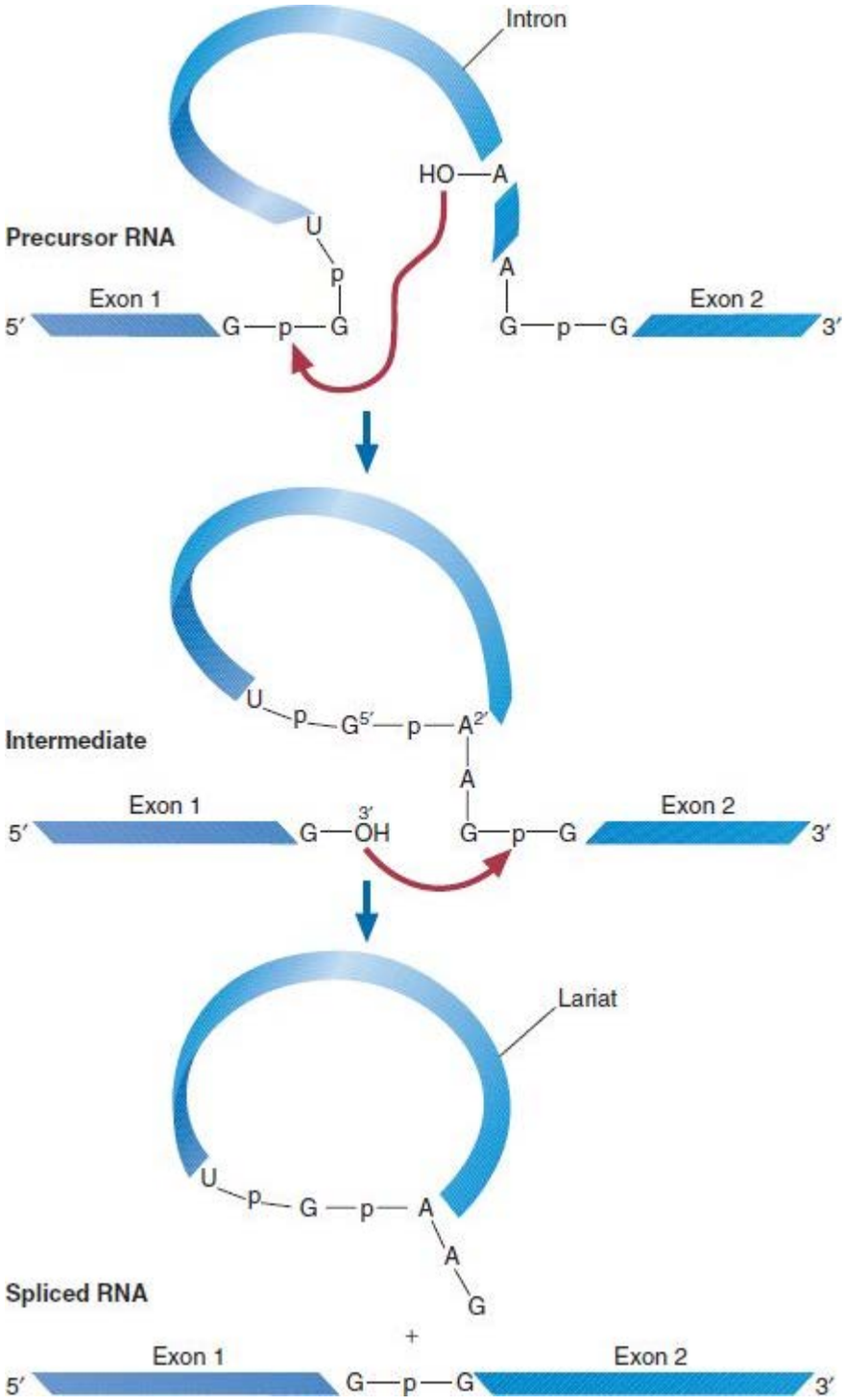


FIGURE 18.46

RNA Splicing

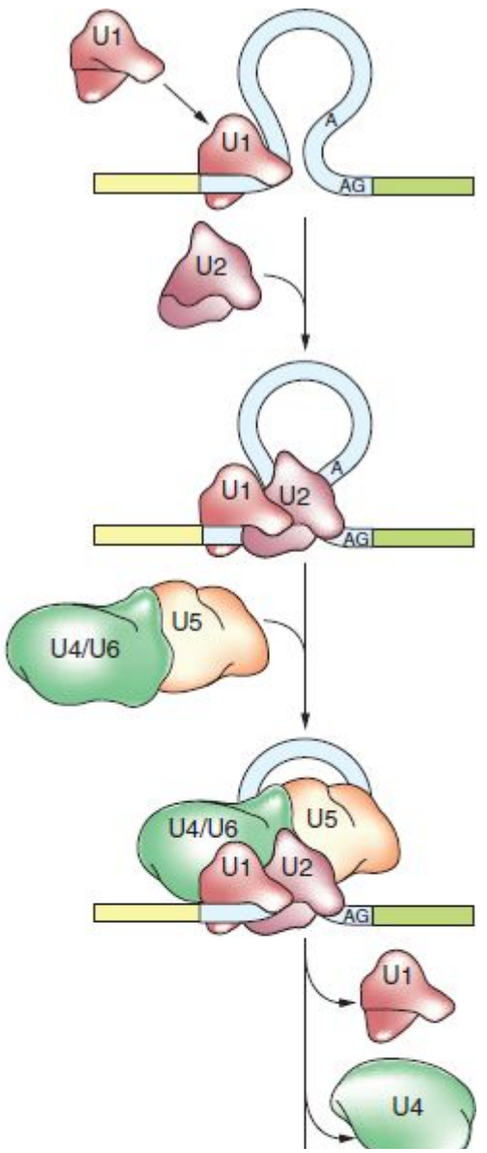
mRNA splicing, which occurs in the nucleus, begins with the nucleophilic attack of the 2′-OH of a specific adenosine on a phosphate in the 5′-splice site. A lariat is formed by a 2′,5′-phosphodiester bond. In the next step, 3′-OH of exon 1 (acting as a nucleophile) attacks a phosphate adjacent to the lariat. This reaction releases the intron and ligates the two exons.

2. The lariat is cleaved and degraded, and the two exons are joined when the 3′-OH, which is bound to U5 of the upstream exon, attacks a phosphate that is adjacent to the lariat.

KEY CONCEPTS

- Transcription in eukaryotes is significantly more complex than its counterpart in prokaryotes.
- In addition to chromatin-remodeling and RNA-processing reactions, gene transcription requires the binding of unique sets of transcription factors to promoter sequences.
- Eukaryotic RNA transcripts undergo several processing reactions.

As a splicing event ends, an *exon junction complex* (EJC) binds to each splice site 20 nt upstream from the exon–exon junction to form the mature messenger ribonucleoprotein (mRNP). EJCs remain associated with mRNAs during nuclear export and cytoplasmic localization until the first round of translation by a ribosome. EJCs play a role in *nonsense-mediated decay* (NMD), an mRNA surveillance mechanism that detects premature termination signals (stop codons) that result in the synthesis of truncated and possibly toxic polypeptides. Caused by RNA splicing errors (as well as random mutations and faulty DNA rearrangements), premature stop codons are distinguished from normal stop codons by their position relative to an EJC. As the mRNA is translated by a ribosome, detection of a stop codon upstream of an EJC (indicating there is another stop codon downstream) causes the surveillance mechanism to initiate the mRNA decay process. NMD occurs within *P bodies*, cytoplasmic structures that consist of enzymes involved in several mRNA degradation or silencing processes (see p. 738).



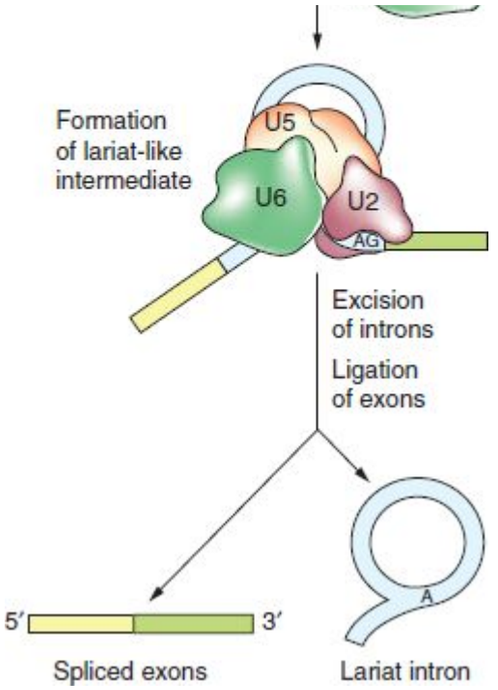


FIGURE 18.47

The Spliceosome

The lariat intermediate forms as a result of the cut made at the 5′-splice site, and the 5′-end of the intron is then connected to the conserved adenine. During this process, U6 dissociates from U4, both U1 and U4 are released from the splicing complex, and a U6/U2 interaction occurs. U5 bound to the 3′-splice site positions 5′-exon close to the 3′-exon so they can be readily ligated.

The majority of mammalian pre-mRNAs, most of which contain multiple introns, are processed by **supraspliceosomes** (21.1 MDa), each formed from four active spliceosomes (**Figure 18.48**). It has been suggested that the supraspliceosome increases the speed and efficiency of transcript splicing and provides opportunities for intron excision proofreading.

Processing of the replication-dependent histone mRNAs of multicellular animals is different in several respects from that of other mRNAs. The histone mRNA transcript has a 5′-cap but no introns or poly(A) tail. Instead of a poly(A) tail, there is an evolutionarily conserved 25-nt stem-loop at the 3′-end that binds to SLBP (stem-loop binding protein). SLBP facilitates histone mRNA transport, translation, and degradation. The histone mRNA stem-loop is essential for the rapid degradation of histone mRNAs when DNA synthesis ends at the conclusion of the S phase of the cell cycle.

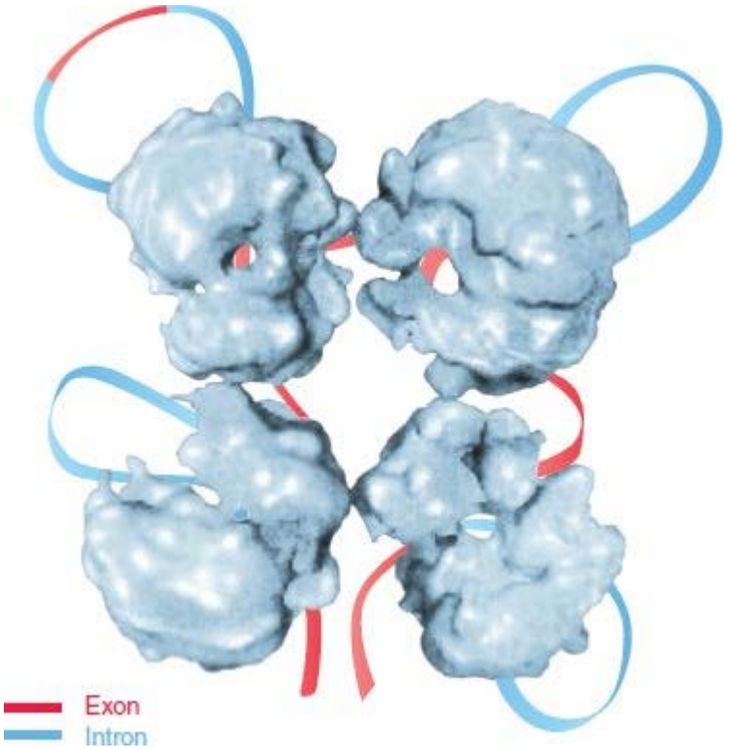


FIGURE 18.48

The Supraspliceosome

In this model of supraspliceosome structure and function, a complex formed by four spliceosomes processes a pre-mRNA. Loops of the introns (blue) being spliced extend outward. Note that an alternative exon (red) (see p. 736) is illustrated in the upper left-hand corner.

Biochemistry IN THE LAB

CRISPR

Why is CRISPR technology, derived from a family of protective DNA sequences within prokaryotes, now recognized as one of the powerful tools in life science research? **CRISPR** (*c*lustered *r*egularly *i*nterspersed *s*hort *p*alindromic *r*epeats), a powerful genome-editing technology now widely used in biomedical research, is adapted from a primitive prokaryotic defense mechanism (a form of acquired immunity). A diverse group of prokaryotes (bacteria and archaea) can recognize and eliminate foreign DNA by inserting fragments of invading viral or plasmid DNA into a genome location called the *CRISPR array* (**Figure 18H**) using certain *Cas* (*C*RISPR-*as*sociated) *proteins* encoded by the cas gene operon. Each organism's CRISPR array is composed of short spacers (DNA sequences from previous foreign DNA exposures) between repeat sequences. The cell can detect subsequent invasions by using Cas proteins with endonuclease activity and guide RNAs, referred to as CRISPR RNA (*crRNA*) that are transcribed from each spacer. In the bacterial type II CRISPR system, the most common one used in scientific genome editing, a second RNA called *tracrRNA* is also transcribed from the CRISPR array. Cas9 is the endonuclease that is used to cleave DNA sequences. A cellular RNase processes the primary crRNA transcript and tracrRNAs to yield crRNA/tracrRNA complexes. Each complex

subsequently associates with a Cas9 protein to form an active ribonucleoprotein complex. The cas9/RNA complex proceeds to interrogate the cell's DNA. When a cas9/RNA complex finds the appropriate sequence motif (i.e., the crRNA sequence matches the interrogated DNA sequence), each of the two DNA strands is cleaved by a different domain in the Cas9 protein. A DSB in the viral or plasmid DNA sequence effectively destroys its infecting capability.

The CRISPR/Cas9 genome-editing technique is a simplified version of the type II CRISPR system. It includes Cas9 and a single guide RNA (*sgRNA*), an engineered fusion of tracrRNA and crRNA. CRISPR editing can be used to make deletions in a genome or insert new DNA sequences. An sgRNA is synthesized to line up with a specific target sequence in the cell's genome. After the sgRNA directs the Cas9 complex to a matching segment of DNA, the Cas9 endonuclease will cut the DNA.

A cell's own DNA repair machinery can be used to add or delete small pieces of DNA, or, by engineering an sgRNA with multiple crRNAs, an existing segment can be removed and then replaced with a customized sequence. For example, a defective copy of a gene (e.g., a protein-coding gene) can be replaced with a normal version.

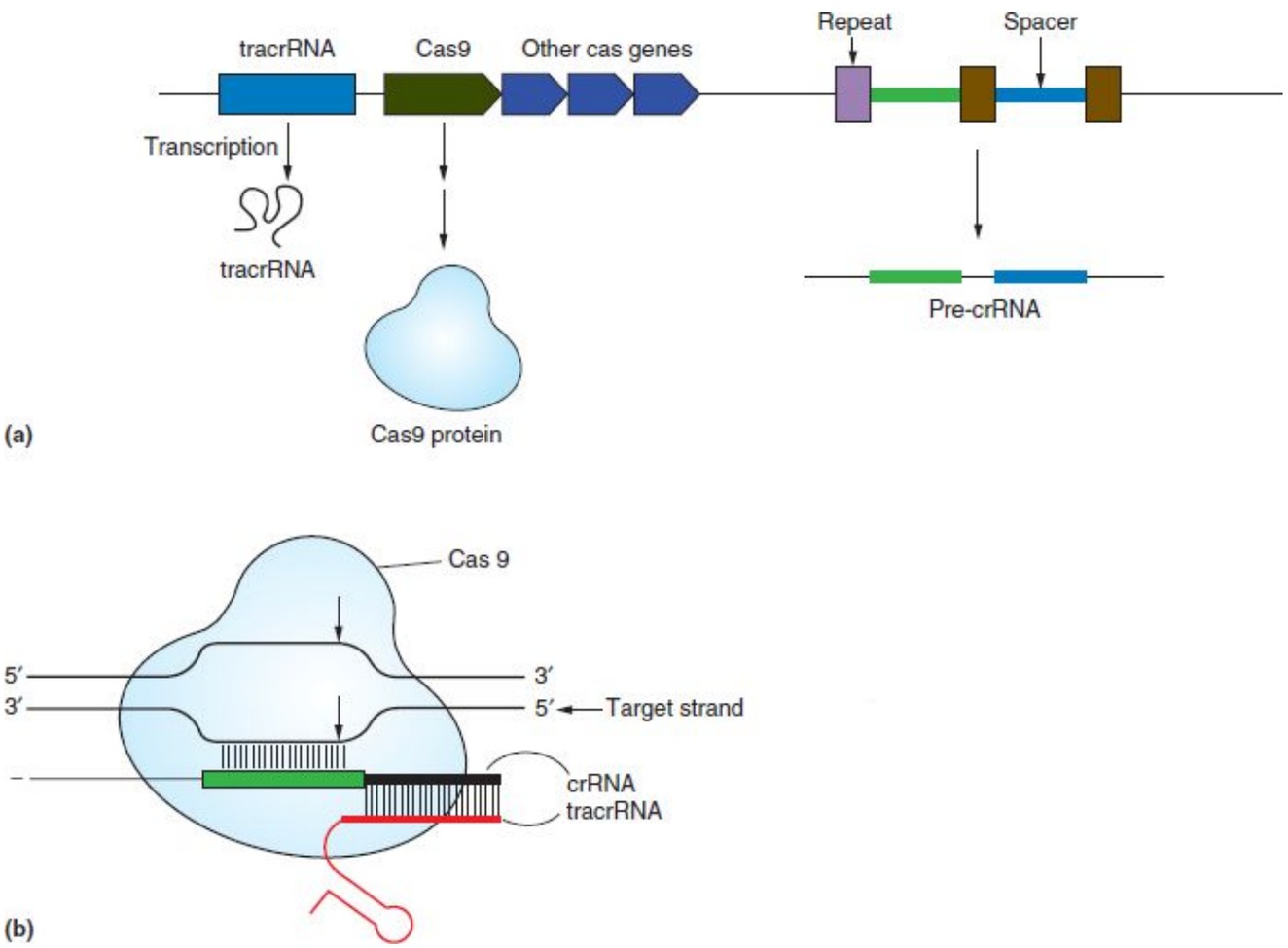


FIGURE 18H

Bacterial CRISPR Type II System

(a) This simplified diagram of a CRISPR locus shows a repeat-spacer array, a series of Cas genes (including that of the endonuclease Cas9), and a tracrRNA gene. (b) The crRNA in the Cas9/cr/tracr complex is bound via complementary base pairing to a DNA target strand. As a result, Cas9 will cleave both DNA strands (indicated by arrows), thus eliminating the threat from a foreign DNA molecule.

SUMMARY: The CRISPR editing system has revolutionized molecular biology research. One of CRISPR's many current benefits is its use in the investigation of the genetic causes and potential treatment of human disease using human cells and animal models.

18.3 GENE EXPRESSION

Ultimately, the internal order most essential to living organisms requires the precise and timely regulation of gene expression. It is, after all, the capacity to switch genes on and off that enables cells to respond efficiently to a changing environment. In multicellular organisms, complex programmed patterns of gene expression are responsible for cell differentiation and intercellular cooperation.

The regulation of genes, as measured by their transcription rates, is the result of a complex hierarchy of control elements that coordinate the cell's metabolic activities. Some genes, referred to as housekeeping genes (p. 654), are routinely transcribed because they code for gene products routinely required for cell function. In the differentiated cells of multicellular organisms, certain specialized protein products of transcribed genes are produced that cannot be detected elsewhere (e.g., hemoglobin in red blood cells). Genes that are expressed only in certain circumstances are referred to as *inducible* (p. 733). For example, the enzymes required for lactose metabolism in *E. coli* are synthesized only when lactose is actually present and glucose, the bacterium's preferred energy source, is absent.

In general, prokaryotic gene expression involves the interaction of specific proteins with DNA in the immediate vicinity of a transcription start site. Such interactions may have either a positive effect (i.e., transcription is initiated or increased) or a negative effect (i.e., transcription is blocked). In an interesting variation, the inhibition of a negative regulator (called a *repressor*) activates affected genes. (The inhibition of a repressor gene is referred to as derepression.) Eukaryotic gene expression uses these mechanisms and several others, including gene rearrangement and amplification, epigenetic mechanisms, and various complex transcriptional, RNA processing, and translational controls. In addition, the spatial separation of transcription and translation inherent in eukaryotic cells provides another opportunity for regulation: mRNA transport control.

This section describes several examples of gene expression control. The discussion of prokaryotic gene expression focuses on operons and riboswitches. An **operon** is a set of genes under the regulation of the same operator and promoter(s). The **operator** is a regulatory sequence that binds to specific repressor or activator proteins that modulate gene expression. The *lac* operon in *E. coli* is the most thoroughly researched example. Bacteria also employ an RNA-based control mechanism called the *riboswitch*, made up of a specific untranslated sequence within the mRNA along with the small metabolite to which it binds. The action of the riboswitch usually represses gene expression by terminating transcription, blocking translation, or causing mRNA self-destruction. Eukaryotic gene expression is less understood than that of prokaryotes such as *E. coli* because of larger genomes and more varied and complex regulatory mechanisms. The section ends with a brief overview of growth factor–triggered gene expression.

Gene Expression in Prokaryotes

The highly regulated metabolism of prokaryotes such as *E. coli* allows these organisms to manage limited resources and to respond to a changing environment. The timely synthesis of enzymes and other gene products only when needed, combined with the rapid destruction of mRNAs by ribonucleases (RNases), prevents dissipation of energy and nutrients. This flexibility is made possible at the genetic level where the control of inducible genes is often effected by groups of linked structural genes and regulatory genes called operons. Investigations of operons, especially the *lac* operon, have provided substantial insight into how gene expression can be altered by environmental conditions. The *lac* operon and several types of riboswitch are described.

THE *LAC* OPERON The *lac* operon (**Figure 18.49**) consists of a control element and structural genes that code for the enzymes of lactose metabolism. The control element contains the promoter site, which overlaps the operator site. The *catabolite gene activator protein* (CAP) site (described shortly), a 16-bp DNA sequence upstream of the promoter, is another regulatory element. The structural genes Z, Y, and A specify the primary structure of *β*-galactosidase, lactose permease, and thiogalactoside transacetylase, respectively. *β*-Galactosidase catalyzes the hydrolysis of lactose, which yields the monosaccharides galactose and glucose, whereas lactose permease promotes lactose transport into the cell. The role of thiogalactoside transacetylase in lactose metabolism is unclear, but it is known that this enzyme transfers an acetyl group from acetylCoA to the 6-O-methyl group of the *β*-glycosides of sugars such as glucose, galactose and lactose. It is believed that this reaction facilitates the transport of nonmetabolizable glycosides out of cells, since lactose metabolism proceeds normally without it. A repressor gene *lacI*, directly upstream of the *lac* operon, codes for the *lac* repressor protein, a tetramer that binds to the operator site with high affinity. (There are about 10 copies of *lac* repressor protein per cell.) The binding of the *lac* repressor to the operator prevents the functional binding of RNA polymerase to the promoter (**Figure 18.50**).

Without its inducer (allolactose, a *β*-1,6-isomer of lactose) the *lac* operon remains repressed because the *lac* repressor binds to the operator. When lactose becomes available, a few molecules are converted to allolactose by *β*-galactosidase. Allolactose then binds to the repressor, changing its conformation and promoting dissociation from the operator. Once the inactive repressor has diffused away from the operator, the transcription of the structural genes begins. The *lac* operon remains active until the lactose supply is consumed or until glucose, the preferred energy source, becomes available. Then the repressor reverts to its active form and rebinds to the operator.

If the *lac* operon is repressed in the absence of lactose, what is the source of allolactose? Transcription is never completely blocked, since the repressor protein occasionally detaches, with resulting synthesis of a low number of operon-coded proteins. So when the bacterial cell encounters lactose, there are a few molecules of lactose permease available to facilitate the transport of lactose into the cell, where it is converted to allolactose.

Glucose is the preferred carbon and energy source for *E. coli*. If both glucose and lactose are available, the glucose is metabolized first. Synthesis of the *lac* operon enzymes is induced only after the glucose has been consumed. (This makes sense because glucose is more commonly available and has a central role in cellular metabolism. Why expend the energy to synthesize the enzymes required for the metabolism of other sugars if glucose is also available?)

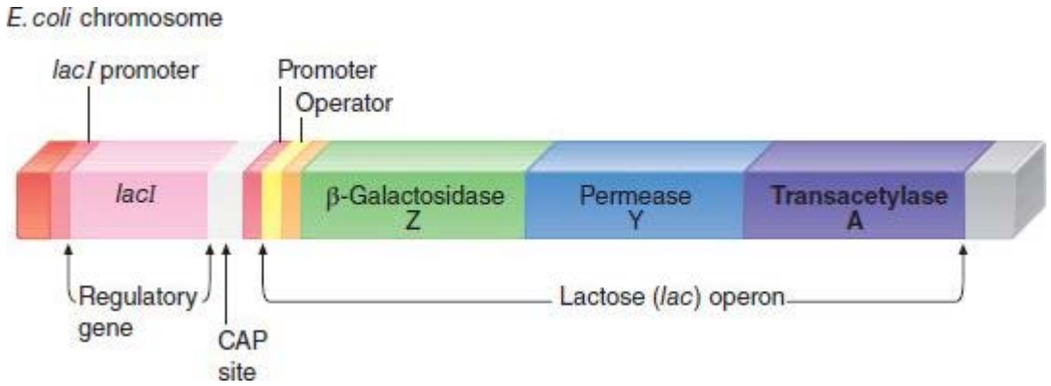


FIGURE 18.49

The *lac* Operon in *E. coli*

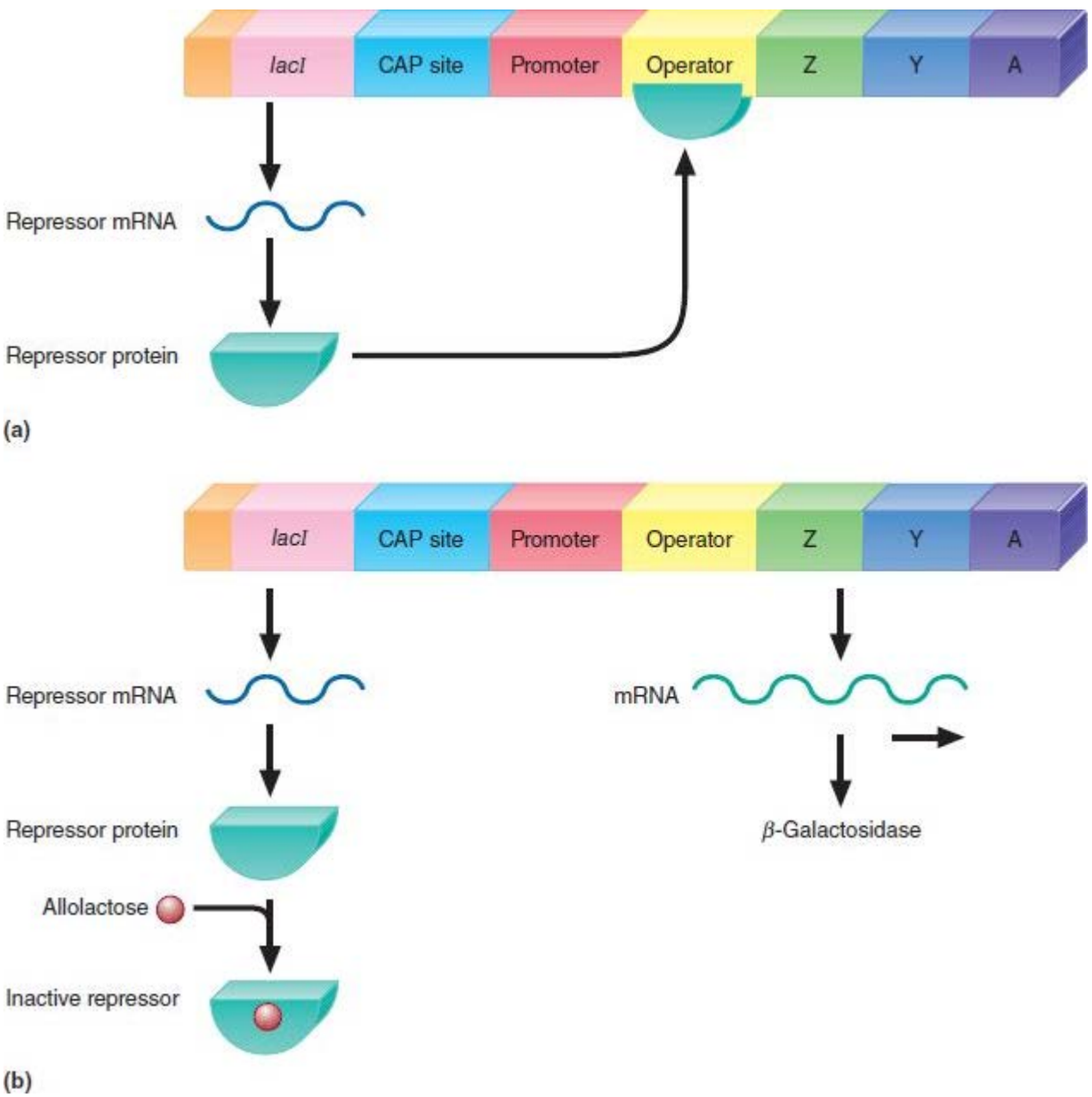


FIGURE 18.50

Function of the *lac* Operon

(a) The repressor gene *lacI* encodes a repressor protein that binds to the operator when lactose (the inducer) is not present. (b) When lactose is present, its isomer allolactose binds to the repressor protein, thereby inactivating it. (Not shown: the effect of glucose on the *lac* operon. Refer to the text for a discussion of this topic.)

The delay in activating the *lac* operon is mediated by CAP. CAP is an allosteric homodimer that binds to the chromosome at a site directly upstream of the lac promoter when glucose is absent. CAP is an indicator of glucose concentration because it binds cAMP. The cell's cAMP concentration is inversely related to glucose levels because glucose transport depresses the activity of adenylate cyclase. The binding of cAMP to CAP, which occurs only when glucose is absent and cAMP levels are high, causes a conformational change that allows the protein to bind to the *lac* promoter. CAP binding promotes transcription by increasing the affinity of RNA polymerase for the *lac* promoter. In other words, CAP exerts a positive or activating control on lactose metabolism.

RIBOSWITCHES A **riboswitch** is a metabolite-sensing domain in the 5′-untranslated region of mRNAs. Found mostly in bacteria, riboswitches monitor cellular concentrations of specific metabolites. Genes that contain riboswitches typically code for proteins that are involved in either the transport or synthesis of molecules that are expensive to produce, such as TPP or FMN. Riboswitches are toggle switch–like devices that act as feedback inhibitors to prevent the wasteful synthesis of molecules that are already present in sufficient concentrations. They are composed of two structural elements: an *aptamer*, which directly binds the metabolite, and an *expression platform*, the gene expression regulator. When the aptamer binds the metabolite, it undergoes a structural change that in turn alters the structure of the expression platform. The most common results of this process are transcription or translation inhibition.

When TPP binds to its aptamer (**Figure 18.51a**), the riboswitch is converted from a structure that has an open translation initiation site to one in which the initiation site is sequestered in a hairpin loop, effectively blocking translation. The FMN riboswitch illustrated in **Figure 18.51b** causes premature transcription termination when it reconfigures to form a terminator hairpin that prevents RNA polymerase binding. Examples of other riboswitches include those for cobalamin (coenzyme B_{12}), purines, and lysine. In rare instances, the riboswitch catalyzes a self-cleavage reaction (**Figure 18.51c**) that results in a decrease in mRNA copy number.

KEY CONCEPTS

- Constitutive genes are routinely transcribed, whereas inducible genes are transcribed only under appropriate circumstances.
- In prokaryotes, inducible genes and their regulatory sequences are grouped into operons.
- Riboswitches are metabolite-sensing domains in mRNAs that regulate the transport or synthesis of certain types of molecules.

Gene Expression in Eukaryotes

Eukaryotic genomes are also substantially more intricately regulated than those of prokaryotes, as the following example confirms. With the exception of mature red blood cells, most human cells contain the same genome. (As red blood cells develop from precursor cells called reticulocytes, they lose their nuclei.) Yet each of more than 200 highly differentiated cell types expresses a unique subset of genes that changes in response to intercellular signaling mechanisms and/or

environmental cues. Eukaryotic diversity is made possible by vastly larger groups of transcription factors and ncRNA than occurs in prokaryotes. Most eukaryotic transcription factors can be classified into the following categories, which are based on DNA-binding specificity or effects on chromatin structure: DNA sequence-specific and general or nonspecific chromatin structure factors, and chromatin remodeling factors. Eukaryotic gene expression is also regulated at more levels than the prokaryotic version. These levels include genomic control, transcriptional control (see pp. 723–27), RNA processing, RNA editing, RNA transport, and translational control. A brief description of these topics is followed by an overview of signal transduction–triggered gene expression.

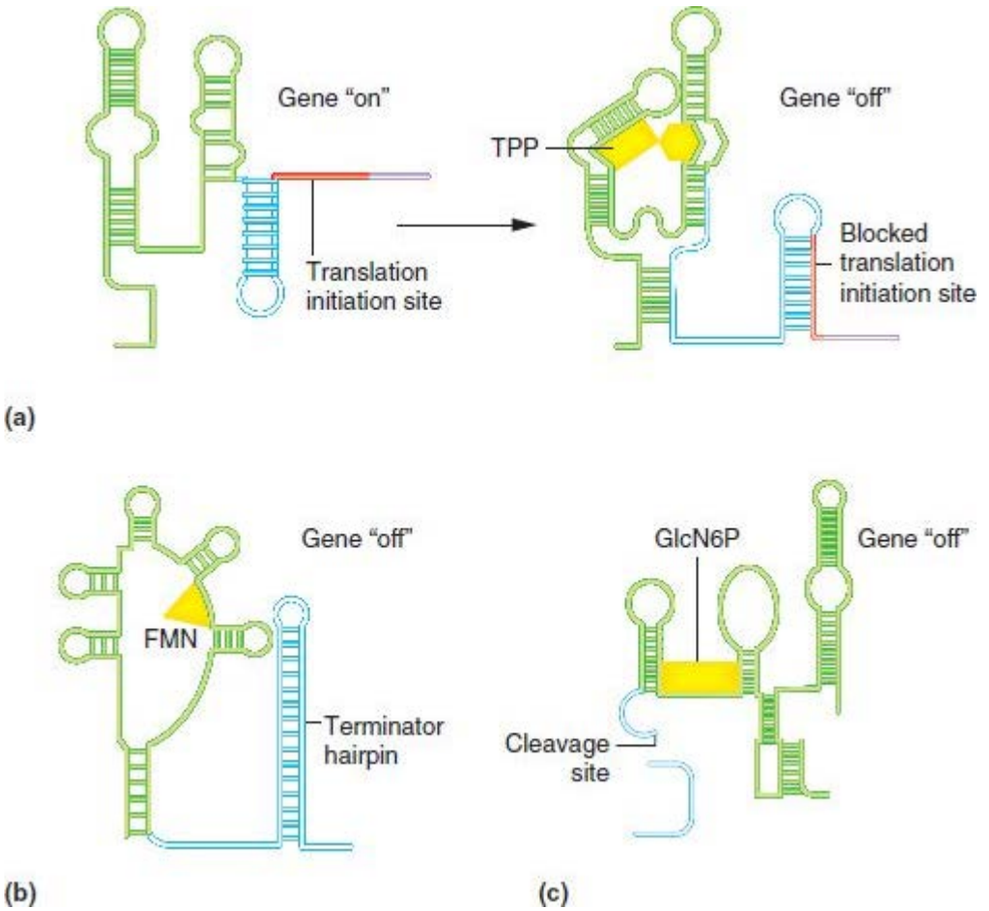


FIGURE 18.51

Riboswitches

(a) Translation prevention. When TPP binds to the aptamer, the riboswitch rearranges so that the expression platform forms a hairpin that blocks translation initiation. (b) Transcription termination. FMN binding causes the formation of a hairpin structure that halts transcription. (c) Self-cleavage. In rare instances, the binding of a metabolite (in this case, the sugar GlcN6P) triggers an mRNA self-cleavage reaction.

GENOMIC CONTROL As described previously, there appears to be two major influences on eukaryotic transcription initiation: chromatin structure and transcription factor–regulated RNA polymerase complex formation. Gene expression is affected by changes in the structural organization of the genome (i.e., chromatin remodeling induced by DNA methylation/demethylation and histone covalent modification). A significant amount of gene regulation also occurs through transcription initiation control (e.g., covalent modification of the pol II C-terminal domain).

Other examples of genomic control include gene rearrangements and gene amplification. The

differentiation of certain cells involves gene rearrangements, for example, the rearrangements of antibody genes in B-lymphocytes. Transposition (see p. 704) is also believed to affect gene regulation. During certain stages in development, the requirement for specific gene products may be so great that the genes that code for their synthesis are selectively amplified. Amplification occurs via repeated rounds of replication within the amplified region. For example, the rRNA genes in various animals (most notably amphibians, insects, and fish) are amplified within immature egg cells (called oocytes).

RNA PROCESSING Several types of eukaryotic RNA processing reactions have already been described (pp. 727–30). Among the most important of these is *alternative splicing*, the joining of different combinations of exons to form cell-specific proteins (**Figure 18.52**). By one estimate, about 37% of human protein-coding genes generate multiple protein isoforms. In addition, alternative splicing can generate mRNAs that differ in their 5′ and 3′ UTRs and in their alternative splice sites. These mRNAs can also possibly change the molecule's *reading frame* (a triplet-base sequence that codes for amino acids during translation).

Mammals have four genes (α, β, γ, and δ) that code for tropomyosin, a protein that regulates actin filaments in muscle and nonmuscle cells (e.g., fibroblasts and brain cells). The α-tropomyosin gene (**Figure 18.53**) consists of 13 to 15 exons. Five of the exons are common to all isoforms of the protein, while the remaining exons are alternately used in different α-tropomyosin mRNAs. For example, rat striated muscle α-tropomyosin mRNA contains exons 3, 11, and 12 but not exons 2 or 13, while the smooth muscle isoform contains exons 2 and 3 but not exons 11 and 12. This variation accounts in part for the differences in contractile fiber structure and function in these two muscle types. Rat brain α-tropomyosin mRNA lacks exons 2, 7, and 11 through 13 and functions in the actin-myosin cytoskeletal system in these noncontractile cells.

The selection of alternative sites for polyadenylation also affects mRNA function. In addition to altering mRNA binding sites for miRNAs, changes in polyadenylation sites also affect mRNA longevity (see p. 727). Polyadenylation site changes can also alter an mRNA's structural and functional properties. The mRNA that codes for the heavy chain of the antibody IgM is a well-researched example. There are two forms of IgM: the membrane-bound antibody and the secreted antibody. During the early phase of B-lymphocyte differentiation, the cell produces the plasma membrane IgM because heavy-chain transcript polyadenylation occurs at a site downstream from two exons that code for membrane anchor domains. Later, the cell synthesizes both the membrane-bound and the secreted versions of IgM. The heavy chain of secreted IgM is a truncated molecule that lacks the membrane anchor because the mRNA transcript is polyadenylated upstream of the membrane domain exons. Selection of the heavy-chain mRNA polyadenylation site depends on expression of the protein CstF, which determines the site that is processed.

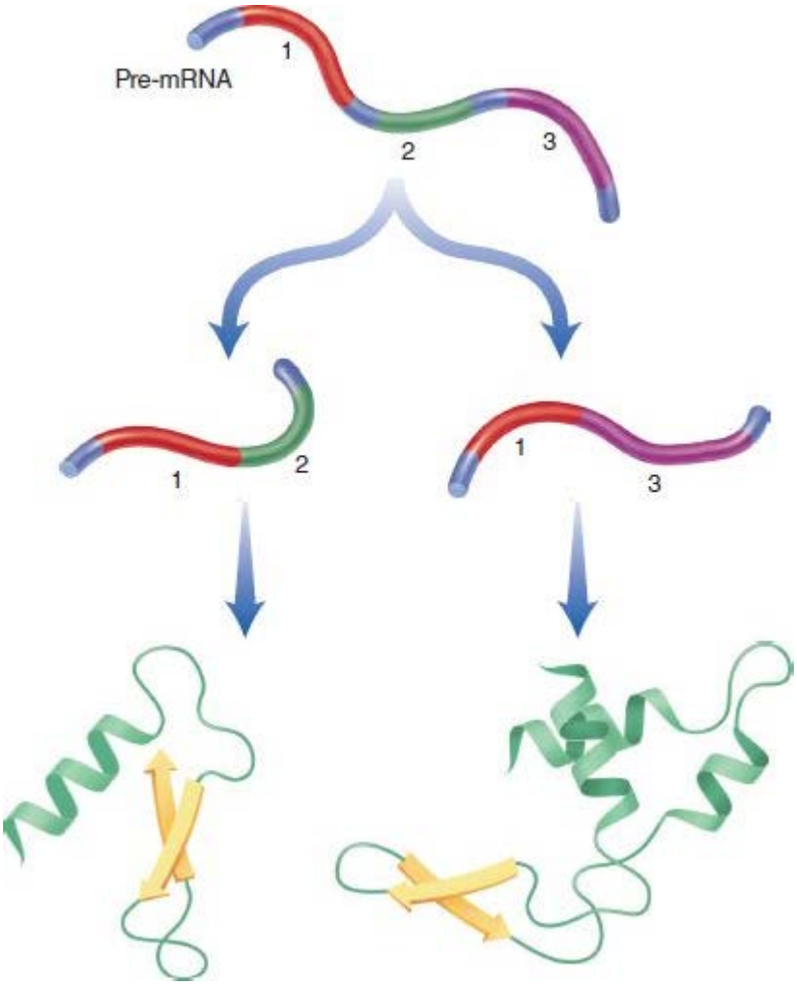


FIGURE 18.52

RNA Processing

The coding properties of an mRNA molecule depend on the types of processing event its precursor undergoes. Different polypeptides can be synthesized from splicing different combinations of exons from the same pre-mRNA transcript.

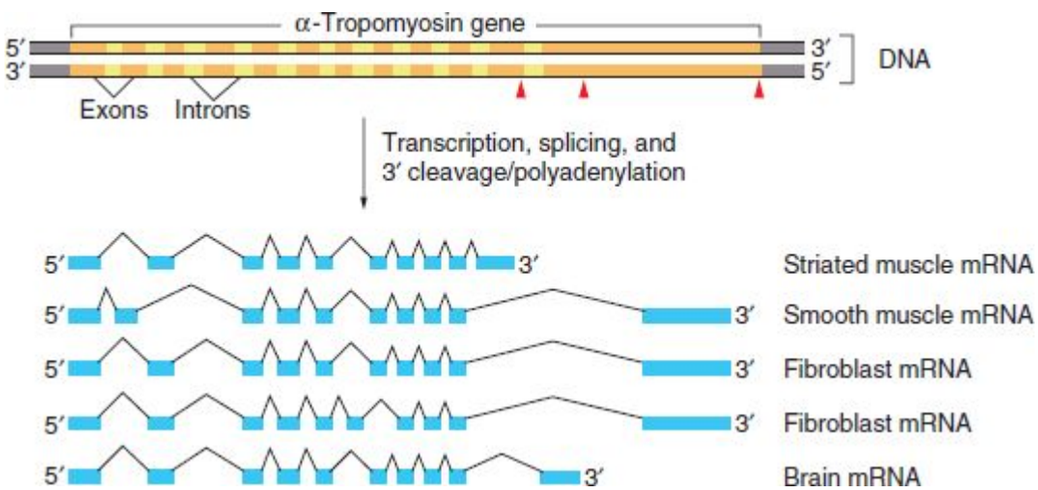


FIGURE 18.53

Alternate Splicing of the α-Tropomyosin Gene

α-Tropomyosin is a protein that regulates actin filaments in both muscle and nonmuscle cells. Alternate splicing of the primary transcript in cells results in alternate versions of *α*-tropomyosin, which are utilized for different purposes. The arrowheads indicate alternate cleavage/polyadenylation sites at the 3′ ends of mRNAs. The blue

rectangles in the mRNA represent exons. The lines between the exons represent intervening sequences that will be removed by splicing.

After transcription, base changes are effected by means of **RNA editing**. Alterations in the mRNA base sequence can have several consequences. When they occur in the 5′ and 3′ UTRs, for example, translation initiation and RNA stability, respectively, may be affected. Other possibilities include the alteration of intron splice sites and changes in the amino acid sequence of the polypeptide product. Among the best-researched examples of RNA editing are C → U and A → I conversions, where I is an abbreviation for inosine (**Figure 14.25**). The mRNA for the apolipoprotein B-100 gene codes for a polypeptide with 4563 amino acid residues, which is a component of VLDL (p. 421). Intestinal cells produce a shorter version of this molecule, namely, apolipoprotein B-48 (2153 amino acid residues), which becomes incorporated into the chylomicron particles produced by these cells (p. 421). The cytosine in a CAA codon that specifies glutamine is converted by cytidine deaminase into a uracil. The new codon, UAA, is a stop signal in translation; hence a truncated polypeptide is produced during translation of the edited mRNA.

The deamination of adenosine to inosine, catalyzed by ADAR (adenosine deaminase acting on RNA), is the most common RNA editing modification. A prominent example of an A → I transition occurs in some brain neurons, where ADAR deaminates a specific adenosine residue in the mRNA for a glutamate receptor subunit. When the edited mRNA is read by a ribosome, the I base (hypoxanthine) is read as a G. As a result, an arginine residue (codon CGA) is substituted for a glutamate (codon CAA), and the ion channel in the receptor becomes less permeable to Ca^{2+}. Transgenic mice in which the A → I transition was prevented were observed to develop a severe form of epilepsy.

POSTTRANSCRIPTIONAL GENE SILENCING **Gene silencing**, a form of posttranscriptional gene regulation in higher eukaryotes, involves short 22-nt miRNAs (p. 674). First discovered in the nematode worm *Caenorhabditis elegans*, miRNAs inhibit the translation of target mRNAs by binding to partially complementary sequences in their 3′ UTRs, although examples of 5′ UTR and coding region binding have been observed. The first miRNA to be investigated, lin-4, regulates an early phase in the worm's larval development. By preventing the translation of lin-14 and lin-28 mRNAs, lin-4 facilitates the progression of developing larvae from an early stage to a later one. MiRNAs, each of which may have hundreds of target mRNAs, have been discovered in organisms as diverse as fruit flies, frogs, rice, and maize. About one-third of human protein-coding genes are regulated by miRNAs. Defects in miRNA regulation have been associated with diseases such as cancer (e.g., leukemia) and heart disease.

Gene silencing (**Figure 18.54**) begins with the RNAP II–catalyzed synthesis of a 70-nt-long ss-RNA precursor called a primary miRNA (pri-miRNA). Each pri-miRNA, which may be thousands of nts in length, is a stem-loop structure that is processed in the nucleus into one or more miRNAs by a protein complex called microprocessor. *Microprocessor*, composed of the s- and dsRNA-binding protein DGCR8 (pasha in *D. melanogaster*) and an RNase called *drosha*, cleaves pri-miRNA 11 nt away from the base of each hairpin structure to release pre-miRNA. After GTP-driven export from the nucleus via the carrier protein exportin-5, an RNase called *dicer* cleaves each pre-miRNA to yield a double-stranded miRNA. One strand of the miRNA duplex, called the guide strand, is incorporated into a ribonucleoprotein complex called RNA-induced silencing complex (RISC). The other miRNA strand (the passenger strand) is degraded. An RISC protein called *argonaute* positions the miRNA so that it can bind the target mRNA, thereby inactivating it.

MiRNA-mediated gene silencing utilizes components of **RNA interference**, a process originally believed to be limited to protection against viruses and transposons. Cells use double-stranded

siRNAs (p. 674) to recognize and then degrade target mRNAs. siRNAs are the products of dicer-induced cleavage of larger RNA molecules (e.g., a viral RNA genome). Once the guide siRNA is incorporated into RISC (**Figure 18.54**), it binds to its complementary sequence on the target mRNA. Because the sequences match exactly, *slicer* (an enzymatic activity in a domain of an argonaute protein) cleaves the mRNA into pieces.

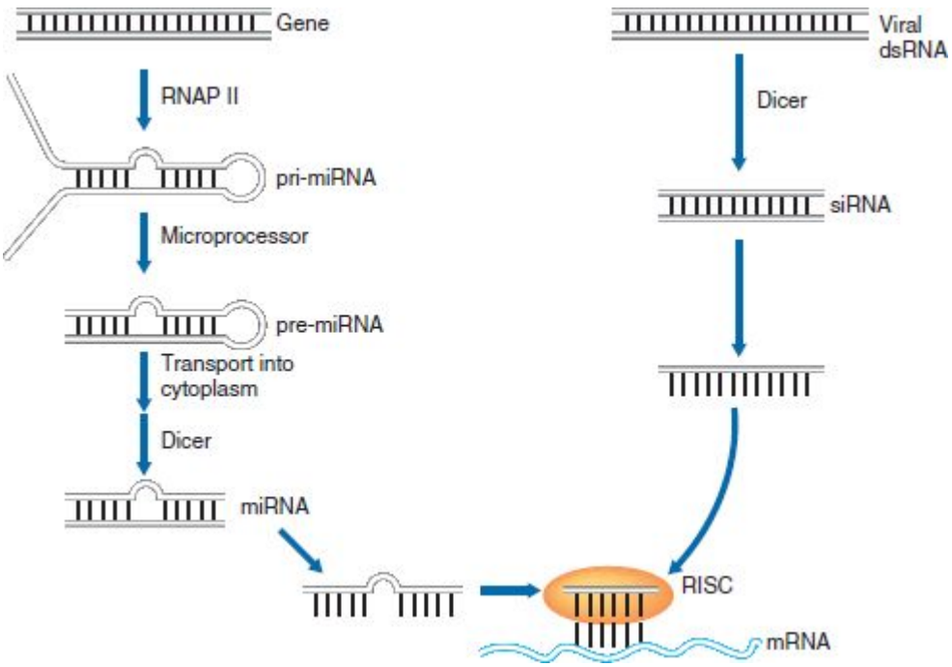


FIGURE 18.54

miRNA and siRNA Processing

In posttranscriptional gene silencing (a), the primary transcript of a miRNA gene, pri-miRNA, is processed by microprocessor, a protein complex containing DGCR8 (pasha) and drosha, and dicer to form miRNA. The miRNA guide strand is then incorporated into the RISC complex, where it binds to a complementary sequence in the 3′ UTR of its target mRNA, referred to as a miRNA response element (MRE). Because these two sequences are not perfectly complementary, the mRNA is silenced but not degraded. In RNA interference (b), dicer cleaves a foreign dsRNA to yield the dsRNA molecule siRNA. Once the guide strand of the siRNA has been positioned within the RISC, it binds to its complementary sequence on the viral mRNA. Because these two sequences are perfectly complementary, the slicer activity of the RISC proceeds to cleave the mRNA into pieces.

RNA TRANSPORT mRNA transport out of the nucleus, a highly regulated process, occurs in three phases: processing reactions, docking and passage through the NPC (p. 56), and release into the cytoplasm. In the first phase, pre-mRNA molecules are simultaneously processed into mRNAs and packaged into ribonucleoprotein complexes (mRNPs). mRNP proteins (e.g., cap-binding protein, EJCs, and poly(A)-binding protein) recruit export factors that allow NPC targeting. The capping and splicing proteins allow binding to TREX, an export protein complex. Once mRNPs are linked via a TREX subunit to Nxf1-Nxt1, a heterodimer nuclear export receptor, they move through the NPC. When an mRNP complex reaches the cytoplasm, the release of export proteins triggers the remodeling of the complex that in turn directs transport to its final destination where translation will occur.

TRANSLATIONAL CONTROL Eukaryotic cells can respond to various stimuli (e.g., heat shock, viral infections, and cell cycle phase changes) by selectively altering protein synthesis. The

covalent modification of several translation factors (nonribosomal proteins that assist in the translation process) has been observed to alter the overall protein synthesis rate and/or enhance the translation of specific mRNAs. For example, when cellular iron levels are low, a repressor protein binds to mRNAs coding for the iron storage protein ferritin. When iron levels rise sufficiently, the binding of iron to the repressor protein triggers a conformational change that causes it to dissociate from mRNA. The ferritin mRNA can then be translated.

SIGNAL TRANSDUCTION AND GENE EXPRESSION All cells respond to signals from their environment in part by altering gene expression patterns. Signal transduction–triggered changes in gene expression are initiated by the binding of a ligand to either a cell-surface receptor or an intracellular receptor. The mechanisms by which signal molecules switch certain genes on or off are an intricate series of reactions and protein conformation changes that transmit information from the cell's environment to specific DNA sequences in the nucleus. The enormous research efforts devoted to investigation of cancer (see Biochemistry in Perspective: Carcinogenesis, p. 743) have shown that the best understood examples of such signal transduction pathways to be those that affect cell division.

In contrast to single-cell organisms in which cell growth and cell division are governed largely by nutrient availability, an elaborate intercellular network of growth-promoting and growth-inhibiting signal molecules regulates the proliferation of cells in multicellular organisms. Several complicating features of intracellular signal transduction mechanisms have been revealed by research efforts in cell proliferation. For example, the mechanisms by which signal molecules alter gene expression often involve the simultaneous activation of several different pathways. Depending on circumstances, the activation of several types of receptor may result in overlapping responses.

In the eukaryotic cell cycle, cells repeatedly progress through each of the four phases (M, G_1, S, and G_2: refer to **Figure 18.13**). Checkpoints occur in G_1 (in yeast cells, it is referred to as START), G_2, and M phases. The cell is prevented from entering the next phase until the conditions are optimal in its current phase (e.g., sufficient cell growth in G_2 or alignment of chromosomes in M). The fixed, rhythmic activities observed in cell division are regulated so that each phase is completed before the next one starts. Progression is accomplished by the alternating synthesis and degradation of a group of regulatory proteins called the **cyclins**. Each type of cyclin binds to and activates specific cyclin-dependent protein kinases (Cdks) (p. 741). Activated Cdks proceed to phosphorylate a variety of proteins that control the passage of cells though specific checkpoints in the cell cycle.

Regulation of cell division involves both positive and negative controls. Positive control is exerted largely by a variety of growth factors that bind to specialized cell receptors. The protein products of *tumor suppressor genes* inhibit cell division. Well-known examples of tumor suppressor genes include *Rb* (so named because of the role played by the loss of *Rb* gene function in the childhood eye cancer *retinoblastoma*) and the *p53* gene (coding for a protein that among its numerous roles acts as cyclin chaperone that arrests cell cycle progression). The arrest of the cell cycle is prolonged when a certain amount of DNA damage has occurred, as in overexposure to radiation. If DNA repair mechanisms are incomplete, a complex mechanism involving p53 leads to programmed cell death, or **apoptosis** (**Figure 2.22**).

The positive effects exerted by growth factors are now believed to include gene expression that specifically overcomes the inhibitions at the cell cycle checkpoints, especially the G_1 checkpoint. The binding of growth factors to their cell-surface receptors initiates a cascade of reactions that induces two classes of genes: early response genes and delayed response genes.

Early Response Genes. These genes, which usually code for transcription factors, are rapidly activated, usually within 15 minutes. Among the best-characterized early response genes are the *jun, fos*, and *myc* protooncogenes. **Protooncogenes** are normal genes that, if mutated, can promote carcinogenesis. (See Biochemistry in Perspective: Carcinogenesis, p. 743.) Each of the *jun* and *fos* protooncogene families code for a series of transcription factors containing leucine zipper domains. Both Jun and Fos proteins form dimers that can bind DNA. Among the best characterized of these is a Jun-Fos heterodimer, referred to as AP-1, which forms through a leucine zipper interaction. Myc is one member of a large class of transcription factors that possess the basic helix-loop-helix-leucine zipper (bHLHZip) DNA-protein binding motif. Members of the Myc protein family form homo- and heterodimers with themselves and with members of certain other transcription factor families. When Myc forms a heterodimer with a protein called Max, the expression of a large number of genes is affected. The products of some of Myc/Max target genes have a stimulatory effect on the cell cycle.

Delayed Response Genes. Delayed response genes are induced by the activities of the transcription factors and other proteins produced or activated during the early response phase. Among the products of the delayed response genes are the Cdks, the cyclins, and other components required for cell division.

As mentioned earlier (see p. 611), many growth factors bind to tyrosine kinase receptors, and some of these are linked via G-protein-like mechanisms to DAG and IP_3 generation (**Figure 16.8**). Epidermal growth factor (EGF; p. 603) is one growth factor of this type, and **Figure 18.55** gives a brief overview of its role in the activation of the transcription factor AP-1. When EGF binds to its cognate receptor, the receptor undergoes dimerization and autophosphorylation. The now phosphorylated cytoplasmic domain then binds numerous proteins. Among these proteins is the SOS/GRB2 dimer. The activation of SOS, a type of guanine nucleotide exchange factor (GEF; p. 606), requires binding first to Grb2 and then to the activated receptor. Subsequently SOS/GRB2 binds to and activates the monomeric G protein Ras, the product of a protooncogene (p. 741). Ras activation ends when GTP is hydrolyzed in a reaction catalyzed by a GTPase activating protein (GAP; p. 607).

PLCγ is also activated when it binds to the EGF receptor. Like its counterpart in G-protein–linked receptors, active PLCγ hydrolyzes PIP_2 to form DAG and IP_3. DAG activates PKC, the enzyme that in turn activates a variety of proteins involved in cell growth and proliferation.

Phosphorylation cascades are induced by both the activation of Ras and increased levels of DAG and IP_3 in the cell following EGF binding. One of the key enzymes activated in the phosphorylation cascade is MAPKK (mitogen-activated protein kinase kinase). Activated MAPKK then phosphorylates both a tyrosine and a threonine of MAP kinase (MAPK). (This unusual reaction appears to ensure that MAP kinase is activated only by MAPKK.) Active MAPK (a serine/threonine kinase) then phosphorylates a variety of cellular proteins. Among those proteins are Jun, Fos, and Myc. Phosphorylated Jun and Fos proteins then combine to form AP-1, which promotes the transcription of several delayed response genes that promote cell division.

QUESTION 18.5

The mechanism by which light influences plant gene expression is referred to as *photomorphogenesis*. Because of serious technical problems with plant cell culture, relatively little is known about plant gene expression. However, certain DNA sequences, referred to as *light-responsive elements* (LREs), have been identified. On the basis of the gene expression

patterns observed in animals, can you suggest (in general terms) a mechanism whereby light induces gene expression? [*Hint*: Review the properties of phytochrome.]

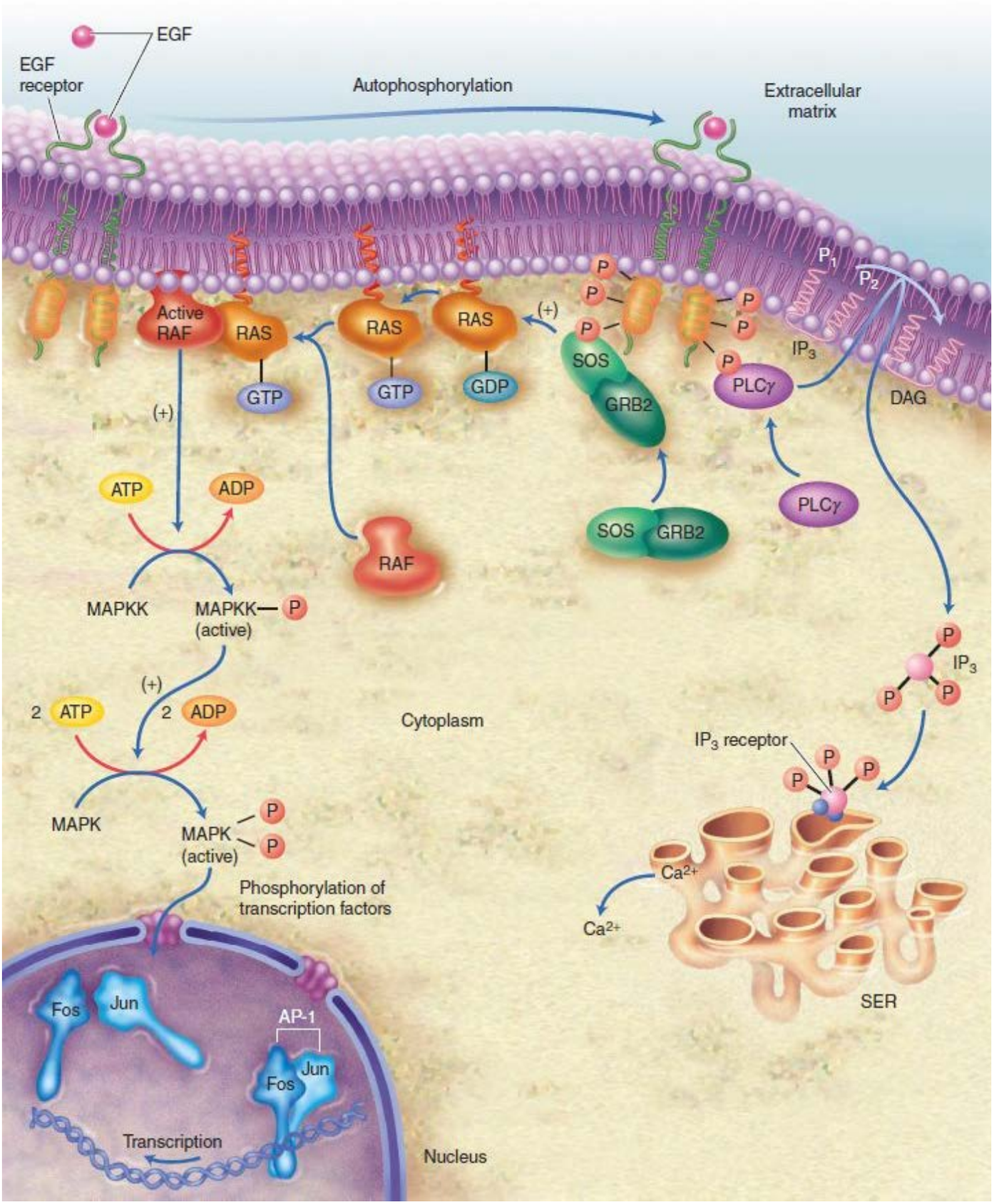


FIGURE 18.55

A Simplified Example of Eukaryotic Gene Expression Triggered by Growth Factor Binding

Signal transduction pathway illustrating select events that are triggered when a growth factor such as EGF binds to its plasma membrane receptor. Subsequent changes in growth factor– or hormone-triggered gene expression that occur are typically mediated by several different mechanisms. This example illustrates only Ras and PLCγ activation. When EGF binds, it promotes dimerization of the receptor and autophosphorylation of the tyrosine residues on its cytoplasmic domain. As a result, the receptor binds a variety of cytoplasmic proteins. One of these proteins, GRB2, bound to SOS (a GEF), is activated when it associates with the EGF receptor. GRB2/SOS then activates Ras by promoting the exchange of GDP for GTP. Activated Ras initiates a phosphorylation cascade by activating the protein kinase RAF (a MAPKKK), which in turn activates MAPKK. MAPKK activates MAPK, which then migrates into the nucleus where it activates a number of transcription factors (e.g., Fos and Jun that form AP-1). After PLCγ binds to the EGF receptor, it catalyzes the cleavage of PIP_2 into IP_3 and DAG. IP_3 stimulates the release of Ca^{2+} into the cytoplasm. As illustrated in **Figure 16.7**, in the presence of Ca^{2+} and DAG, protein kinase C activates yet another series of protein kinases that in turn affect the function of several cell growth regulatory proteins.

Biochemistry IN PERSPECTIVE

Carcinogenesis

What is cancer, and what are the biochemical processes that facilitate the transformation of normal cells into those with cancerous properties? Cancer is a group of diseases in which genetically damaged cells proliferate autonomously. Such cells cannot respond to normal regulatory mechanisms that ensure the intercellular cooperation required in healthy multicellular organisms. Consequently, they continue to proliferate, thereby robbing nearby normal cells of nutrients and eventually crowding surrounding healthy tissue. Depending on the damage they have sustained, abnormal cells may form either benign or malignant tumors. Benign tumors, which grow slowly and are limited to a specific location, are not considered cancerous and rarely cause death. In contrast, *malignant* tumors are often fatal because they can undergo *metastasis*, migration through blood or lymph vessels to distant locations throughout the body. Wherever new malignant tumors arise, they interfere with normal functions. When life-sustaining processes fail, patients die.

Cancers are classified by the tissues affected. The vast majority of cancerous tumors are *carcinomas* (tumors derived from epithelial tissue cells such as skin, various glands, breasts, and the lining of most internal organs). *Leukemias* (lymphoblastic or myelogenic) are cancers of various blood-forming cells. *Lymphomas* are solid tumors of lymphoid tissue (lymph nodes and spleen). Tumors arising in connective tissue are called *sarcomas*. Despite the differences among these diseases, they also have several common characteristics, including genomic instability, growth factor-independence, growth suppression signal insensitivity, and apoptosis evasion. Each tumor originates from a single damaged cell. In other words, a tumor is a clone derived from a cell in which heritable changes have occurred. The genetic damage consists of mutations (e.g., point mutations, deletions, and inversions) and chromosomal rearrangements or losses. Such changes result in the loss or altered function of molecules involved in cell growth or proliferation. Tumors typically develop over a long time and involve several independent types of genetic damage. As a

result, the risk of many types of cancer increases with age.

Transformation, the process in which an apparently normal cell is converted or "transformed" into a malignant cell, is a complex process that may involve numerous genes in diverse cellular pathways. Although a cancerous cell may contain thousands of mutations, typically only a small set (e.g., 9 in breast cancer and 12 in colorectal cancer) is required for transformation. There are three phases in transformation: initiation, promotion, and progression.

During the *initiation* phase of carcinogenesis, a destabilizing change in a cell's genome provides it with a growth advantage over its neighbors. Most initiating mutations affect protooncogenes or tumor suppressor genes. **Protooncogenes** code for a variety of growth factors, growth factor receptors, enzymes, or transcription factors that promote cell growth and/or cell division. It is the mutated versions of proto-oncogenes, called **oncogenes**, that promote abnormal cell proliferation (**Table 18A**). **Tumor suppressor genes** code for proteins that actively protect cells from progressing toward cancer. For example, in cells with damaged DNA the tumor suppressor RB1 codes for a protein (Rb) that inhibits a transcription factor complex required for cell cycle progression. Other tumor suppressors are DNA repair genes such as BRCA1 and BRCA 2 (p. 709), which are involved in the repair of double-stranded DNA breaks. The tumor suppressor protein p53, coded for by the TP53 gene on chromosome 17, has multiple functions that include DNA repair protein activation, cell cycle regulation, and apoptosis initiation. In response to cellular stresses such as DNA damage and oxidative stress, p53 is activated by phosphorylation of its N-terminal domain by specific kinases. The subsequent conformational change converts p53 into a transcription regulatory protein and allows it to bind transcription activator proteins.

In hereditary cancers, the initiation event is often caused by genomic instability attributed to a germ line mutation in a DNA repair gene. For example, about 5% of breast cancer cases in the United States are caused by inherited heterozygous mutations in either BRCA1 or BRCA2 (p. 709). Many hereditary cancers are genetically recessive (i.e., both copies of the gene must be inoperable for the cancer to occur). Because mutations occur throughout life, as an individual grows older, the risk of a mutation in the second gene copy increases. By some estimates, the cumulative cancer risk of a carrier of one copy of mutated BRCA1 is 80% by age 70. The cause of genomic instability in sporadic cancers is less clear. However, it is believed that several factors contribute to destabilizing DNA damage. Examples include oxidative stress, proteotoxic stress (an inadequate unfolded protein response), and chronic inflammation. Initiating events are caused by damage that alters the expression or function of protooncogenes and/or tumor suppressors. Examples of this damage include the following:

TABLE 18A Selected Oncogenes*

Oncogene	Protooncogene Function
Sis	Platelet-derived growth factor
erbB	Epidermal growth factor receptor
Src	Tyrosine-specific protein kinase
Raf	Serine/threonine-specific protein kinase
Ras	GTP-binding protein
Jun	Transcription factor

Fos	Transcription factor
Myc	Transcription factor

*Abnormal versions of protooncogenes that mediate cancerous transformations.

1. **Carcinogenic chemicals**. Most cancer-causing chemicals are mutagenic; that is, they alter DNA structure. Some carcinogens (e.g., nitrogen mustard) are highly reactive electrophiles that attack electron-rich groups in DNA (as well as RNA and protein). Other carcinogens (e.g., benzo[*a*]pyrene) are actually procarcinogens, which are converted to active carcinogens by one or more of the body's enzyme-catalyzed reactions.
2. **Radiation**. Some radiation (UV, X-rays, and γ-rays) is carcinogenic. As noted, the damage inflicted on DNA includes single- and double-strand breaks, pyrimidine dimer formation, and the loss of both purine and pyrimidine bases. Radiation exposure causes the formation of ROS, which may be responsible for most of radiation's carcinogenic effects.
3. **Viruses**. Viruses appear to contribute to the transformation process in several ways. Some introduce oncogenes into a host cell chromosome as they insert their genome. Viruses can also affect the expression of cellular protooncogenes through insertional mutagenesis, a random process in which viral genome insertion inactivates a regulatory site or alters the protooncogene's coding sequence. Only a few human cancers have been proven to be associated with viral infection. Examples include human papilloma virus and hepatitis B virus, which induce cervical cancer and liver cancer, respectively.

Chemicals that do not alter DNA structure can also promote tumor development. So-called **tumor promoters** contribute to carcinogenesis by two principal methods. By activating components of intracellular signaling pathways, some molecules (e.g., the phorbol esters) provide the cell a growth advantage over its neighbors. (Recall that phorbol esters [see p. 614] activate PKC because they mimic the actions of DAG.) The effects of many other tumor promoters are unknown but may involve transient effects such as increasing cellular Ca^{2+} levels or increasing synthesis of the enzymes that convert procarcinogens into carcinogens. Unlike initiating agents, the effects of tumor promoters are reversible. They produce permanent damage only with prolonged exposure after an affected cell has undergone an initiating mutation.

Following initiation and promotion, cells go through a process referred to as progression. During *progression*, genetically vulnerable precancerous cells, which already possess significant growth advantages over normal cells, are further damaged. Eventually, the continued exposure to carcinogens and promoters makes further random mutations inevitable. If these mutations affect cellular proliferative or differentiating capacity, then an affected cell may become sufficiently malignant to produce a tumor.

Cancer is not only a genetic disease; it is also currently appreciated as an epigenetic disease. Chromatin-associated gene silencing has been implicated in the progress of a number of cancers. Epigenetic analysis of DNA methylation utilizes DNA sequencing methods on bisulfite-treated genomic DNA. (Bisulfite [HSO_3^-] converts unmethylated cytosines to uracil, but does not affect methylated cytosines.) This work has revealed that DNA in cancer cells is, in general, hypomethylated in comparison to normal cells. Hypomethylation of repetitive sequences, which becomes more pronounced as normal cells are transformed into cancerous cells, results in genomic

instability, transposable element activation, and the expression of normally repressed genes. Lung cancer caused by smoking provides an example of cancerous epigenetic change. Cigarette smoke has been definitively proven to cause demethylation of the γ-synuclein gene, which codes for an unfolded, soluble protein. Normally expressed only in neurons, the γ-synuclein gene is an oncogene in lung tumors where it facilitates metastasis. Epimutations also include hypermethylation of CpG islands near normally expressed gene sequences that render them inactive. For example, hypermethylation silences the *WRN* gene that codes for a protein with DNA helicase and exonuclease activity. In the absence of the WRN protein, affected cells have unstable chromosomes, with large numbers of deletions, inversions, and translocations.

LncRNAs and Cancer

Extensive genome-wide cancer surveys have revealed that about 80% of cancer-linked genetic variants occur in noncoding DNA sequences, many of which are transcribed into lncRNAs (p. 675). In addition, a significant number of lncRNAs have been functionally associated with various forms of cancer. In cancerous cells, lncRNA dysfunction can contribute to uncontrolled cell division, apoptosis and tumor suppressor resistance, angiogenesis induction (improved blood supply to tumors), and metastasis. HOTAIR is a well-researched example.

HOTAIR (HOX transcript antisense RNA) is transcribed from the *HOXC* gene on chromosome 12. (*HOX* genes control body plan segments such as vertebrae during embryogenesis.) Under normal conditions, the primary identified function of HOTAIR is regulation of the *HOX* gene cluster on chromosome 2. Acting as a chromatin scaffold, HOTAIR silences these genes by recruiting PRC2, a chromatin remodeling methyltransferase complex that initiates silencing by trimethylating H3K27, and LSD1, a demethylase that inactivates the active histone mark H3K4me2. Excessive synthesis of HOTAIR in several types of cancer (e.g., breast, gastric, colorectal, liver, and ovarian cancers, and glioblastoma) results in the genome-wide targeting of PRC2 to hundreds of genes, including those with tumor suppressor functions. Under these conditions, HOTAIR also facilitates LSD1's activation of c-Myc, one of three transcription factor genes in the myc family, as well as its coactivator HBXIP. Together a complex formed by HOTAIR, HBXIP, and LSD1 contributes to cMyc transcriptional activation. Since c-Myc controls hundreds of genes involved in cell proliferation and cellular metabolism, any increase in its transcription increases the risk of carcinogenesis. HOTAIR also promotes tumor progression by acting as a decoy for miRNAs. In metastatic breast cancer, for example, the overexpression of HOTAIR results in the sequestration of a miRNA that suppresses the translation of an epidermal growth factor receptor coding mRNA.

SUMMARY: Carcinogenesis is the process whereby cells with a growth advantage over their neighbors are transformed by mutations in the genes that control cell division into cells that no longer respond to regulatory signals.

Chapter Summary

1. Living organisms must possess efficient mechanisms for rapid and accurate DNA synthesis. DNA replication occurs by a semiconservative mechanism; that is, each of the two parental strands serves as a template to synthesize a new strand. Enzymatic activities required in DNA replication are DNA

unwinding, primer synthesis, polynucleotide synthesis, supercoiling control, and ligation. Although the basic features of DNA replication are similar in prokaryotes and eukaryotes, there are significant differences (e.g., replication time and rate, replication origin numbers, Okazaki fragment size, and replication machinery structure).

2. There are several types of DNA repair mechanism. These include photoreactivation repair, base excision and nucleotide excision repair, double-strand break repair, and recombinational repair.
3. Genetic recombination, a process in which DNA sequences are exchanged between different DNA molecules, occurs in two forms. In general recombination, the exchange occurs between sequences in homologous chromosomes. In site-specific recombination, the exchange of sequences requires only short homologous sequences. DNA-protein interactions are principally responsible for the exchange of largely nonhomologous sequences.
4. Transposition, the movement of genetic elements (transposons) from one location to another within a genome, can cause genetic changes such as insertions, deletions, and translocations. The movement of retrotransposons, found in large numbers in eukaryotic genomes, can cause disease or provide opportunities for genetic diversity.
5. The synthesis of RNA, referred to as DNA transcription, requires a variety of proteins. Transcription initiation involves binding an RNA polymerase to a specific DNA sequence called a promoter. Regulation of transcription differs significantly between prokaryotes and eukaryotes. Examples of eukaryotic transcription processes not observed in prokaryotes include RNA-processing features such as capping, poly(A) tail synthesis, and RNA splicing.
6. Prokaryotic gene expression regulation involves regulatory units called operons and RNA-based structures called riboswitches. Eukaryotes have a wide variety of mechanisms that control gene expression. Prominent examples include DNA methylation, histone covalent modification and chromatin remodeling, RNA-processing reactions such as alternative splicing, and RNA editing, RNA transport, posttranscriptional gene silencing, and translational controls.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on genetic information to help you prepare for exams.

Chapter 18 Review Quiz

Suggested Readings

Cao J, et al. 2015. Three-dimensional regulation of transcription. Protein Cell 6(4):241–53.

Ehrenberg R. 2016. Cigarettes cause telltale DNA damage. Sci News (Nov 26).

Eidem TM, et al. 2016. Noncoding RNAs: regulators of the mammalian transcription machinery. J Mol Biol 428:2652–9.

Fabre E, Zimmer C. 2018. From dynamic chromatin architecture to DNA damage repair and back. Nucleus 9:161–70.

Fang Y, Fullwood MJ. 2016. Roles, functions and mechanisms of long non-coding RNAs in cancer. Genomics Proteomics Bioinformatics 14:42–54.

Galganski L, et al. 2017. Nuclear speckles: molecular organization, biological function and role in disease 45(18), DOI:10:1093/nar/gkx759.

Goyal N, Kesharwani D, Datta M. 2018. Lnc-ing non-coding RNAs with metabolism and diabetes: roles of lncRNAs. Cell Mol Life Sciences 75:1827–37.

Ledford H. 2017. CRISPR’s mysteries. Nature 541:280–2.

Long Y, et al. 2017. How do lncRNAs regulate transcription? Sci Adv, DOI10.1126/sciadv.aao2110.
Mestel R. 2017. The original CRISPR. Sci News (Apr 15).
Michel B, Sandler SJ. 2017. Replication restart in bacteria. J Bacteriol 199(13):102–17.
Thurtle-Schmidt DM. 2018. Molecular biology at the cutting edge: A review on CRISPR/Cas9 gene editing for undergraduates. Biochem Mol Biol Education 46(2):195–205.
Zhao M, et al. 2018. MALAT1: a long non-coding RNA highly associated with human cancers. Oncology Letters 16:19–26.

Key Words

alternative splicing, *736*
annotation, *717*
apoptosis, *740*
apurinic site, *700*
apyrimidinic site, *700*
attachment site, *709*
base excision repair, *700*
bioinformatics, *718*
cell division cycle (Cdc) proteins, *697*
cDNA library, *715*
chromatin-remodeling complexes, *723*
chromosomal jumping, *715*
chromosomal walking, *715*
clamp loader, *690*
coding strand, *718*
colony hybridization technique, *714*
composite transposon, *709*
conjugation, *708*
consensus sequence, *720*
contig, *715*
core promoter element, *724*
cosmid, *713*
CRISPR, *731*
cyclin-dependent protein kinase, *697*
DNA glycosylase, *700*
DNA ligase, *692*
DNA microarray, *716*
DNA topoisomerase, *692*
double-strand break repair model, *706*
enhancer, *724*
exon junction complex, *729*
exonuclease, *690*

functional genomics, *712*
general recombination, *704*
general transcription factor, *725*
gene silencing, *738*
genomics, *712*
global genomic repair, *700*
GTPase activating protein, *741*
guanine nucleotide exchange factor, *741*
helicase, *690*
insertional element, *709*
insulator, *725*
intrinsic termination, *721*
inverted repeat, *709*
lagging strand, *693*
leading strand, *693*
MCM complex, *697*
Mediator, *725*
Meselson–Radding model, *706*
mismatch repair, *702*
nonsense-mediated decay, *729*
nucleotide excision repair, *700*
Okazaki fragment, *693*
oncogene, *743*
operator, *732*
operon, *732*
origin of replication complex, *697*
palindrome, *709*
photoreactivation repair, *699*
polymerase chain reaction, *712*
preinitiation replication complex, *697*
primase, *690*
primer, *690*
primosome, *690*
processivity, *690*
promoter, *720*
protooncogene, *741*
recombinant DNA technology, *712*
recombination, *689*
replication, *689*
replication factor C, *699*
replication factory, *689*
replication fork, *693*

replication licensing factor, *697*
replication protein A, *698*
replication slippage, *702*
replicon, *693*
replisome, *690*
rho-dependent termination, *721*
rho-independent termination, *721*
riboswitch, *735*
RNA editing, *737*
RNA interference, *738*
RNA splicing, *727*
semiconservative replication, *689*
shotgun cloning, *715*
silencer, *724*
single-stranded DNA-binding protein, *694*
site-specific recombination, *704*
SOS response, *703*
spliceosome, *728*
supraspliceosome, *730*
synthesis-dependent strand annealing model, *707*
telomerase, *699*
telomere end-binding protein, *699*
telomere repeat-binding factors, *699*
transcription coupled repair, *701*
transcription factory, *723*
transcription preinitiation complex, *725*
transcription start site, *724*
transduction, *708*
transfection, *713*
transformation, *708*
transposable element, *704*
transposition, *704*
transposon, *704*
tumor promoter, *744*
tumor suppressor gene, *743*
vector, *712*
yeast artificial chromosomes, 713

Review Questions

SECTION18.1

Comprehension Questions

1. Define the following terms:
    a. replication
    b. semiconservative
    c. replication factory
    d. primosome
    e. clamp loader
2. Define the following terms:
    a. processivity
    b. replisome
    c. exonuclease
    d. DNA ligase
    e. replicon
3. Define the following terms:
    a. Okazaki fragment
    b. ter region
    c. tus protein
    d. transcription preinitiation complex
    e. ORC
4. Define the following terms:
    a. licensing factors
    b. RPA
    c. TEBP
    d. TRF
    e. RFC
5. Define the following terms:
    a. transposition
    b. DNA glycosylase
    c. apurinic site
    d. apyrimidinic site
    e. mismatch repair
6. Define the following terms:
    a. nonreplicative transposition
    b. transposable element
    c. bacterial transformation
    d. transduction
    e. conjugation
7. Define the following terms:
    a. transfection
    b. replicative transposition
    c. composite transposition
    d. retrotransposon

e. insertional element

8. Define the following terms:
   a. PCR
   b. cosmid
   c. electroporation
   d. transgenic animal
   e. colony hybridization technique
9. Define the following terms:
   a. inverted repeat
   b. DNA microarray
   c. chromosomal jumping
   d. genome project
   e. bioinformatics
10. Define the following terms:
   a. tus-ter complex
   b. chimera
   c. MHC antigens
   d. helices
   e. primer

Fill in the Blanks

11. The process whereby DNA molecules are cut and rejoined in new combinations is called _____ .
12. A _____ is a multienzyme complex that synthesizes RNA primers in *E. coli* DNA replication.
13. An enzyme called _____ catalyzes the formation of a covalent phosphodiester bond between the 3′-OH end of one segment of a DNA strand and the 5′-phosphate end of an adjacent segment during DNA replication.
14. In prokaryotes, Okazaki fragments are approximately _____ nucleotides long.
15. _____ is a mechanism that removes and replaces individual nucleotides in DNA when bases have been damaged.
16. The two forms of recombination are general and _____.
17. General recombination occurs between _____ DNA molecules.
18. DNA sequences called _____ can move from one site in a genome to another.
19. DNA _____ are enzymes that prevent tangling of DNA strands during DNA replication.
20. The average of a number of closely related but nonidentical sequences is referred to as a(n) _____.

Short-Answer Questions

21. List and describe the steps in prokaryotic DNA replication. How does this process appear to differ from eukaryotic DNA replication?
22. List the stage of DNA replication when each of the following enzymes is active.
   a. helicase
   b. primase

c. DNA polymerases
d. ligase
e. topoisomerase
f. DNA gyrase

23. DNA is polymerized in the 5′ → 3′ direction. Demonstrate with the incorporation of three nucleotides into a single strand of DNA how the 5′ → 3′ directionality is derived.
24. Explain the significance of “jumping genes.”
25. Describe two forms of genetic recombination. What functions do they fulfill?
26. Although genetic variation is required for species to adapt to changes in their environment, most genetic changes are detrimental. Explain why genetic variations are rarely beneficial.
27. General recombination occurs in bacteria where it is involved in several types of intermicrobial DNA transfer. What are these types of transfer, and by what mechanisms do they occur?
28. Compare and contrast the mechanisms of replicative and nonreplicative transposition.
29. Within cells, cytosine slowly converts to uracil. What type of mutation would this cause in DNA molecules? What impact would the same modification have if it occurred in RNA molecules at the RNA or polypeptide product level?
30. A correlation has been found among species between life span and efficiency of DNA repair systems. Suggest a reason for this phenomenon.
31. What are the similarities and differences between cellular DNA replication and PCR?
32. Describe the purpose of marker genes in recombinant DNA technology.
33. List the steps in the processing of a typical eukaryotic mRNA precursor to prepare it for its functional role.
34. Estimate the magnitude of amplification of a single DNA molecule that can be obtained using PCR during five cycles.
35. During base excision repair, DNA glycosylase cleaves the N-glycosidic link between the altered base and the deoxyribose component of the nucleotide. Draw a typical nucleotide and indicate which bond is cleaved.
36. Cells exposed to ultraviolet light develop thymine dimers. Some organisms use visible light to reverse this damage. Explain how this form of DNA repair occurs.
37. RNA molecules are more reactive than DNA molecules. Explain.
38. Describe the function of telomere end-binding proteins.
39. Describe the function of telomere repeat-binding factors.
40. Describe the role of mismatch repair (MMR) in DNA repair. Why can mutations in MMR increase the risk of human cancer?
41. What role does a stalled RNA polymerase play in transcription-coupled nucleotide excision repair.
42. Telomerase activity has been found to be 10 to 20 times more active in cancer cells than in normal somatic cells. What is the significance of this circumstance?

Critical-Thinking Questions

43. Infection caused by a rare virulent strain of group A streptococcus has appeared relatively recently. In approximately 25 to 50% of these cases (reported in Great Britain and the United States), infection resulted in necrotizing fasciitis, a rapidly spreading destruction of flesh often accompanied by hypotension (low blood pressure), organ failure, and toxic shock. If antibiotic treatment is not initiated within three days of exposure to the bacterium, gangrene and death may result. Similar cases were reported in the 1920s. However, these earlier cases had a significantly lower fatality rate, even though antibiotics were not then available. (Physicians reported washing the affected areas with acidic solutions.) Group A streptococci are converted

into the pathogenic form by becoming infected with a certain virus. This virus's genome contains a gene that codes for a tissue-destroying toxin. Can you describe in a general term how a viral infection might cause a permanent change in the pathogenicity of a group A streptococcus bacterium? Considering the apparent difference in virulence between the bacterium in the 1920s and the present, is there any method for determining whether the same strain of group A streptococcus is responsible for both sets of cases? Note that preserved specimens of infected tissue from these early cases are available.

44. LINE1 (L1) elements compose approximately 17% of the human genome. The 5′ UTR of the L1 element contains an RNA polymerase II promoter. Describe how L1 element transposition occurs.
45. Because of overuse of antibiotics and/or weakened governmental surveillance of infectious disease, several superbug diseases that had been thought to be no longer a threat to human health (e.g., pneumonia and tuberculosis) are rapidly becoming unmanageable. In several instances, so-called superbugs (microorganisms that are resistant to almost all known antibiotics) have been detected. How did this circumstance arise? What will happen if this process continues?
46. Use the techniques identified in this chapter to describe in general terms how a researcher can map the genome of a newly discovered organism.
47. It was once thought that the DNA polymerase machinery moves along DNA in a manner analogous to a train on a track. Current evidence indicates that the polymerizing machinery is instead stationary and that the DNA strands are pumped through the complex. What advantages does this stationary mechanism have?

SECTION 18.2

Comprehension Questions

48. Define the following terms:

    a. promoter
    b. consensus sequence
    c. operon
    d. chromatin-remodeling complex
    e. general transcription factors

49. Define the following terms:

    a. RNA splicing
    b. spliceosome
    c. operator
    d. riboswitch
    e. lac operon

50. Define the following terms

    a. TATA box
    b. transcription factory
    c. P body
    d. CTCF
    e. supraspliceosomes

Fill in the Blanks

51. The creation of mRNA from DNA is called _____.
52. The TATA box is the binding site of _____.
53. Promoters aid in initiating _____.
54. The eukaryotic mRNA is protected from degradation by a 3′ _____.

Short-Answer Questions

55. What would you expect the functional properties of a riboswitch specific for the amino acid lysine to be?
56. The retinoblastoma tumor suppressor gene Rb (RB1) codes for the retinoblastoma protein (pRB). pRB prevents the progression of the cell cycle through G_1 if DNA has been damaged. It does so in part because it binds a transcription-activating dimer referred to as E2F-DP. The pRB/E2F-DP recruits a histone deacetylase to chromatin. Explain.
57. The tumor suppressor pRB also binds to and suppresses the activity of retinoblastoma-binding protein 2 (RBP2), a histone demethylase that removes methyl groups from di- and trimethylated lysines in histone 3. What is the possible consequence of an inactivating mutation in RB1 that causes an inability of pRB to bind RBP2?
58. All eukaryotes possess a surveillance pathway referred to as nonsense-mediated mRNA decay (NMD). Its principal function is to eliminate mRNA transcripts with premature stop codons. Such faulty transcripts are detected during translation and subsequently destroyed by removal of the 5′ -cap, followed by degradation by a nuclease. Describe how premature stop codons are detected and what type of error causes them.
59. Define and describe the roles of the following in transcription:

    a. transcription factors
    b. RNA polymerase
    c. promoter
    d. sigma factor
    e. enhancer
    f. TATA box

Critical-Thinking Questions

60. In eukaryotes, the DNA replication rate is 50 nucleotides per second. How long would the replication of a chromosome of 150 million base pairs take if eukaryotic chromosomes were replicated like those of prokaryotes? Actually, eukaryotic replication takes only several hours. How do eukaryotes achieve this high rate?
61. Explain how a reverse transcriptase activity within a cell can result in gene amplification.

SECTION 18.3

Comprehension Questions

62. Define the following terms:

    a. gene silencing
    b. RNA interference
    c. tumor suppressor gene
    d. protooncogene

e. GEF

63. Define the following terms:

a. cell transformation
b. oncogene
c. apoptosis
d. early response gene
e. delayed response gene

Fill in the Blanks

64. An RNA metabolite-sensing control mechanism is called a _____.
65. The lac operon consists of the control elements and _____ that code for the enzymes of lactose metabolism.
66. _____ is the term used for the programmed death of a cell.
67. Proteins that actively protect cells from progressing toward cancer are called _____.

Short-Answer Questions

68. Describe the advantages and disadvantages for organisms that arrange genes in operons.
69. How can viruses cause mutations?
70. Many genes generate different products depending on the type of cell expressing the gene. How is this phenomenon accomplished?
71. During certain stages of cell growth, the requirement for certain gene products may require gene amplification. What purpose does gene amplification serve?
72. How are the genes referred to in Question 71 amplified?
73. Mutations are caused by chemical and physical phenomena. Indicate the type of mutation that each of the following reactions or molecules might cause:

a. ROS
b. intercalating agents
c. a small alkylating agent
d. a large alkylating agent
e. nitrous acid

Critical-Thinking Questions

74. Phorbol esters have been observed to induce the transcription of AP-1 influenced genes. Explain how this process could occur. What are the consequences of AP-1 transcription? What role does intermittent exposure to phorbol esters have in an individual's health?
75. There appears to be insufficient genetic material to direct all the activities of certain types of eukaryotic cell (e.g., such as T lymphocytes). What genetic processes solve this problem?
76. Retinoblastoma is a rare cancer in which tumors develop in the retina of the eye. The tumors arise because of the loss of the *Rb* gene, which codes for a tumor suppressor. Explain why hereditary retinoblastomas usually appear in children who have only one functional copy of *Rb* and why the nonhereditary form of retinoblastoma usually occurs later in life.
77. Adjacent pyrimidine bases in DNA form dimers with high efficiency after exposure to UV light. If these dimers are not repaired, skin cancers can result. Melanin is a natural sunscreen produced by melanocytes, a type of skin cell, when the skin is exposed to sunlight. Individuals who spend long periods developing a tan eventually acquire a thick and highly wrinkled skin. Such individuals are also at high risk for skin cancer. Can you explain, in general terms, why these phenomena are related?

78. Explain the difference between the potential effects on an individual organism of errors made during replication and those made during transcription.
79. Explain how features of a riboswitch can change when activated by excess metabolite concentration.

MCAT Study Questions

80. The enzyme responsible for the joining of Okazaki fragments is

    a. helicase
    b. primase
    c. DNA ligase
    d. DNA topoisomerase

81. The leading strand in DNA replication requires only one primer for its synthesis. The lagging strand requires several primers because

    a. the lagging strand is prepared in short segments
    b. the lagging strand lacks a topoisomerase
    c. the leading strand serves as a template for the lagging strand
    d. All DNA polymerases read from the 5′ end.

82. The stress of supercoiling during replication is released by

    a. DNA polymerase
    b. DNA topoisomerase
    c. exonuclease
    d. DNA helicase

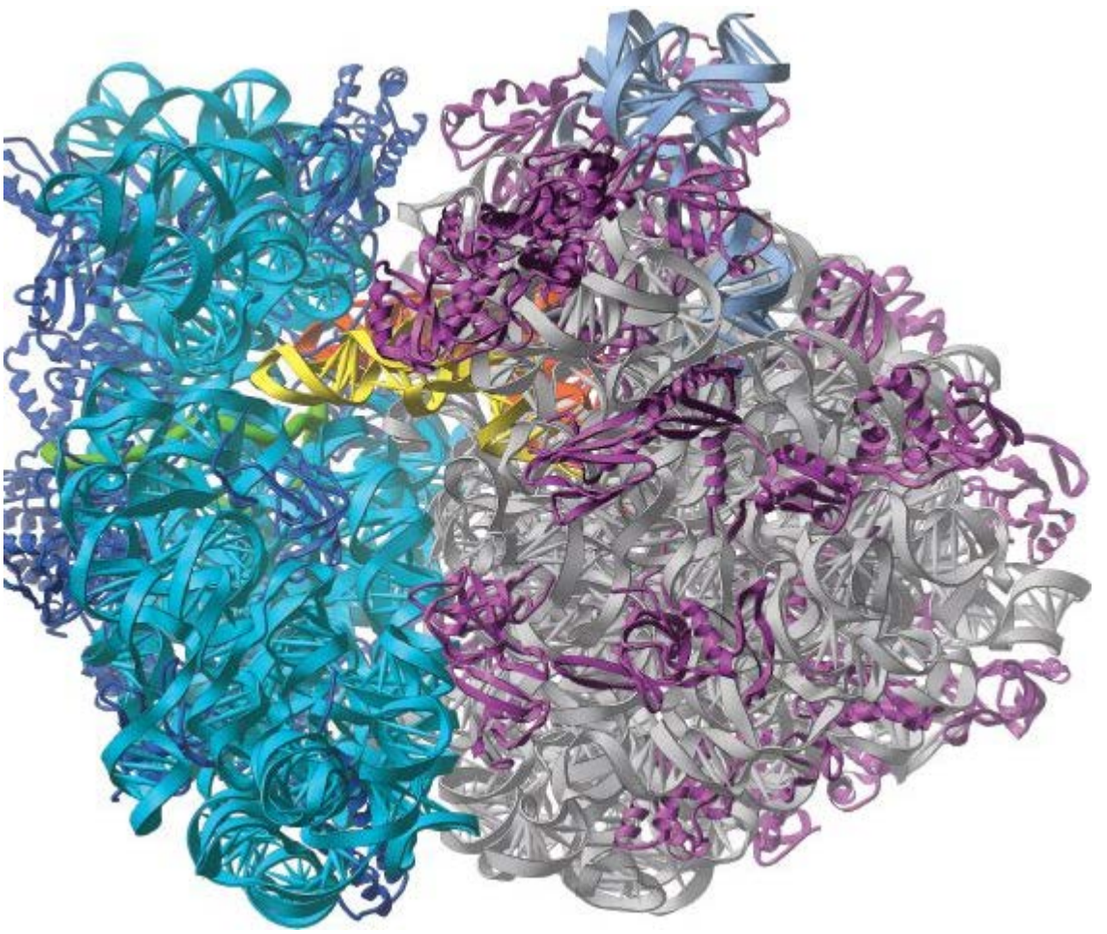
83. Enzymes or proteins involved in DNA repair include

    a. RAD52
    b. RuvA
    c. UvrA
    d. MAPK

84. Molecules in higher eukaryotes that are directly involved in gene silencing include all of the following except

    a. miRNA
    b. argonaute
    c. microprocessor
    d. glycosylase

CHAPTER 19 Protein Synthesis



The Ribosome Ribosomes are ribonucleoprotein molecular machines that synthesize proteins in all living cells. In this illustration of the high-resolution structure of a complete bacterial 70S ribosome, proteins are shown in dark blue and magenta, and rRNA molecules in cyan and gray. The tRNAs are orange and yellow. Note that the ribosome is primarily composed of rRNA, which performs most catalytic activities. Protein molecules largely serve supporting roles.

OUTLINE

MRSA: THE SUPERBUG

19.1 THE GENETIC CODE

Codon–Anticodon Interactions

The Aminoacyl-tRNA Synthetase Reaction

19.2 PROTEIN SYNTHESIS

Prokaryotic Protein Synthesis
Eukaryotic Protein Synthesis

19.3 THE PROTEOSTASIS NETWORK
The Heat Shock Response
The Proteostasis Network and Human Disease

Biochemistry in Perspective

Trapped Ribosomes: RNA to the Rescue!

Biochemistry in Perspective

Context-Dependent Coding Reassignment

Biochemistry in the Lab

Proteomics

AVAILABLE ONLINE

Biochemistry in Perspective

EF-Tu: A Motor Protein

MRSA: The Superbug

If you were to ask Sam about his leg infection, it is unlikely that he could tell you how or even exactly when it began. Skin abrasions are very frequent with physically active teenagers, especially on high school football teams. The torn skin on his left lower leg probably occurred during a practice game. After a quick toweling off in the locker room, this busy, stressed high school student was off to his next activity. The leg didn't become bothersome until several days later, when the abrasion became red and itchy, followed later by slight swelling and warmth. Still, Sam was not overly concerned: his experiences with cuts and scrapes had never amounted to anything that couldn't be remedied with a few Band-Aids.

About a week later, when Sam mentioned it to his parents, the now abscessed (pus-releasing) lesion had become very painful. When an antibiotic prescription failed to heal the leg and red streaks just below the skin began to appear, it was apparent to everyone that Sam had a very serious infection. On the morning of Sam's first appointment with an infection specialist, four days after the initial visit to his family doctor, his symptoms (fever, chills, weakness, and nausea), though still relatively mild, indicated that he was in the beginning stages of septicemia (infection of the bloodstream by microorganisms). After the specialist sent a specimen to be cultured, Sam was diagnosed with an MRSA-caused infection.

MRSA (methicillin-resistant *Staphylococcus aureus*) is a strain of the Gram-positive bacterium *S. aureus* that is resistant to the β-lactams, a class of β-lactam ring–containing antibiotics that includes the penicillins and penicillin-related antibiotics (e.g., the cephalosporins). The β-lactams kill susceptible bacteria by stopping transpeptidation, a key reaction in cell wall synthesis. Resistant bacteria synthesize β-lactamase, an enzyme that hydrolyzes the β-lactam ring, thus inactivating the antibiotic.

Antibiotic resistance is not restricted to the β-lactams. It is an evolutionary phenomenon caused by a selection pressure, in this case the unrestricted use of antibiotics. The widespread use of a specific antibiotic provides any bacterium with a plasmid that contains an antibiotic-resistance gene, allowing it to evade the antibiotic's action against the pathogen with a significant growth advantage. In other words, the antibiotic removes the resistant bacterium's competition for resources. Some strains of bacteria, called superbugs, are resistant to several antibiotics because they possess several resistance genes.

In most cases of MRSA infection, the drug of choice is vancomycin. Vancomycin kills Gram-positive bacteria by interfering with transglycosylation, yet another reaction in cell wall formation. To the horror

of the hospital physicians and Sam's parents, lab cultures of the organism replicating in Sam's bloodstream proved it to be insensitive to vancomycin as well. One of the few remaining antibiotic treatments for MRSA infections is linezolid, a molecule that disrupts the initiation phase of bacterial protein synthesis. Within a few days after beginning intravenous therapy with linezolid, Sam began to recover. After 1 week, he was released from the hospital with a prescription for linezolid tablets. Both Sam and his parents were warned that he had to adhere rigorously to the medication schedule (one tablet every 12 hours for 7 days). Failure to do so would create the risk of a relapse, a serious threat to Sam's health. Any further treatment with linezolid could result in toxic side effects such as an irreversible peripheral neuropathy. It could even lead to resistance of the organism to linezolid. Without any antibiotics proven effective against this particular strain of MRSA, Sam's life would be threatened.

Sam's parents waited until he recovered from the infection to tell him that Jake, a fellow member of his team, had also been infected with MRSA. For unknown reasons, Jake's infection was more aggressive than Sam's. Despite the heroic efforts of his doctors, Jake died after two weeks in the hospital. Afterward, the school district temporarily closed the high school so that it could be thoroughly sanitized. The students were also informed about prevention strategies. In addition to routine cleaning of sports equipment, students were warned that such simple behaviors as hand washing, the covering of open cuts and abrasions, and not sharing towels or razors could prevent this tragedy from happening again.

Overview

PROTEINS ARE THE MOST DYNAMIC, NUMEROUS, AND VARIED CLASS OF BIOMOLECULES. THE UNIQUENESS OF EACH CELL TYPE IS CAUSED ALMOST entirely by the proteins it produces. It is not surprising, therefore, that a relatively large amount of cellular energy is used in protein synthesis. Because of their strategic importance in the cellular economy, the synthesis of proteins is a regulated process. Although control is also of major importance at the transcriptional level, control of the translation of genetic messages allows for additional opportunities for regulation. This is especially true in multicellular eukaryotes, whose complex lifestyles require amazingly diverse regulatory mechanisms.

Protein synthesis is the process in which genetic information encoded in the nucleic acids is translated into the 20 standard amino acid "alphabet'" of polypeptides. In addition to translation (the mechanism by which a nucleotide base sequence directs the polymerization of amino acids), protein synthesis includes the processes of posttranslational modification and targeting. *Posttranslational modification* consists of the chemical alterations cells use to prepare polypeptides for their functional roles. Several modifications assist in *targeting*, which directs newly synthesized molecules to a specific intracellular or extracellular location.

In all, at least 100 different molecules are involved in protein synthesis. Among the most important of these molecules are the components of the ribosomes, large ribonucleoprotein machines that synthesize polypeptides. Each ribosome "reads" the base sequence of an mRNA and, fueled by GTP, rapidly and precisely converts this information into the amino acid sequence of a polypeptide. Speed is required because organisms must respond expeditiously to ever-changing environmental conditions. In prokaryotes such as *E. coli*, for example, a polypeptide of 100 residues is synthesized in less than 6 s. Eukaryotes are slower, at about 2 residues per second. Precision in mRNA translation is critical because the proper functioning of each polypeptide is

determined not only by the molecule's primary sequence but also by its accurate folding.

As a result of the great expense involved in protein synthesis, living organisms employ mechanisms that coordinate energy in the form of nutrient availability to translation regulation. In *E. coli*, for example, one consequence of low amino acid levels is the synthesis of the GTP derivative guanosine 5'-diphosphate-3'-diphosphate (ppGpp), which immediately inhibits the transcription of translation component genes (e.g., genes coding for rRNAs, ribosomal proteins, amino-acyl-tRNA synthetases, and t-RNAs). In eukaryotes, low amino acid levels trigger the homodimerization of GCN2, a protein that phosphorylates the initiation factor eIF2α, thereby inhibiting the translation of most mRNAs. Other molecules that regulate eukaryotic protein synthesis based on nutrient availability are AMPK (p. 470) and the mTORC1 complex (p. 612).

As the result of decades of intense work, an increasingly more detailed understanding of ribosome structure and function is emerging. One of the unexpected findings is that rRNA performs the critical functions of the ribosome. For example, the catalytic activity that forms peptide bonds resides in an RNA molecule. In addition, rRNA molecules also have roles in tRNA-mRNA docking, ribosomal subunit association, proofreading, and the binding of translation factors. Ribosomal proteins, for the most part, have supporting roles. In recent years, it has also become apparent that the ribosome is a major component of the cellular protein quality control process.

Chapter 19 provides an overview of protein synthesis. The chapter begins with a discussion of the genetic code, the mechanism by which nucleic acid–base sequences specify the amino acid sequences of polypeptides. This is followed by discussions of protein synthesis as it occurs in both prokaryotes and eukaryotes and a description of the mechanisms that convert polypeptides into their biologically active conformations. A critical feature of this process, protein folding, was described in Section 5.3. The chapter ends with a section devoted to the role of the proteostasis network in protein folding and an introduction to **proteomics**, a technology used to characterize the protein products of the genome.

19.1 THE GENETIC CODE

Translation is fundamentally different from the transcription process that precedes it. During transcription, the language of DNA sequences is converted to the closely related dialect of RNA sequences. During protein synthesis, however, a nucleic acid–base sequence is converted to a clearly different language (i.e., an amino acid sequence), hence the term *translation*. Researchers were at first at a loss to explain how an mRNA base code could be converted into an amino acid polymer. Then Francis Crick realized that a series of adaptor molecules must mediate the translation process. This role was eventually assigned to tRNA molecules (**Figure 17.25**).

Before adaptor molecules could be identified, however, a more important problem had to be solved: deciphering the genetic code. The **genetic code** can be described as a coding dictionary that specifies a meaning for each base sequence. Once the importance of the genetic code was recognized, investigators speculated about its dimensions. Only four different bases (G, C, A, and U) occur in mRNA, but 20 amino acids must be specified. It therefore appeared reasonable that a combination of bases codes for each amino acid. A sequence of two bases would specify only a total of 16 amino acids (i.e., $4^2 = 16$). However, a three-base sequence, or **codon**, provides more than sufficient base combinations for translation (i.e., $4^3 = 64$).

The codon assignments for the 64 possible trinucleotide sequences, determined by Marshall Nirenberg, Heinrich Matthaei, and Har Gobind Khorana, are presented in **Table 19.1**. Of these, 61 code for amino acids. Four codons serve as punctuation signals. UAA, UAG, and UGA are *stop*

(polypeptide chain terminating) signals. AUG, the codon for methionine, also serves as a *start* signal (sometimes referred to as the *initiating codon*). The genetic code possesses the following properties.

1. **Specific**. Each codon is a signal for a specific amino acid.
2. **Degenerate**. Any coding system in which several signals have the same meaning is said to be degenerate. The genetic code is partially degenerate because most amino acids are coded for by several codons. For example, leucine is coded for by six different codons (UUA, UUG, CUU, CUC, CUA, and CUG). In fact, methionine (AUG) and tryptophan (UGG) are the only amino acids that are coded for by a single codon.

TABLE 19.1 The Genetic Code

First position (5′ end)	Second Position: U	Second Position: C	Second Position: A	Second Position: G	Third position (3′ end)
U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA STOP UAG* STOP	UGU Cys UGC Cys UGA* STOP UGG Trp	U C A G
C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	U C A G
A	AUU Ile AUC Ile AUA Ile AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U C A G
G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U C A G

*The stop codons UGA and UAG are also used by some organisms and under specific conditions to specify molecules other than the standard amino acids. The Biochemistry in Perspective: Context-Dependent Coding Reassignment, p. 775) describes how UGA is used to insert selenocysteine into a polypeptide sequence.

3. **Nonoverlapping**. The mRNA coding sequence is "read" by a ribosome starting from the initiating codon (AUG) as a continuous sequence taken three bases at a time until a stop codon is reached. The set of contiguous triplet codons in an mRNA that code for the amino acids in a polypeptide is called an **open reading frame**.
4. **Almost universal**. The genetic code is used by the vast majority of organisms, but there are a few exceptions. Minor deviations have been observed in mitochondria. For example, instead of six arginine codons, there are only four. The remaining two codons (AGA and AGG) are instead used as stop codons. Since UGA, normally a stop codon, instead codes for tryptophan, mitochondria have four stop codons: UAG, UAA, AGA, and AGG. Similar minor changes have been observed in a few species of prokaryotes and eukaryotes such as protozoa and yeast.

It appears that the degenerate genetic code evolved to diminish the deleterious effects of point mutations; that is, base substitutions often result in the incorporation of the same or similar amino acids, as the following two examples involving leucine illustrate. Since all the codons with CU in the first two positions code for leucine, base substitutions at position 3 will have no effect. (Refer to the discussion of the wobble hypothesis, p. 756, for another aspect of this phenomenon.) An

analysis of the genetic code also reveals that all of the codons with U in the second position code for hydrophobic amino acids. If the first base in the CUU codon (leucine) is replaced with A, another hydrophobic amino acid residue, isoleucine, will be substituted.

In certain circumstances, living organisms are not limited to the codon assignments of the genetic code. In one example, the nonstandard amino acid selenocysteine is coded for by a stop codon. This context-dependent codon reassignment is described later in the Biochemistry in Perspective essay on p. 775.

KEY CONCEPTS

The genetic code is a mechanism by which ribosomes translate nucleotide base sequences into the primary sequence of polypeptides.

QUESTION 19.1

As described earlier, DNA damage can cause deletion or insertion of base pairs. If a nucleotide base sequence of a coding region changes by any number of bases other than three base pairs, or multiples of 3, a *frameshift mutation* occurs. Depending on the location of the sequence change, such mutations can have serious effects. The following synthetic mRNA sequence codes for the beginning of a polypeptide:

5′-AUGUCUCCUACUGCUGACGAGGGAAGGAGGUGGCUUAUC-AUGUUU-3′

First, determine the amino acid sequence of the polypeptide. Then determine the types of mutation that have occurred in the following altered mRNA segments. What effect do these mutations have on the polypeptide products?

a. 5′-AUGUCUCCUACUUGCUGACGAGGGAAGGAGGUGGCUUAUCA-UGUUU-3′
b. 5′-AUGUCUCCUACUGCUGACGAGGGAGGAGGUGGCUUAUCAU-GUUU-3′
c. 5′-AUGUCUCCUACUGCUGACGAGGGAAGGAGGUGGCCCUUAUC-AUGUUU-3′
d. 5′-AUGUCUCCUACUGCUGACGGAAGGAGGUGGCUUAUCAU-GUUU-3′

Codon–Anticodon Interactions

Transfer RNA molecules are the “adaptors” required for the translation of the genetic message. Each type of tRNA carries a specific amino acid (at the 3′ terminus) and possesses a three-base sequence called the **anticodon**. Codon–anticodon base pairing is responsible for the actual translation of mRNAs. Although codon–anticodon pairings are antiparallel, both sequences are given in the 5′ → 3′ direction. For example, the codon UGC binds to the anticodon GCA (**Figure 19.1**).

Once the genetic code had been determined, researchers anticipated the identification of 61 types of tRNAs in living cells. Instead, they discovered that cells often operate with substantially fewer tRNAs than expected. Most cells possess about 50 tRNAs, although lower numbers have been observed; bacteria such as *E. coli* have at least 30 RNAs. Further investigation of tRNAs revealed that the anticodon in some molecules contains inosinate (I), which typically occurs at the third anticodon position. (The base adenine is deaminated to form hypoxanthine; see **Figure 14.25** on p. 556.) As tRNAs were investigated further, it became increasingly clear that some molecules

recognize several codons. In 1966, after reviewing the evidence, Crick proposed a rational explanation, the wobble hypothesis.

KEY CONCEPTS

- The genetic code is translated through base-pairing interactions between mRNA codons and tRNA anticodons.
- The wobble hypothesis explains why cells usually have fewer tRNAs than expected.

The **wobble hypothesis** allows for multiple codon–anticodon interactions by individual tRNAs during translation. It is based principally on the following observations.

1. The first two base pairings in a codon–anticodon interaction confer most of the specificity required during translation since most redundant codons specifying a certain amino acid possess identical nucleotides in the first two positions. These interactions are standard (i.e., Watson–Crick) base pairings.
2. The interactions between the third codon and anticodon nucleotides are less stringent. In fact, nontraditional base pairs (i.e., non-Watson–Crick) often occur. For example, tRNAs containing G in the 5′ (or "wobble") position of the anticodon can pair with two different bases; that is, G can interact with either C or U. The same is true for U, which can interact with A or G (**Figure 19.2a**). When I is in the wobble position of an anticodon, a tRNA can base-pair with three different codons because I can interact with U or A or C (**Figure 19.2b**).

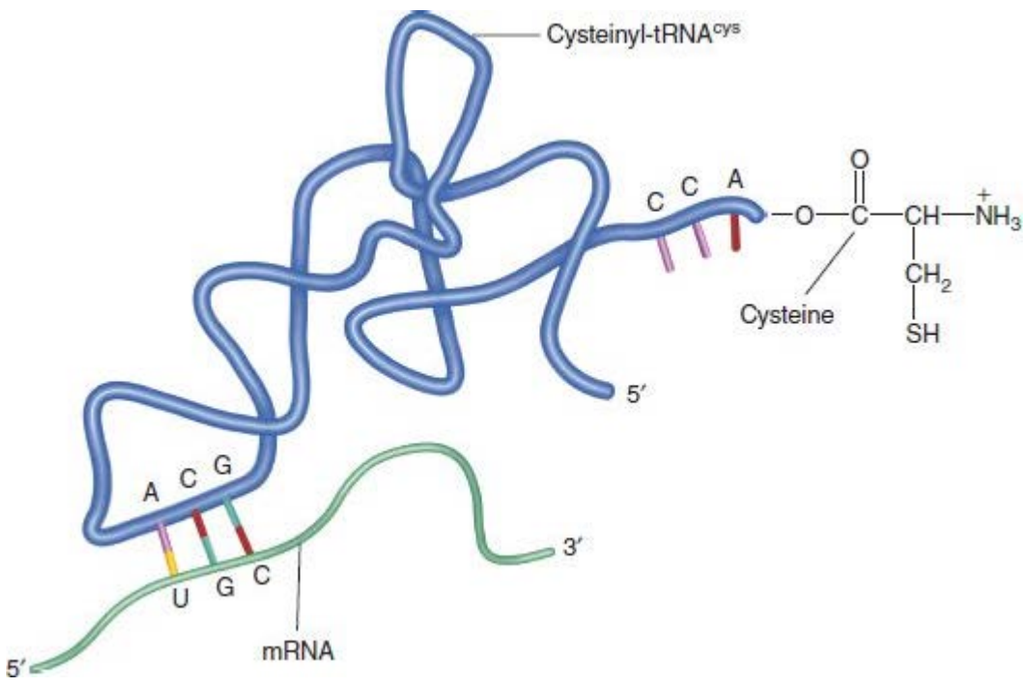


FIGURE 19.1

Codon–Anticodon Base Pairing of Cysteinyl-tRNA^{cys}

The pairing of the codon UGC with the anticodon GCA ensures that the amino acid cysteine will be incorporated into a growing polypeptide chain.

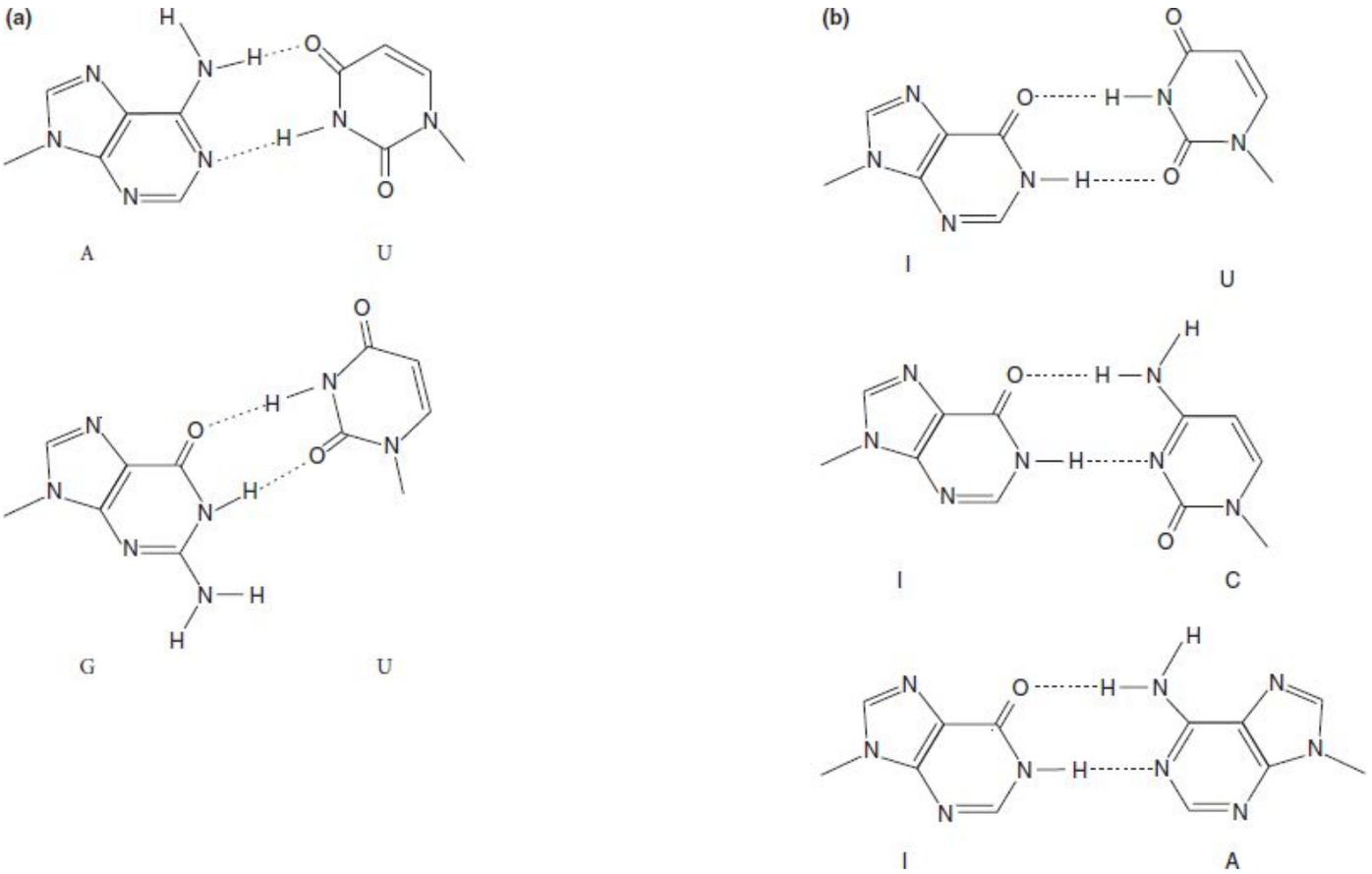


FIGURE 19.2

Wobble Base Pairs

Nonstandard base pairing is critical for the translation of the genetic code. Examples of wobble base pairs include (a) AU and GU base pairs and (b) IU, IC, and IA base pairs.

A careful examination of the genetic code and the "wobble rules" indicates that all 61 codons can be translated with a minimum of 31 tRNAs. An additional tRNA for initiating protein synthesis brings the total to 32 tRNAs.

QUESTION 19.2

The sequence of a DNA segment is GGTTTA. What is the sequence of the tRNA anticodons?

QUESTION 19.3

The amino acid sequence for a short peptide is Tyr-Leu-Thr-Ala. What are the possible base sequences of the mRNA and the transcribed DNA strand that code for it? What are the anticodons?

The Aminoacyl-tRNA Synthetase Reaction

Although the accuracy of translation (approximately one error per 10^4 amino acids incorporated) is

lower than that of DNA replication and transcription, it is remarkably higher than one would expect of such a complex process. The principal reasons for the accuracy with which amino acids are incorporated into polypeptides include codon–anticodon base pairing and the mechanism by which amino acids are attached to their cognate tRNAs. The attachment of amino acids to tRNAs, considered the first step in protein synthesis, is catalyzed by a group of enzymes called the *aminoacyl-tRNA synthetases*.

In most organisms, there is at least one aminoacyl-tRNA synthetase for each of the 20 amino acids. Each enzyme links its specific amino acid to an appropriate tRNA. The process that links an amino acid to the 3′ terminus adenosine residue of the correct tRNA consists of two sequential reactions (**Figure 19.3**), both of which occur within the active site of the synthetase.

1. **Activation**. The synthetase first catalyzes the formation of aminoacyl-AMP. This reaction, which activates the amino acid by forming a high-energy mixed anhydride bond, is driven to completion through the hydrolysis of its other product, pyrophosphate. (An **anhydride** is a molecule containing two carbonyl groups linked through an oxygen atom. The term **mixed anhydride** describes an anhydride formed from two different acids, for example, a carboxylic acid and phosphoric acid.)
2. **tRNA linkage**. A synthetase binds an appropriate tRNA's D arm and the amino acid is transferred from the aminoacyl-AMP to a ribose hydroxyl group of the tRNA's 3′-terminal nucleotide forming an ester linkage. (Depending on the synthetase, aminoacylation may be through the 2′-OH or 3′-OH of the ribose moiety of the nucleotide. Subsequently, the aminoacyl group can migrate between the 2′-OH and 3′-OH groups. Only the 3′-aminoacyl esters are used during translation.) Although the aminoacyl ester linkage to the tRNA is lower in energy than the mixed anhydride of aminoacyl AMP, it still possesses sufficient energy to participate in acyl transfer reactions (peptide bond formation).

The sum of the reactions catalyzed by the aminoacyl-tRNA synthetases is as follows:

$$\text{Amino acid} + \text{ATP} + \text{tRNA} \rightarrow \text{aminoacyl-tRNA} + \text{AMP} + \text{PP}_i$$

The product PP_i is immediately hydrolyzed with a large loss of free energy. Consequently, tRNA charging is irreversible. Because AMP is a product of this reaction, the metabolic price for the linkage of each amino acid to its tRNA is the equivalent of the hydrolysis of two molecules of ATP to ADP and P_i.

The aminoacyl-tRNA synthetases are a diverse group of enzymes that vary in molecular weight, primary sequence, and number of subunits. Each enzyme efficiently produces a specific aminoacyl-tRNA product with relative accuracy. The specificity with which each of the synthetases binds the correct amino acid and its cognate tRNA is crucial for the fidelity of the translation process. Some amino acids can easily be differentiated by their size (e.g., tryptophan vs. glycine) or the presence of positive or negative charges in their side chains (e.g., lysine vs. aspartate). Other amino acids, however, are more difficult to discriminate because their structures are similar. For example, isoleucine and valine differ only by a methylene group. Despite this difficulty, isoleucyl-$tRNA^{ile}$ synthetase usually synthesizes the correct product. However, this enzyme occasionally also produces valyl$tRNA^{ile}$. Isoleucyl-$tRNA^{ile}$ synthetase, as well as several other synthetases, can correct such a mistake because it possesses a separate *proofreading* site. Because of its size, this site binds valyl-$tRNA^{ile}$ and excludes the larger isoleucyl-$tRNA^{ile}$. After its binding in the proofreading site, the ester bond of valyl-$tRNA^{ile}$ is hydrolyzed.

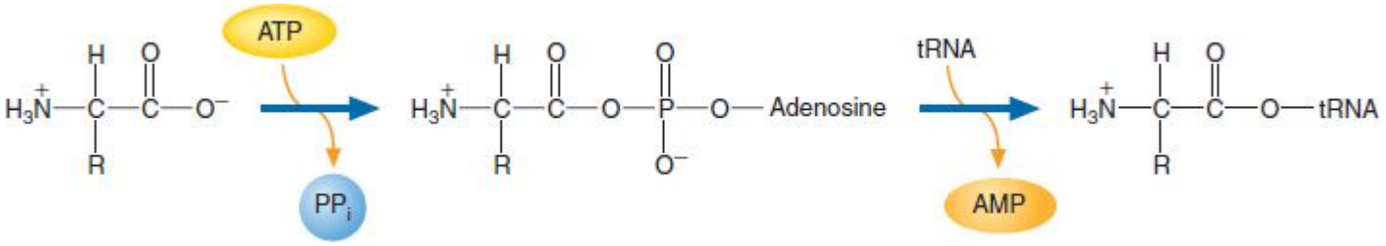


FIGURE 19.3

Formation of Aminoacyl-tRNA

Each aminoacyl-tRNA synthetase catalyzes two sequential reactions in which an amino acid is linked to the 3′-terminal ribose residue of the tRNA molecule.

19.2 PROTEIN SYNTHESIS

An overview of protein synthesis is illustrated in **Figure 19.4**. Despite its complexity and the variations among species, the translation of a genetic message into the primary sequence of a polypeptide can be divided into three phases: initiation, elongation, and termination.

1. **Initiation**. Translation begins with **initiation**, when the small ribosomal subunit binds an mRNA. The anticodon of a specific tRNA, referred to as an *initiator tRNA*, then base-pairs with the initiation codon AUG on the mRNA. Initiation ends as the large ribosomal subunit combines with the small subunit. There are two sites on the complete ribosome for the codon–anticodon interactions involved in translation: the P (peptidyl) site (now occupied by the initiator tRNA), the A (aminoacyl) site, and an E or exit site. The E site is occupied by an uncharged tRNA before it is released from the ribosome. In both prokaryotes and eukaryotes, mRNAs are read simultaneously by numerous ribosomes. An mRNA with several ribosomes bound to it is referred to as a **polysome**. In actively growing prokaryotes, for example, the ribosomes attached to an mRNA molecule may be separated from each other by as few as 80 nucleotides.
2. **Elongation**. During the **elongation** phase, the polypeptide is synthesized according to the specifications of the genetic message. The mRNA base sequence is read in the 5′ → 3′ direction, and polypeptide synthesis proceeds from the N-terminal to the C-terminal. The elongation cycle is composed of three steps: codon–anticodon pairing in the A site, peptide bond formation, and the transfer of the peptidyl-tRNA to the P site. Elongation begins as the next aminoacyl-tRNA binds to the A site as the result of codon–anticodon pairing. Peptide bond formation is then catalyzed by peptidyl transferase. In this *transpeptidation* reaction, the α-amino nitrogen of the A site amino acid (the nucleophile) attacks the carbonyl group of the P site amino acid (**Figure 19.5**). As a result of peptide bond formation, the growing peptide chain is now attached to the A site tRNA. Finally, **translocation** occurs. As the ribosome moves along the mRNA one triplet length, the peptidyl chain linked to the A site tRNA is shifted to the P site, and the uncharged tRNA in the P site is released from the ribosome (eukaryotes) or shifted to the E or exit site (prokaryotes) where it is subsequently released from the ribosome. As the A site is vacated, the next codon, now positioned in the A site, binds to its cognate tRNA anticodon. This elongation cycle is repeated until a stop codon enters the A site.
3. **Termination**. During **termination**, the polypeptide chain is released from the ribosome. Translation terminates because a stop codon cannot bind an aminoacyl-tRNA. Instead, a protein-releasing factor binds to the A site. Subsequently, peptidyl transferase (acting as an

esterase) hydrolyzes the bond connecting the now-completed polypeptide chain and the tRNA in the P site. Translation ends as the ribosome releases the mRNA and dissociates into the large and small subunits.

In addition to the ribosomal subunits, mRNA, and aminoacyl-tRNAs, translation requires an energy source (GTP) and a wide variety of protein factors. These factors perform several roles. Some have catalytic functions; others stabilize specific structures that form during translation. Translation factors are classified according to the phase of the translation process they affect, that is, initiation, elongation, or termination. The major differences between prokaryotic and eukaryotic translation appear to be largely a result of the identity and functioning of these protein factors.

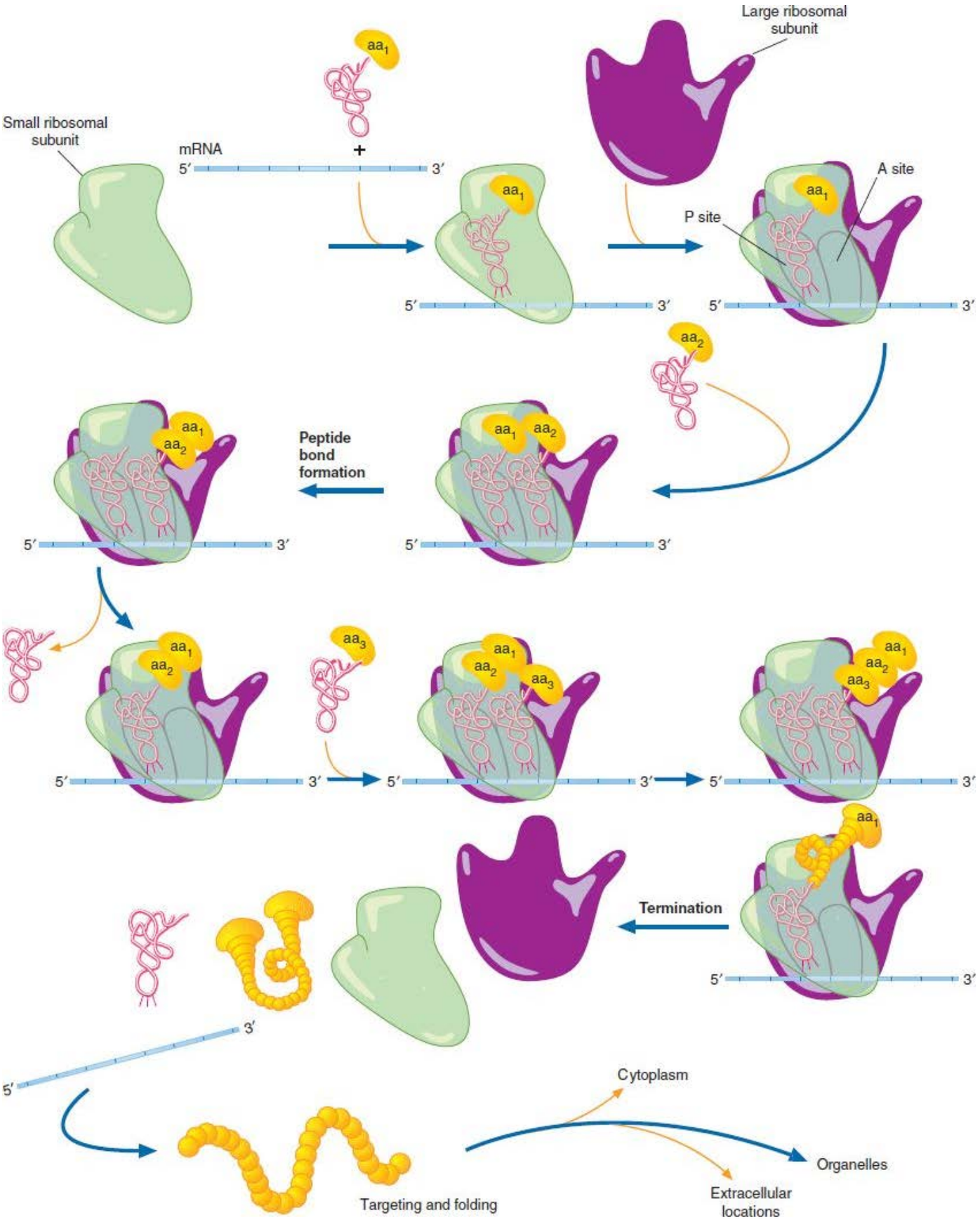


FIGURE 19.4

Protein Synthesis

No matter what the organism, translation consists of three phases: initiation, elongation, and termination. The elongation reactions, which include peptide bond formation and translocation, are repeated many times until a stop codon is reached. The numerous protein factors that facilitate each step in protein synthesis are different in

prokaryotes and eukaryotes. Posttranslational reactions and targeting processes vary according to cell type.

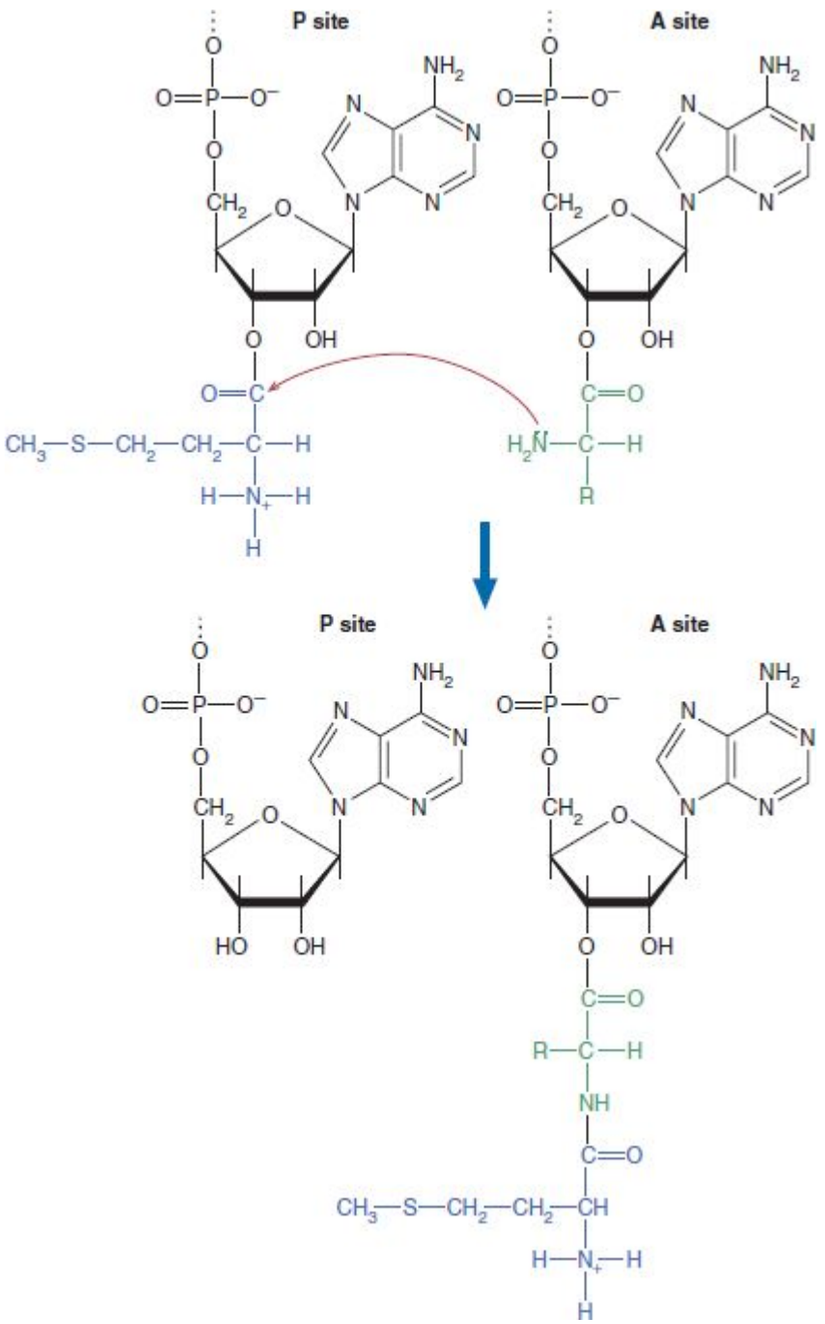


FIGURE 19.5

Peptide Bond Formation

Elongation begins when a peptide bond forms because of the nucleophilic attack of the A-site amino acid's amino group on the carbonyl carbon of the methionine residue linked to the initiating tRNA in the P site. Because a peptide bond has formed, both amino acids are now attached to the A site tRNA.

Regardless of the species, immediately after translation, some polypeptides fold into their final form without further modifications. Frequently, however, newly synthesized polypeptides are modified. These alterations, referred to as **posttranslational modifications**, can be considered the fourth phase of translation. They may include removal of portions of the polypeptide by proteases,

chemical modification of the side chains of certain amino acid residues, and insertion of cofactors. Often, individual polypeptides then combine to form multisubunit proteins.

Posttranslational modifications appear to serve two general purposes: (1) to prepare a polypeptide for its specific function and (2) to direct a polypeptide to a specific location, a process referred to as **targeting**. Targeting is an especially important process in eukaryotes because proteins must be precisely directed to a vast array of possible destinations. In addition to cytoplasm and the plasma membrane (the principal destinations in prokaryotes), eukaryotic proteins may be sent to a variety of organelles (e.g., mitochondria, chloroplasts, lysosomes, or peroxisomes) or they may be secreted from the cell.

KEY CONCEPTS

- Translation consists of three phases: initiation, elongation, and termination.
- After synthesis, many proteins are chemically modified and targeted to specific cellular or extracellular locations.

Although there are many similarities between prokaryotic and eukaryotic protein synthesis, there are also notable differences. Consequently, the details of prokaryotic and eukaryotic processes are discussed separately. Each discussion is followed by a brief description of mechanisms that control translation.

Prokaryotic Protein Synthesis

Protein synthesis in bacteria takes place on 2.4-MDa ribosomes capable of polymerizing amino acids at a rate of approximately 20 per second. The 70S bacterial ribosome is composed of a large 50S subunit and a small 30S subunit (**Figure 19.6**). The large subunit (about 1.5 MDa) consists of 23S and 5S rRNAs and 34 proteins. The small subunit (about 0.8 MDa) contains a 16S rRNA and 21 proteins. In addition to the P, A, and E sites, there are three other functional centers: the decoding center, the peptidyl transferase center, and the GTPase-associated region.

The **decoding center**, located at the A site on the 30S subunit, is where an mRNA codon is matched to an incoming tRNA anticodon. Three highly conserved 16S rRNA bases (A1492, A1493, and G 530) are contiguous with the codon–anticodon base pair triplet. When correct Watson–Crick base pairs have formed between the first two base pairs of the codon–anticodon base pair triplet, a conformational change occurs in A1492 and A1493 that accelerates the tRNA selection phase of the elongation cycle.

The **peptidyl transferase center** (PTC), where peptide bond formation occurs, is located in a cleft in the large subunit containing a domain of 23S rRNA. The core of the PTC, which consists of five conserved bases (A2451, U2505, U2585, C2452, and A2602), binds to the 3′ ends of the aminoacyl- and peptidyl-tRNAs. Peptide bond formation is the result of a concerted proton shuttling mechanism (pp. 764–65) that is triggered when aminoacyl-tRNA binding occurs in the precisely organized rRNA nucleotides within the PTC active site.

The **GTPase associated region** (GAR) is a set of overlapping binding sites on the 50S subunit composed of 23S rRNA structural elements. GAR acts as a GAP (p. 607). When translation factors with GTPase activity interact with GAR, the resulting GTP hydrolysis drives a conformational change in the protein that affects a translation event. The L12 stalk of the large subunit recruits the GTPase translation factors and facilitates their binding to GAR.

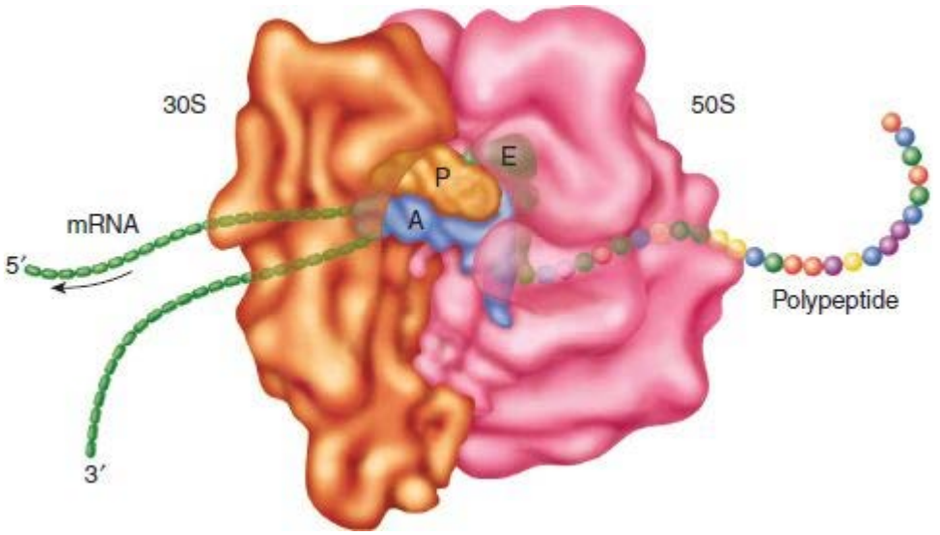


FIGURE 19.6

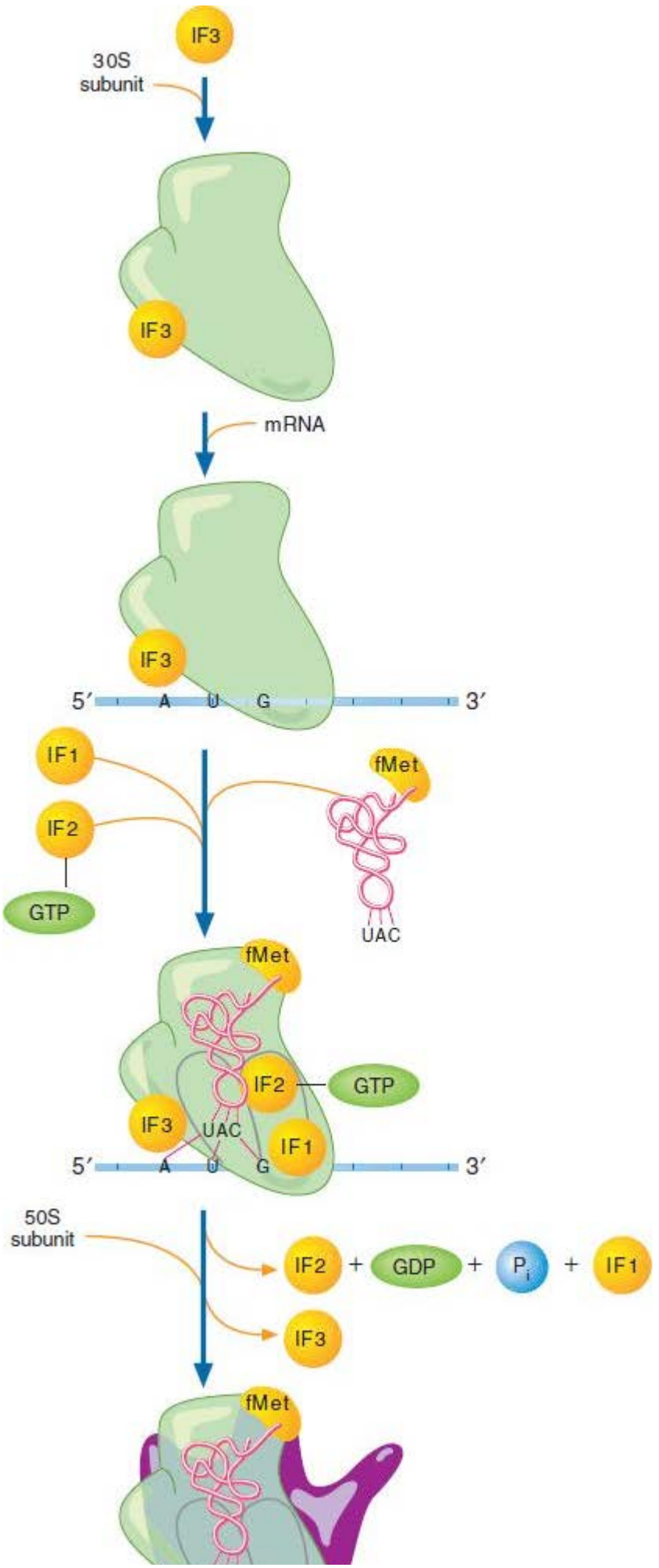
The Functional Ribosome

In this three-dimensional reconstruction of an *E. coli* ribosome during protein synthesis, the large and small subunits are shown in pink and orange, respectively. The relative positions of the mRNA, tRNAs, and the growing polypeptide chain are also illustrated. The tRNAs are identified as A and P to indicate their positions within the acyl and peptidyl sites where peptide bond formation occurs. The tRNA labeled E is in the exit position; that is, having discharged its amino acid during ongoing protein synthesis, it is in the process of leaving the ribosome. The movement of the mRNA through the ribosome is indicated by an arrow.

INITIATION Translation begins with the formation of an initiation complex (**Figure 19.7**). In prokaryotes such as *E. coli*, this process requires three initiation factors: IF1, IF2, and IF3. IF3 binding to the 30S subunit prevents that subunit from binding prematurely to the 50S subunit and promotes rapid codon–anticodon pairing for the initiator tRNA bound to IF2. IF1 binds to the A site of the 30S subunit, thereby blocking it during initiation. As an mRNA binds to the 30S subunit, it is guided into a precise location (so that the initiation codon AUG is correctly positioned) by a purine-rich sequence referred to as the **Shine–Dalgarno sequence**. The Shine–Dalgarno sequence (AGGAGG), which occurs a short distance upstream from AUG, binds to a complementary sequence, referred to as the anti–Shine–Dalgarno sequence, which is contained in the 16S rRNA component of the 30S subunit. Base pairing between the Shine–Dalgarno sequence and the anti-Shine–Dalgarno sequence provides a mechanism for distinguishing a start codon from an internal methionine codon. It also increases the stability of the ribosome-mRNA complex.

Each gene on a polycistronic mRNA possesses its own Shine–Dalgarno sequence and an initiation codon. The translation of each gene appears to occur independently; that is, translation of the first gene in a polycistronic message may or may not be followed by the translation of subsequent genes.

In the next step in initiation, IF2 (a GTPase with a bound GTP) binds to the initiating tRNA and facilitates its entry into the P site. The initiating tRNA in bacteria is a *N*-formylmethionine-tRNA (fmet-$tRNA^{fmet}$), which is synthesized in the following process. After an initiator tRNA ($tRNA^{fmet}$) is charged with methionine, the amino acid residue is formylated by an N^{10}-formyl THF–requiring enzyme. The initiation phase ends as IF2-GTP promotes the joining of the two subunits via its binding to the GAR site of the 50S subunit. The subsequent GAR-initiated GTP hydrolysis causes a conformational change that results in the simultaneous joining of the two subunits and the release of the three initiation factors, GDP, and P_i. The 70S ribosome is now primed for the elongation phase of protein synthesis.


IF3
30S
subunit
IF3
mRNA
IF3
5′
A
U
G
3′
IF1
IF2
GTP
fMet
UAC
fMet
IF2
GTP
IF3
UAC
IF1
5′
A
U
G
3′
50S
subunit
IF2
+
GDP
+
P_i
+
IF1
IF3
fMet

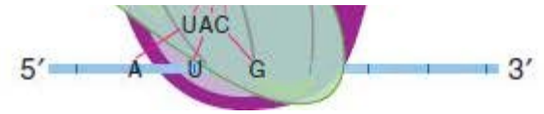


FIGURE 19.7

Formation of the Prokaryotic Initiation Complex

The initiation phase of translation begins with the binding of the initiation factor IF3 to the 30S subunit. After the mRNA binds to the 30S subunit, the GTPase IF2 binds to fmet-tRNA^{fmet} and subsequently promotes the association of the small and large subunits. GTP hydrolysis triggers the release of the initiation factors, GDP, and P_i. The fully functional 70S ribosome is now ready to enter the elongation phase.

ELONGATION As noted, elongation consists of three steps, referred to collectively as an elongation cycle: (1) positioning an aminoacyl-tRNA in the A site, (2) peptide bond formation, and (3) translocation.

The prokaryotic elongation process begins when an aminoacyl-tRNA, specified by the next codon, binds to the now-empty A site. Before aminoacyl-tRNAs can enter the A site, they must bind the elongation factor EF-Tu-GTP. EF-Tu-GTP is a motor protein that positions its cargo within the A site so that the tRNA anticodon is free to interact with an mRNA codon. EF-Tu-GTP prevents unregulated peptide bond formation and serves to protect the aminoacyl linkage to the tRNA from hydrolysis. Entry of the aminoacyl-tRNA into the A site requires the binding of EF-Tu-GTP to yield the EF-Tu-GTP-aminoacyl-tRNA ternary complex that then binds to the 50S subunit GAR, a process that is assisted by the subunit's L12 stalk.

After the anticodon of the aminoacyl-tRNA has been correctly paired to the mRNA codon, GTP hydrolysis releases EF-Tu-GDP from the ribosome. Then another elongation factor (called EF-Ts), acting as a GEF (p. 606), promotes EF-Tu regeneration by displacing its GDP moiety. EF-Ts is then itself displaced by an incoming GTP molecule (**Figure 19.8**). The newly formed EF-Tu-GTP can then bind to a new aminoacyl-tRNA. (The structural and functional properties of EF-Tu are described online in Biochemistry in Perspective: EF-Tu: A Motor Protein.)

After EF-Tu has delivered an aminoacyl-tRNA to the A site, the formation of a peptide bond is catalyzed by the PTC within the 23S rRNA that is located in a cleft in the 50S ribosomal subunit on the side facing the 30S subunit. The mechanism whereby the peptide bond is formed occurs via intrasubstrate proton shuttling (**Figure 19.9**). The ribosome facilitates the reaction in several ways. These include precise positioning of the substrates within the PTC (i.e., the acceptor ends of the A-site and P-site tRNAs become fixed in place via interactions with nucleotide residues of 23S rRNA) and an electrostatic environment that assists the proton shuttle process. The ribosome also reduces the free energy requirements of the reaction by providing a relatively anhydrous environment in the active site that is essential for formation of a highly polar transition state. The energy required to drive this reaction is provided by the high-energy ester bond linking the P-site amino acid to its tRNA.

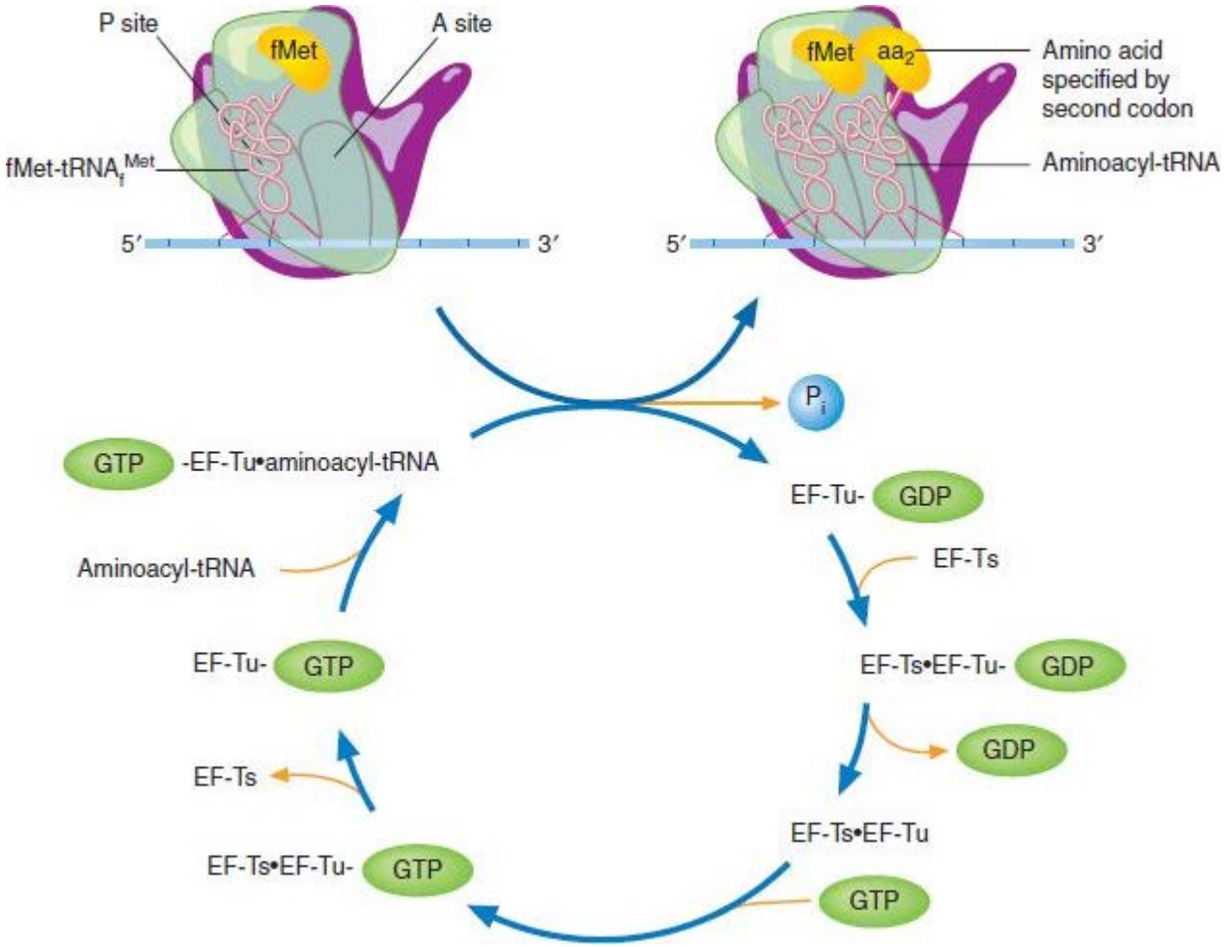


FIGURE 19.8

The EF-Tu-EF-Ts Cycle in *E. coli*

Before EF-Tu can bind an aminoacyl-tRNA, its GDP moiety must be replaced by GTP. The binding of EF-Ts to EF-Tu (GDP) displaces GDP. Then EF-Ts is itself displaced by an incoming GTP. Then EF-Tu (GTP) associates with an aminoacyl-tRNA to form an EF-Tu (GTP) aminoacyl-tRNA complex, which proceeds to deliver the aminoacyl-tRNA to the A site of the ribosome.

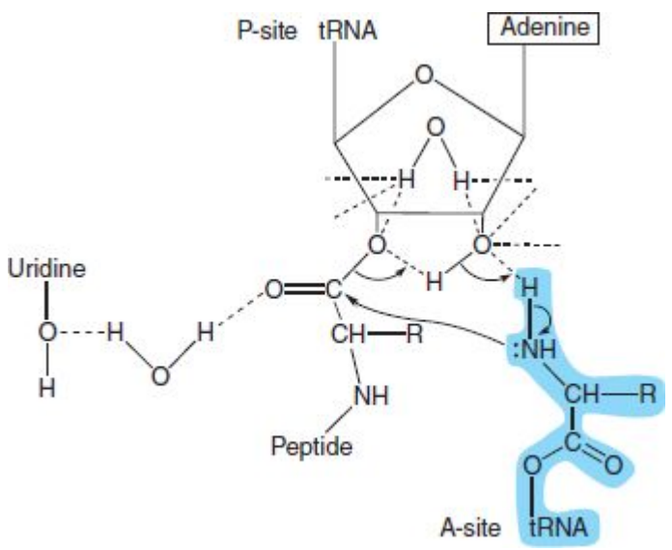


FIGURE 19.9

The Peptidyl Transferase Proton Shuttle Mechanism

The reaction begins with the nucleophilic attack of the α-amino nitrogen on the A-site aminoacyl group

carbonyl carbon of the peptidyl chain (linked to P-site tRNA by an ester bond to 3′-OH group of the ribose of residue 76). During the first elongation cycle, a single *N*-formylmethionyl amino acyl group is linked to the P-site tRNA. A six-membered transition state is thus created in which the 2′-OH group of the A-site ribose donates its proton to the adjacent 3′-oxygen atom as the latter receives an amino proton. The precise alignment of the substrates within the active site is fostered by hydrogen bonds between the ribose oxygens and a bridging water molecule (black dashed lines) and hydrogen bonds between the ribose 2′-hydroxyl oxygen and the bridging water molecule and certain cytosine and adenosine residues (not shown) of 23S rRNA (dashed lines). The carbonyl oxygen of the peptidyl group of the P-site tRNA is stabilized by a hydrogen bond, via a water molecule bridge, to the hydroxyl oxygen of a uridine residue of 23S rRNA.

Immediately after peptide bond formation, the tRNA in the A site (referred to as the peptidyl tRNA because it is linked to the growing peptide chain) is still in the A site, and the now-deacylated tRNA is in the P site. This phase of elongation is referred to as the *pretranslocation state*. Translocation, the shifting of the base-paired tRNAs by one codon position (putting the uncharged tRNA in the E site, the peptidyl-tRNA in the P site, and the new codon in the empty A site), requires the interaction of another GTPase, referred to as EF-G, with a loop of the 23S rRNA in the large subunit. When EF-G-GTP binds near the A site assisted by L12 stalk motion, GTP hydrolysis is triggered. The resulting conformational change causes the two subunits to rotate in opposite directions in relation to each other. The rotational movement, which creates a space between the two subunits, accommodates the ratchet-like shift of the tRNAs along the mRNA strand. Note that recent research has revealed that translocation is an intrinsic feature of ribosomes. The movement of the mRNA/A site tRNA within the ribosome is apparently accelerated approximately 50,000-fold by the binding of EF-G-GTP to the ribosome. GTP hydrolysis primarily causes EF-G release.

In the *posttranslocation state*, the deacylated tRNA is in the E site. The release of the deacylated tRNA from the E site occurs as the incoming EF-Tu aminoacyl-tRNA complex enters the A site and codon–anticodon interactions are initiated. After the release of EF-G-GDP, the ribosome is ready for the next elongation cycle. As the polypeptide elongates, it passes through an irregularly shaped 10- to 20-Å-wide and 100-Å-long exit tunnel in the 50S subunit. Each **nascent** (newly synthesized) polypeptide emerges from the tunnel when it is between 30 and 50 amino acid residues in length, depending on the polypeptide's propensities for forming secondary structures (α-helices) in the 20-Å-wide vestibule near the subunit's surface. Elongation continues until a stop codon enters the A site.

TERMINATION The termination phase begins when a termination codon (UAA, UAG, or UGA) enters the A site. Three **releasing factors** (RF1, RF2, and RF3) are involved in termination. Both RF1 and RF2 resemble tRNAs in shape and size. RF1 recognizes the stop codons UAA and UAG, and RF2 recognizes UAA and UGA. RF3 is a GTPase that promotes the binding of RF1 and RF2 to the ribosome. The binding of RF3-GDP to the RF1- or RF2-ribosome complex induces RF3 to exchange GDP for GTP. RF binding alters ribosomal function. Hydrolysis of the GTP-bound to RF3 triggers the release of RF1 or RF2. Releasing factor binding causes a change in the orientation of the decoding center bases (A1492, A1493, and G530), which triggers a conformational change within the PTC that transiently transforms peptidyl transferase into an esterase. The subsequent hydrolysis reaction cleaves the bond linking the completed polypeptide and the P-site tRNA. Following the polypeptide's release from the ribosome, the mRNA and tRNA also dissociate. The termination phase ends with the polypeptide's release from the ribosome, followed by release of mRNA and tRNA and the separation of the ribosomal subunits. The latter process requires **ribosome recycling factor**, a tRNA-shaped protein that binds to the A site. The energy required for termination is supplied by EF-G-GTP.

POSTTRANSLATIONAL MODIFICATIONS The folding process begins as each nascent polypeptide emerges from the exit tunnel, where it first encounters a molecular chaperone called trigger factor. *Trigger factor* (TF), a 48 kDa protein, is transiently linked to the ribosome via an attachment between its N-terminal domain and ribosomal protein L23. Trigger factor's C-terminal domain, an elongated, narrow, and flexible flap-like structure, is positioned at the ribosome exit site. TF provides a specialized surface that guides the early steps of the folding process. Downstream chaperones (e.g., DNAK [hsp70] and GroES-GroEL, p. 172) further assist the folding process if needed. As mentioned, most polypeptides also undergo a series of modifying reactions that prepare them for their functional role. Most of the information concerning posttranslational modifications has been obtained through research on eukaryotes. However, prokaryotic polypeptides are known to undergo several types of covalent modifications.

The best-researched examples are proteolytic processing reactions. These include removing the formylmethionine residue and signal peptide sequences. **Signal peptides**, or leader peptides, are short peptide sequences, typically near the amino terminal, that determine a polypeptide's destination. In bacteria, for example, a signal peptide is required to insert a polypeptide into the plasma membrane.

Posttranslational chemical modifications of prokaryotic proteins include methylation, phosphorylation, glycosylation, and lipidation, (covalent linkage to lipid molecules). In *E. coli*, chemotaxis is regulated by methylation and phosphorylation of signal transduction proteins. (*Chemotaxis* is the process in which cells alter their movements in response to chemical stimulus such as nutrient molecules in their environment.)

Lipoproteins are fairly common in prokaryotes. B1c, a lipoprotein found in the outer membrane of *E. coli*, is a type of lipocalin (a protein that binds hydrophobic ligands) that is produced under stressful conditions such as starvation. B1c forms covalent linkages with fatty acids and phospholipids that play a role in membrane biogenesis and repair.

TRANSLATIONAL CONTROL MECHANISMS Protein synthesis is an exceptionally expensive process costing four high-energy phosphate bonds per peptide bond (i.e., two bonds expended during tRNA charging and one each during A-site–tRNA binding and translocation). It is perhaps not surprising that enormous quantities of energy are involved. For example, approximately 90% of *E. coli* energy production used in the synthesis of macromolecules may be devoted to the manufacture of proteins. Although the speed and accuracy of translation require high-energy input, the cost would be even higher without metabolic control mechanisms. These mechanisms allow prokaryotic cells to compete with each other for limited nutritional resources.

In prokaryotes such as *E. coli*, most protein synthesis control processes occur at the level of transcription initiation. (Refer to Section 18.3 for a discussion of the principles of prokaryotic transcriptional control.) This circumstance makes sense for several reasons. First, transcription and translation are spatially and temporally coupled; that is, translation is initiated shortly after transcription begins (**Figure 19.10**). Second, the lifetime of prokaryotic mRNA is usually relatively short. With half-lives of between 1 and 3 minutes, the types of mRNA produced in a cell can be quickly altered as environmental conditions change. Most mRNA molecules in *E. coli* are degraded by two exonucleases, referred to as RNase II and polynucleotide phosphorylase.

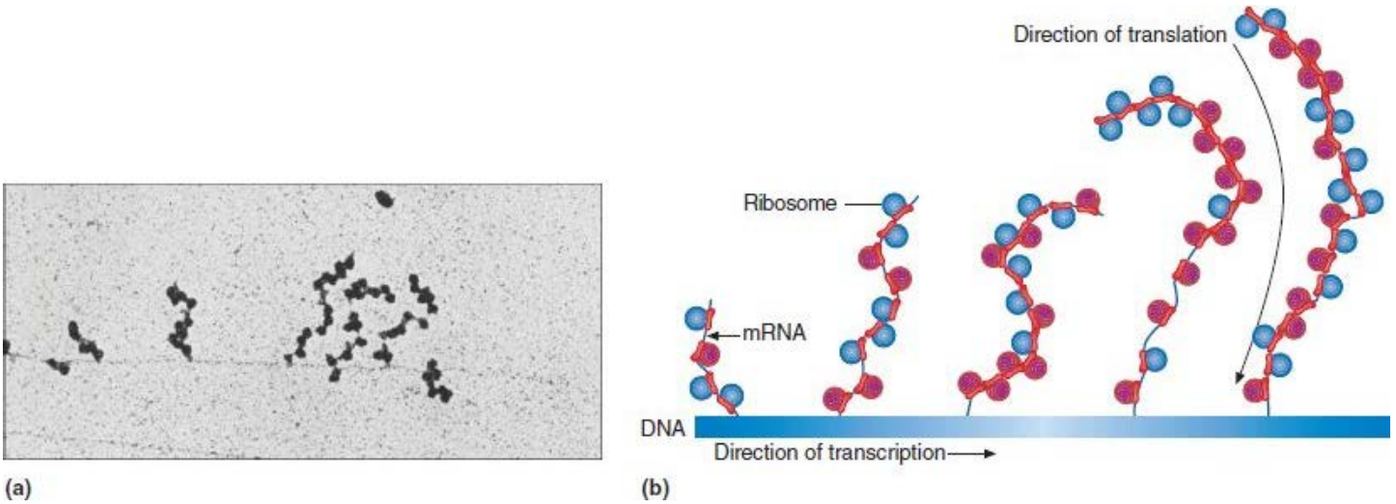


FIGURE 19.10

Transcription and Translation in *E. coli*

(a) An electron micrograph of *E. coli* transcription and translation. In *E. coli*, as in other prokaryotes, transcription and translation are directly coupled. (b) Diagram of (a). Note polyribosomes.

The rates of prokaryotic mRNA translation are variable, owing in part to differences in Shine–Dalgarno sequences. Because Shine–Dalgarno sequences facilitate the selection of the initiation codon, sequence variations may affect the translation rate. For example, the gene products of the *lac* operon (*β*-galactosidase, galactose permease, and thiogalactoside transacetylase) are not produced in equal quantities. Thiogalactoside transacetylase is produced at approximately one-fifth the rate of *β*-galactosidase.

KEY CONCEPTS

- Prokaryotic protein synthesis is a rapid process involving several protein factors.
- Although most prokaryotic gene expression appears to be regulated by transcription initiation, several types of translational regulation have been detected.

The structural and functional differences between prokaryotic and eukaryotic protein synthesis are the basis of the therapeutic and research uses of antibiotics, antimicrobial molecules used to treat infections. The actions of several antibiotics are listed in **Table 19.2**.

TABLE 19.2 Select Antibiotic Inhibitors of Protein Synthesis

Antibiotic	Action
Chloramphenicol	Blocks prokaryotic A site
Cycloheximide	Inhibits eukaryotic peptidyl transferase
Erythromycin	Blocks prokaryotic exit site
Lincosamide	Binding to 23S rRNA of the 50S subunit
Streptomycin	Blocks binding of fmet-$tRNA_i$ to 30S P site
Streptogramins	Premature release of polypeptide chain

Tigecycline	Binding to 16S rRNA prevents entry of aa-tRNA to A site

Eukaryotic Protein Synthesis

Although eukaryotic protein synthesis resembles bacterial translation, there are distinct differences in the two processes. At 4.3 MDa, the eukaryotic ribosome, also known as the 80S ribosome, is larger than the bacterial version. The large 60S ribosomal subunit is composed of 28S, 5S, and 5.8S rRNAs and 47 proteins. The small 40S subunit is composed of an 18S rRNA and 32 proteins. In addition, a significantly larger number of protein factors assist a vastly more sophisticated translation process. The posttranslational modifications of eukaryotic polypeptides are remarkably more numerous than those observed in prokaryotes. Eukaryotic polypeptide targeting mechanisms are also quite intricate.

INITIATION Many of the major differences between the prokaryotic and eukaryotic versions of protein synthesis can be observed during the initiation phase. Among the reasons for the additional complexity of eukaryotic initiation are the following.

1. **mRNA secondary structure**. Eukaryotic mRNA is processed with the addition of a methylguanosine cap (**Figure 18.45**) and a poly(A) tail and the removal of introns. In addition, eukaryotic mRNA does not associate with a ribosome until it has left the nucleus complexed with several proteins.
2. **mRNA scanning**. In contrast to prokaryotic mRNA, eukaryotic molecules lack Shine–Dalgarno sequences, which allow for the identification of the initiating AUG sequence. Instead, eukaryotic ribosomes "scan" mRNAs for the AUG start codon embedded in the *Kozak sequence* (named for Marilyn Kozak) A*CCA*CCAUGA*, where the asterisk indicates each of the adenine bases that can be replaced with a guanine.

Eukaryotes use a more complex spectrum of initiation factors than prokaryotes. There are at least 12 eukaryotic initiating factors (eIFs), several of which possess numerous subunits. The functional roles of most of these factors are still under investigation.

Eukaryotic initiation begins with the assembly of the preinitiation complex (PIC) (**Figure 19.11**). The **43S preinitiation complex** is composed of the 40S subunit, met-$tRNA_i^{met}$, and several initiation factors: eIF1, eIF1A, eIF2-GTP, eIF3, and eIF5. The small (40S) subunit is momentarily prevented from binding to the large (60S) subunit because it is associated to eIF3 and the large subunit is bound to eIF6. PIC construction begins when eIF1 binds to the small subunit near the P site, and eIF1A binds to and blocks the A site during initiation. eIF3 is a multisubunit protein complex that facilitates the subsequent mRNA binding process and prevents premature small subunit binding to the large subunit. eIF5 is a GAP that is specific for the GTPase eIF2 that binds to met-$tRNA_i^{met}$ to yield the eIF2-met-$tRNA_i^{met}$ ternary complex. With the ternary complex bound in the P site, the PIC is complete and ready to bind mRNA.

Initiation proceeds as the 5'-methylated m^7G cap of the mRNA binds to a **cap-binding complex** (CBC). There are two types of CBCs: CBC20/80 (in metazoans) or Cbc1/Cbc2 (in yeast) and eIF4F. When an mRNP leaves the nucleus, its 5'-methylated (m^7G) cap is bound to a nuclear cap-binding complex (CBC20/80). During the first round of translation, called the *pioneer round*, the CBC20/80 cap facilitates surveillance of the mRNA's exon junction complexes (p. 729) in conjunction with nonsense-mediated decay (NMD) proteins (p. 729). As the ribosome encounters an exon junction complex (EJC, p. 729), it removes it and proceeds to translate the mRNA unless

a premature STOP codon (a nonsense mutation) is encountered ahead of an EJC. Once this event occurs, NMD proceeds to transfer the mRNA to P bodies (p. 730) where it is degraded. In the absence of premature STOP codons, the 5’ mRNA cap binds the eIF4F CBC, and ribosomes proceed with the mRNA’s subsequent translations.

The eIF4F CBC consists of eIF4E (cap-binding protein), eIF4A (an RNA helicase), and eIF4G (a scaffold protein). eIF4E is the rate-limiting factor in eukaryotic protein synthesis. Some cellular proteins and viruses can evade the necessity of using eIF4E, a protein regulated by mTORC1 (p. 612) and MAPK (pp. 741–42) signal transduction pathways. The mRNAs for proteins such as the heat shock proteins (hsps) and amyloid precursor protein and viruses such as poliovirus possess internal ribosome entry site sequences (IRES) that allow eIF4E-independent initiation.

The binding of eIF4F CBC to the mRNA cap involves an ATP-dependent RNA helicase (eIF4A), assisted by eIF4B, which removes any secondary structures in the 5′-UTR that can interfere with the mRNA scanning process. The 3′-poly(A) tail of the mRNA is then brought into close proximity to the 5′capped end by interactions between eIF4G and multiple copies of **poly(A)-binding protein (PABP)** to form a circular mRNA. The completed **48S initiation complex** proceeds to scan the mRNA in search of the 5′-AUG-3′ near the 5′ end (**Figure 19.12**). Both eIF1 and eIF1A assist in the scanning process. When the initiation codon is reached, a change in the conformation of the scanning complex causes it to lock on to the mRNA. The conformation change also induces eIF5, a GAP, to mediate the hydrolysis of the GTP bound to eIF2. Once the GTP is hydrolyzed, eIF2-GTP is regenerated by eIF2B, a GEF. The subsequent binding of eIF5B, a ribosome-dependent GTPase homologous to IF2, promotes the joining of the 60S subunit (now devoid of eIF6) to the 48S initiation complex and displacement of eIF2, eIF3, and eIF5. Hydrolysis of the GTP bound to eIF5B allows the dissociation of eIF5B-GDP and eIF1A from the active 80S ribosome.

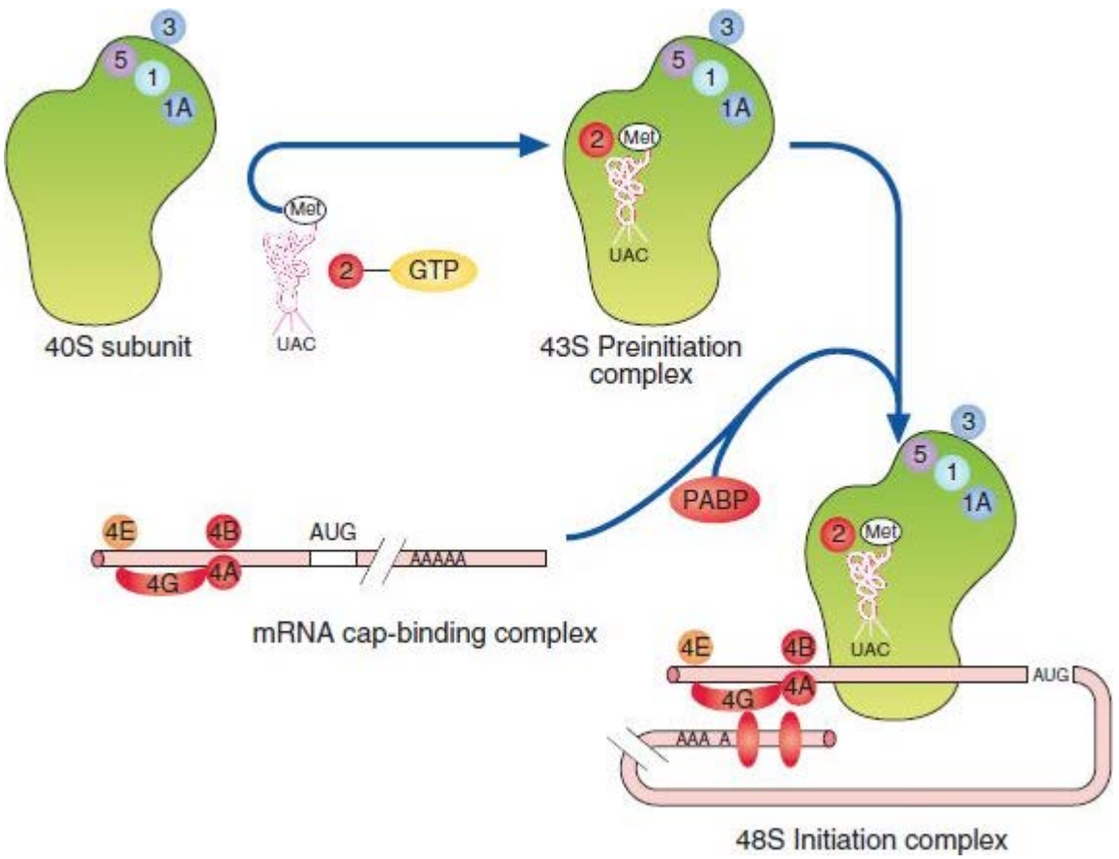


FIGURE 19.11

Eukaryotic Initiation: Assembly of the 43S Preinitiation Complex and the 48S Initiation Complex

In preparation for its role in polypeptide synthesis, the 40S ribosomal subunit, previously bound to eIF1, eIF1A, eIF3, and eIF5, now binds a complex of eIF2-GTP and the initiation tRNA, met-$tRNA_i$. Once the 40S subunit binds to the met-$tRNA_i$-eIF2-GTP complex, it is referred to as the 43S preinitiation complex. Once eIF4B, a helicase, has removed secondary structure from the 5′-UTR of the mRNA, the cap-binding complex, composed of eIF4A, eIF4B, eIF4E, and eIF4G, binds to the 5′-cap structure. The 48S initiation complex forms as the 3′-poly(A) tail of the mRNA is linked to the 5’-capped end by interactions between eIF4G and multiple copies of PABP. Note that the eukaryotic initiation process involves at least 30 proteins. Only the most important are mentioned in the text and illustrated in this figure.

Trapped Ribosomes: RNA to the Rescue!

How are ribosomes bound to damaged mRNAs retrieved so that they can be recycled? What happens when a ribosome attempts to translate an mRNA that is either truncated or missing a stop codon? The answer is that translation stalls and ribosomes bound to such defective mRNAs become trapped. They cannot move the process forward, nor can they eject the mRNA. Because of the critical importance of efficient protein synthesis, bacteria solve this seemingly intractable problem with a unique 10S RNA molecule called tmRNA. tmRNA is so named because it possesses a tRNA-like domain (TLD) and an mRNA-like domain (MLD). The TLD, formed by the 5′ and 3′ ends of tmRNA, mimics the acceptor stem, T-stem, variable stem, and 3′CCA end of an alanyl-tRNA. The MLD contains an open reading frame (ORF) that specifies the amino acid sequence (about 10 residues long) of the peptide tag that the ribosome will eventually add to the stalled polypeptide.

Ribosome rescue (**Figure 19A**) begins with charging of the TLD with an alanine, catalyzed by alanyl-tRNA synthetase. The alanyl-charged TLD then binds to EF-Tu and SmpB, a small protein that mimics the anticodon arm of a tRNA. The TLD of the tmRNA-protein complex then binds at the A site, thus positioning its bound alanine within the PTC. The alanyl residue is then covalently bound to the nascent polypeptide. The nearby MLD is poised for the ribosome to switch from the damaged mRNA to its ORF base sequence. Although this transition is structurally awkward, translation soon resumes and continues until the MLD stop codon (UAA) is reached. Soon after its release from the ribosome, the damaged mRNA is degraded in a process that is facilitated by tmRNA. The newly released polypeptide is also degraded, having been targeted for destruction by protease binding sites within the newly synthesized peptide tag.

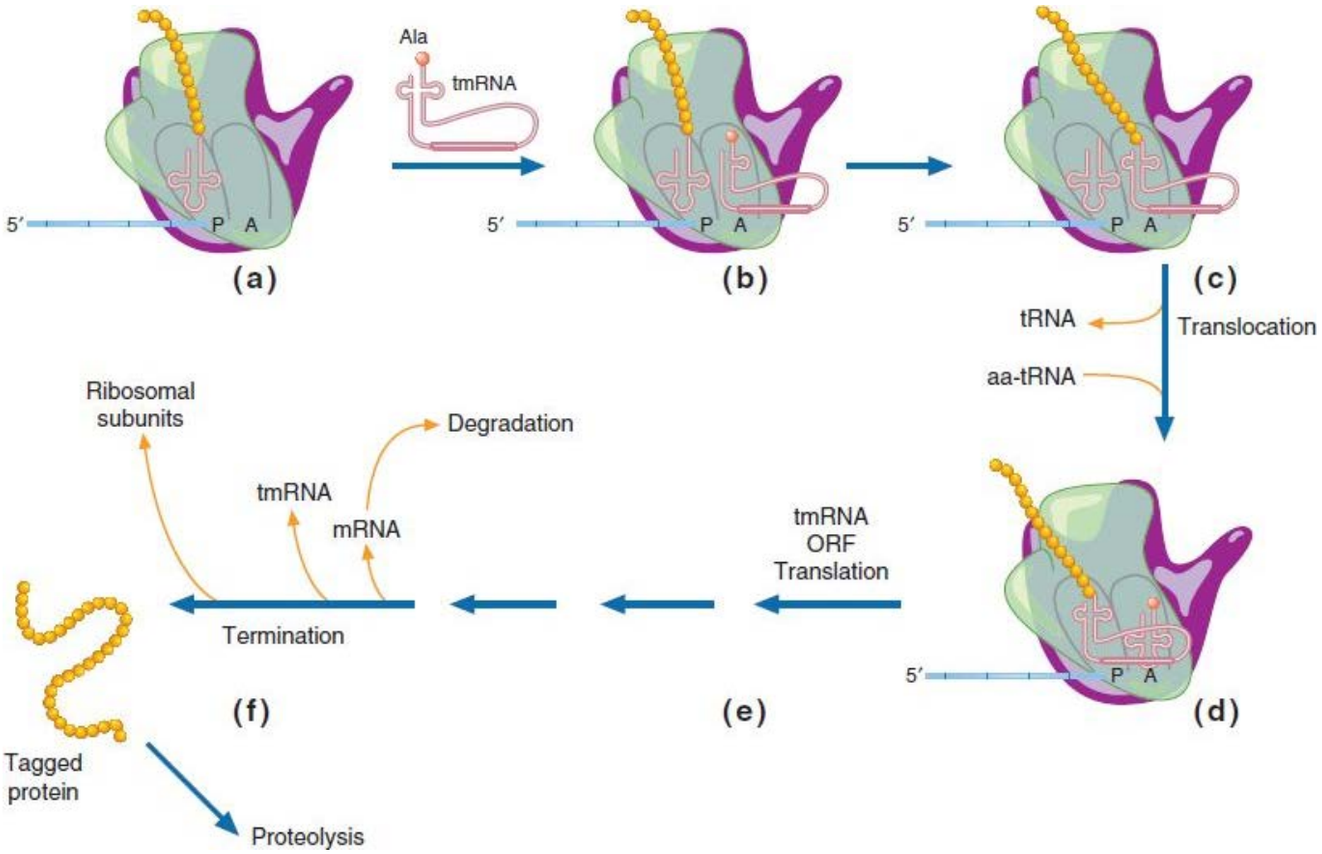


FIGURE 19A

Trapped Ribosome Rescue

(a) A tmRNA linked to an alanine residue binds via codon–anticodon interactions in the A site of the stalled ribosome. (b) The ribosome-catalyzed addition of the alanine to the C-terminus of the nascent polypeptide is followed by translocation (c). (d) The MLD of the tmRNA moves into position so that the first codon in its ORF is in the A site. (e) The MLD ORF is then translated by the ribosome until the stop codon is reached. (f) Releasing factor-mediated termination releases the tmRNA and mRNA and causes ribosomal subunit separation. The defective mRNA and the polypeptide are then degraded.

Eukaryotes do not have a ribosome rescue operation comparable to the tm-tRNA system. When eukaryotic mRNAs lack a stop codon, the ribosome translates the poly(A) tail, one AAA codon at a time, until the 3′ end is reached, thereby producing a C-terminal polylysine segment. An mRNA surveillance system called *nonstop-mediated mRNA decay* detects and rescues the stalled ribosome. A protein called Ski7 binds to the vacant A site, thereby causing the ribosome to release the mRNA and the polypeptide. The defective mRNA is then destroyed. The polypeptide is degraded because the polylysine peptide tag on its C-terminal end is a target for proteases.

SUMMARY Living organisms, faced with the high metabolic cost of protein synthesis, have evolved the means of ensuring the efficiency of the process. Both prokaryotes and eukaryotes have methods for rescuing ribosomes trapped by association with damaged mRNAs.

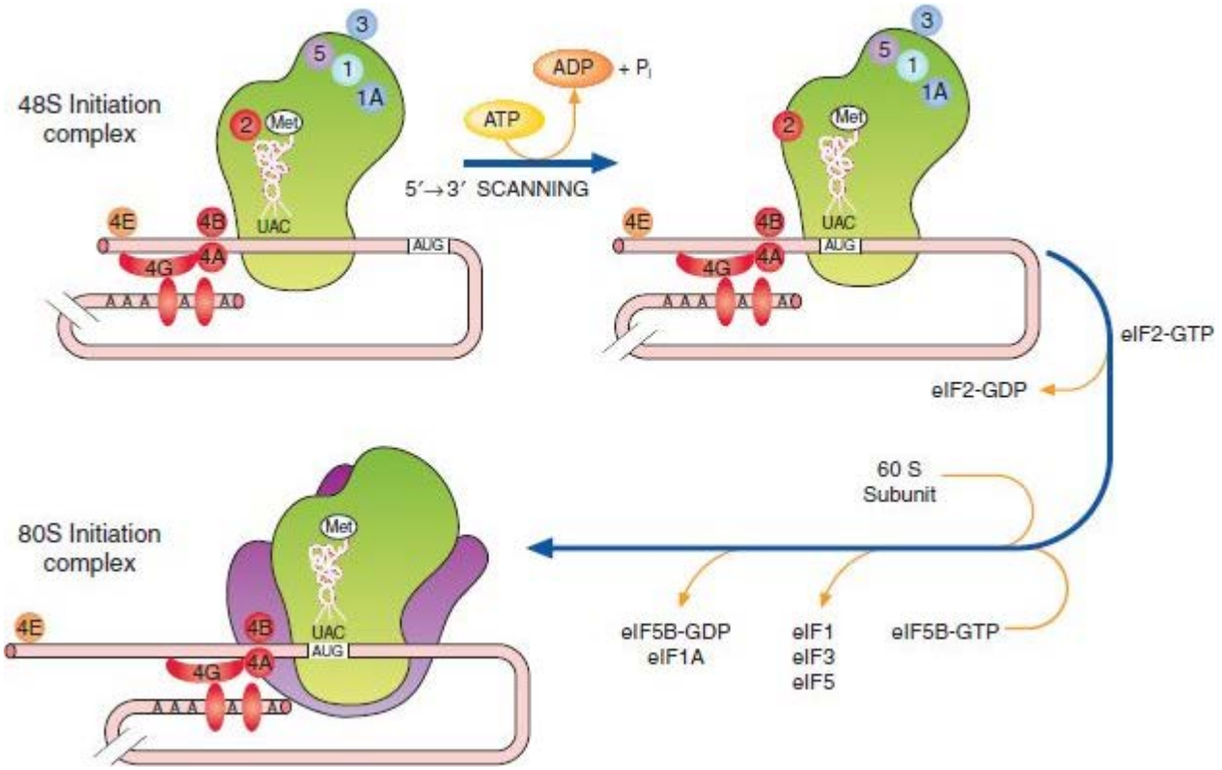


FIGURE 19.12

Eukaryotic Initiation: Assembly of the 80S Initiation Complex

The newly formed 48S complex moves in an ATP-requiring scanning process along the mRNA in the 5′ → 3′ direction in search for the start codon. Once correct base pairing between the AUG codon and the anticodon of met-$tRNA_i$ has occurred, eIF5 triggers GTP hydrolysis and the release of eIF2-GDP. The subsequent binding of eIF5B-GTP to the CBC-mRNA complex facilitates the joining of the complex with the 60S subunit, an event that involves the displacement of eIF1, eIF3, and eIF5. Hydrolysis of the GTP linked to eIF5B causes the dissociation of eIF5B-GDP and eIF1A from the elongation competent 80S ribosome.

ELONGATION **Figure 19.13** illustrates the eukaryotic elongation cycle. Several elongation factors (eEFs) are required during this phase of translation. *eEF1α*, a 50-kDa polypeptide, is the eukaryotic equivalent of EF-Tu; that is, it is a GTPase that binds aminoacyl-tRNAs and delivers them to the A site. If correct codon–anticodon pairing occurs, eEF1*α* hydrolyzes its bound GTP and exits the ribosome, leaving its aminoacyl-tRNA behind. If correct pairing does not occur, the complex leaves the A site, thereby preventing incorrect amino acid residues from being incorporated.

During the next elongation step (peptide bond formation), the peptidyl transferase activity of the large ribosomal subunit catalyzes the nucleophilic attack of the A-site *α*-amino group on the carboxyl carbon of the P-site amino acid residue. eEF1*α*-GDP dissociates from the ribosome immediately before transpeptidation. eEF1*β* and eEF1*γ* are GEFs that mediate the regeneration of eEF-1*α*-GTP by promoting an exchange of GDP for GTP.

Translocation in eukaryotic ribosomes requires a 100-kDa polypeptide referred to as *eEF2*, which is also a GTPase. During translocation, eEF2-GTP binds to the ribosome. GTP is then hydrolyzed to GDP, and eEF2-GDP is released. As noted, GTP hydrolysis provides the energy needed to physically move the ribosome along the mRNA. At the end of translocation, a new codon is exposed in the A site. As the polypeptide grows, it passes through an exit tunnel with the same dimensions as the bacterial tunnel. The exit tunnel's size excludes large domain folding but can allow *α*-helix formation.

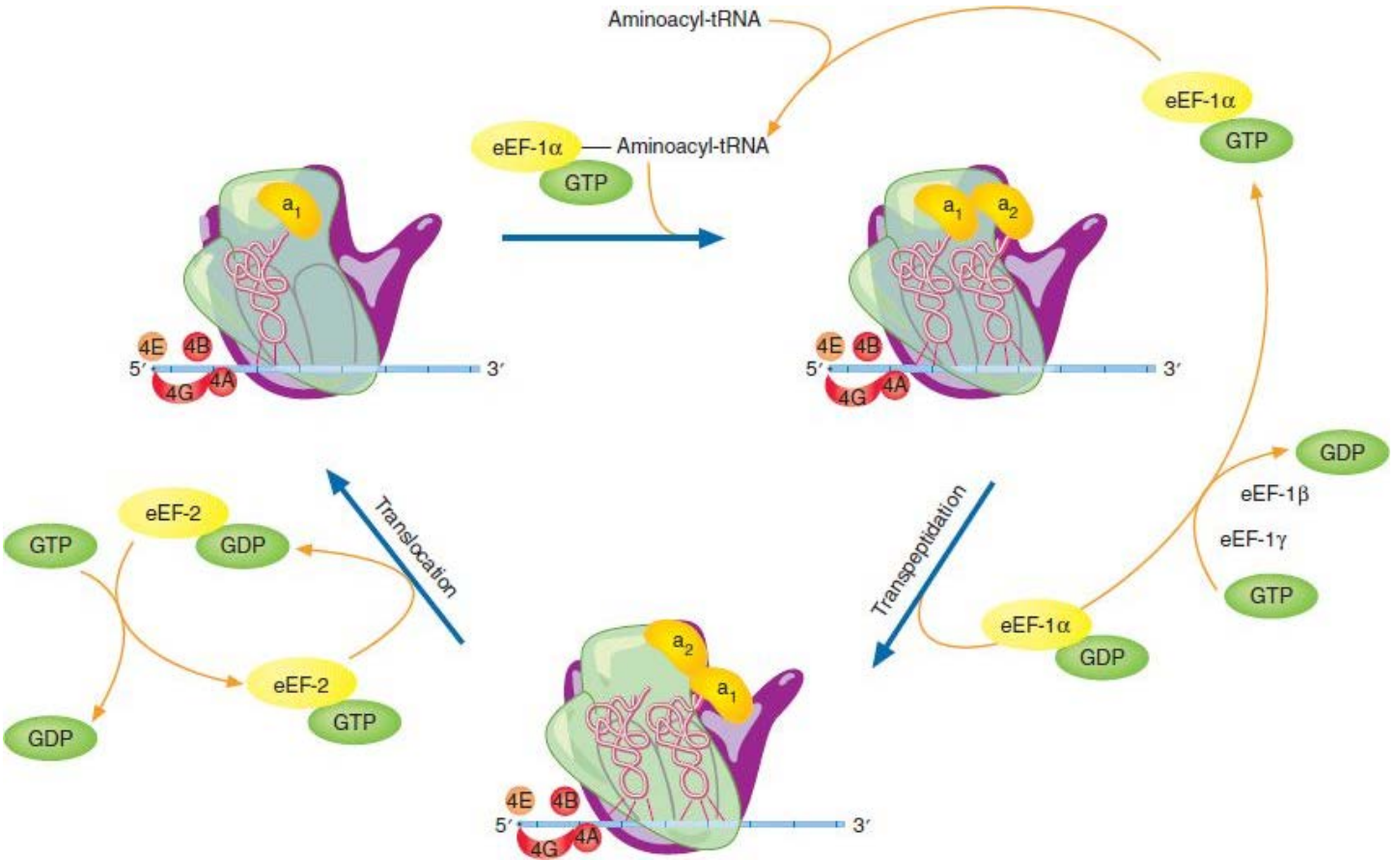


FIGURE 19.13

The Elongation Cycle in Eukaryotic Translation

Elongation comprises three phases: binding of an aminoacyl-tRNA to the A site, transpeptidation, and translocation.

TERMINATION In eukaryotic cells, two releasing factors mediate the termination process: eRF1 (a molecule that resembles the size and overall shape of a tRNA and recognizes and binds to stop codons) and eRF3 (a GTPase). When a stop codon (UAG, UGA, or UAA) enters the A-site, eRF1 binds to it (**Figure 19.14**). As a result of this binding process, peptidyl transferase catalyzes the hydrolysis of the ester linkage between the polypeptide and the P-site tRNA. The hydrolysis of the GTP bound to eRF3 is believed to then trigger the dissociation of eRF1 from the ribosome. In eukaryotes, the dissociation of ribosomal subunits has also been attributed to eIF3, eIF1, and eIF1A.

The efficiency of eukaryotic translation (i.e., the number of polypeptides that can be synthesized per unit of time) is made possible to a large extent by the circular conformation of eukaryotic polysomes (**Figure 19.15**), As soon as ribosomal subunits and their associated protein factors are released, they are optimally positioned for recruitment into new ribosomes.

KEY CONCEPTS

- Eukaryotic protein synthesis, like its prokaryotic counterpart, has three phases: initiation, elongation, and termination.
- Features that are unique to eukaryotic translation include an abundance of protein factors that facilitate each step, cap-binding protein, and the formation of circular mRNA polysomes.

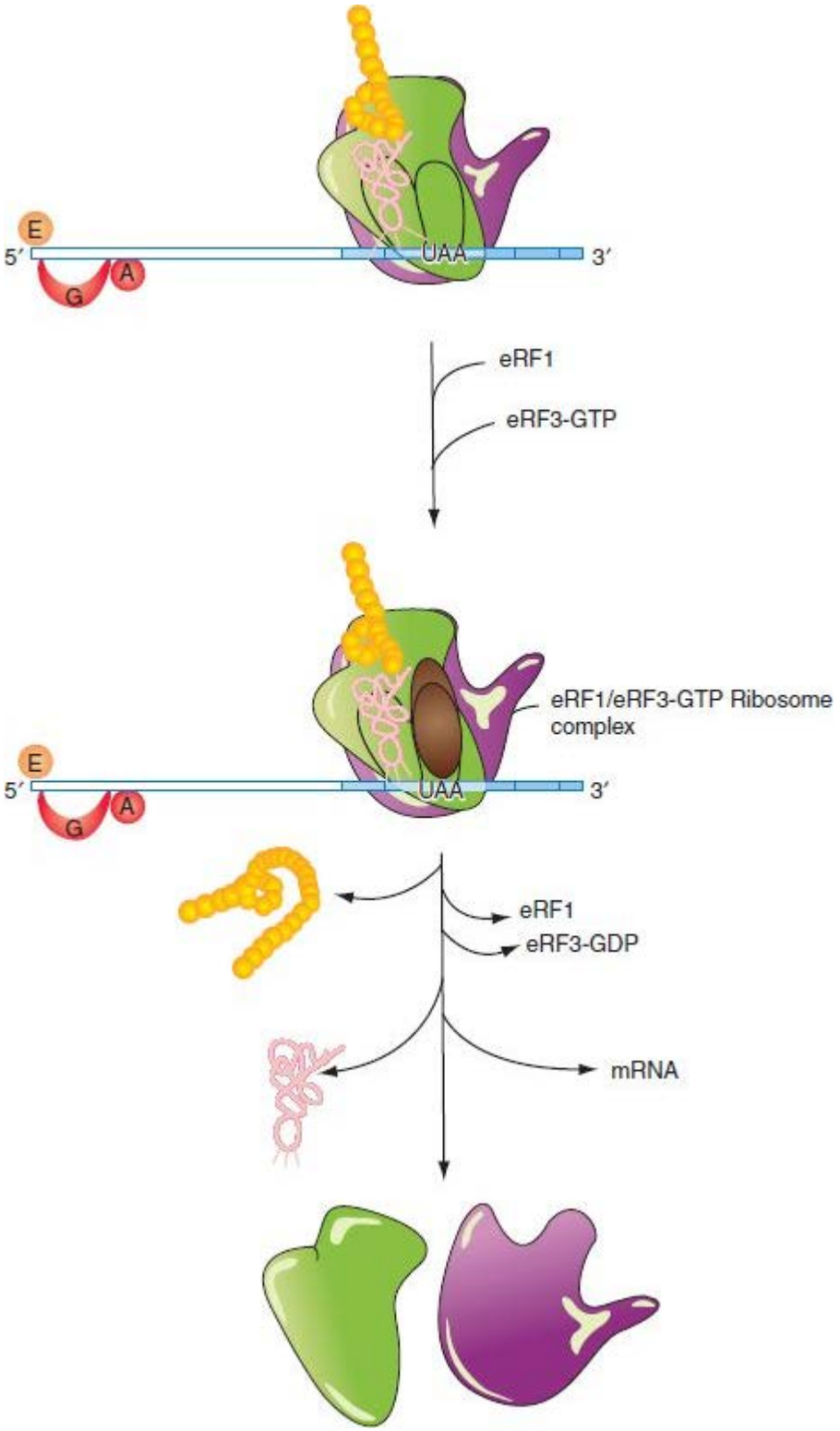


FIGURE 19.14

Eukaryotic Protein Synthesis Termination

As a stop codon (UAG, UGA, or UAA) enters the A site, releasing factor eRF1 recognizes the stop codon and binds to the A site. eRF1 in association with eRF3-GTP promotes the conversion of the peptidyl transferase to a hydrolase that catalyzes the hydrolysis of the ester bond linking the now complete polypeptide chain to the P-site tRNA. Release of the polypeptide is followed by the dissociation of eRF1 and eRF3-GDP from the ribosome. The dissociation of the ribosomal subunits is mediated principally by eIF3, eIF1, and eIF1A.

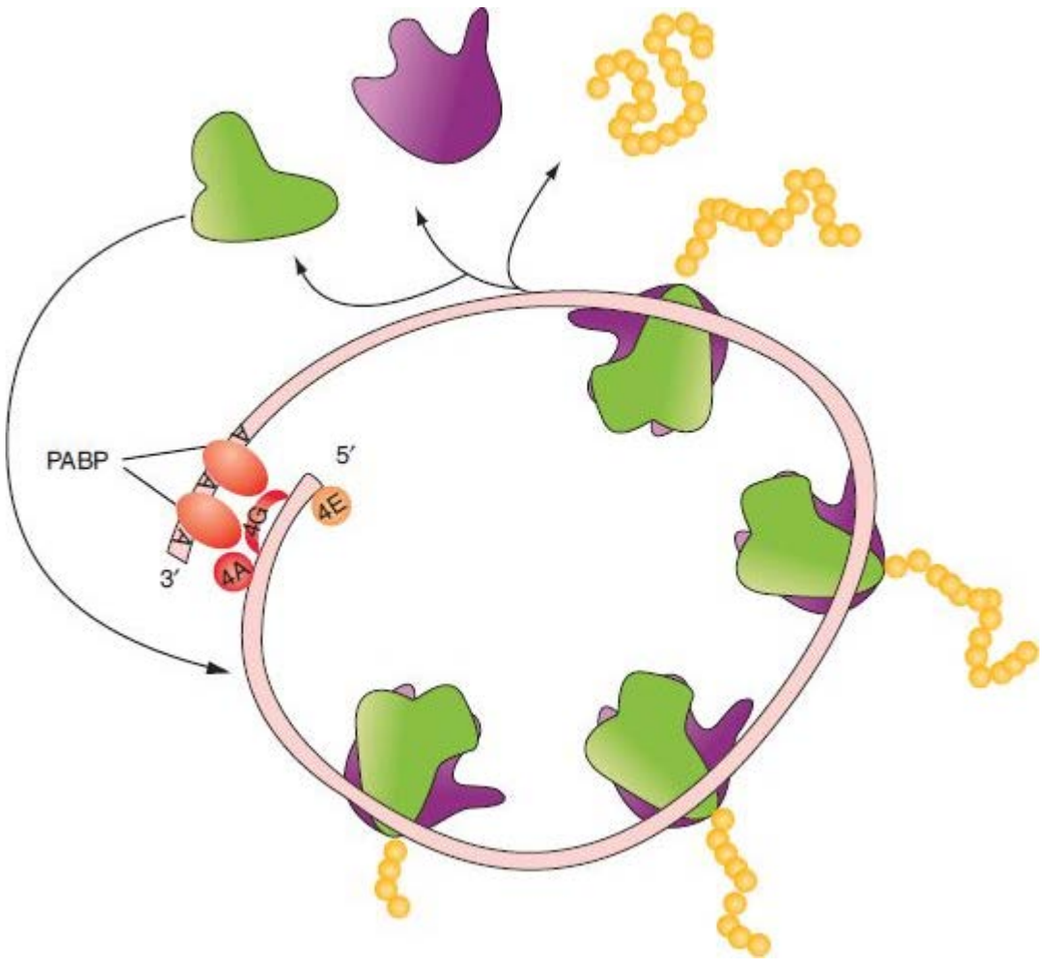


FIGURE 19.15

The Eukaryotic mRNA Polysome

Eukaryotic mRNA is circularized via an interaction between the 5′ UTR and the 3′-poly(A) tail that is mediated by PABP, the poly(A)-binding protein that links the poly(A) sequence to the CBC. As a result of this structural feature, when a polypeptide's synthesis is completed and the ribosomal subunits are released, the close proximity of these subunits to the 5′ cap facilitates immediate recruitment for another round of protein synthesis.

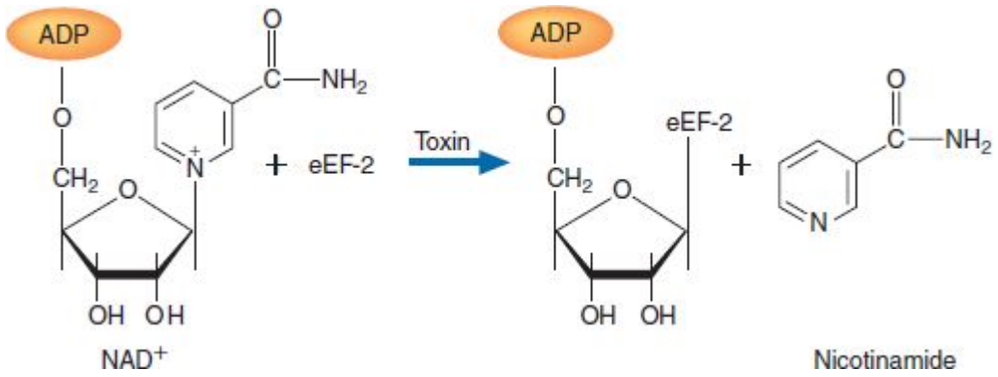
Diphtheria

QUESTION 19.4

Diphtheria is a highly contagious and potentially fatal respiratory tract illness. A throat lesion called a *pseudomembrane*, which is formed from bacterial cells and damaged throat epithelial cells, causes suffocation. Once considered a serious threat to children (e.g., the 1735–1740 epidemic in colonial New England killed a substantial proportion of children under the age of 16), diphtheria is now completely preventable because of a highly effective vaccine.

Diphtheria is caused by an exotoxin-producing pathogenic strain of the bacterium *Corynebacterium diphtheriae*. The exotoxin is a protein that is coded for by a genetic element introduced into the bacterial cell by a bacteriophage. After the diphtheria exotoxin is released

by the bacterium, it kills host cells by forming diphthamide, a specific ADP-ribosylated histidine residue in eEF2. Cells die because they cannot synthesize proteins. The mechanism by which eEF2 function is affected by ADP-ribosylation is unknown. Can you suggest any possibilities?

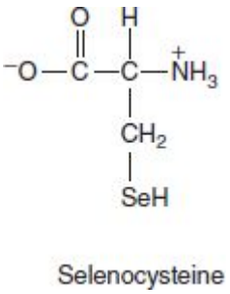


Biochemistry IN PERSPECTIVE

Context-Dependent Coding Reassignment

How is selenocysteine, a nonstandard amino acid, incorporated into polypeptides during protein synthesis? Selenium (Se), a trace mineral essential in small amounts, exerts its biological functions as the nonstandard amino acid selenocysteine (**Figure 19B**). Selenocysteine, now referred to as 21st amino acid, can be incorporated into polypeptides because of a mechanism called *context-dependent codon reassignment* that utilizes the stop codon UGA. The incorporation of selenocysteine into at least 25 selenoproteins (e.g., glutathione peroxidase and thioredoxin reductase) is widespread and has been observed in Eubacteria, Archaea, and Eukarya. The following discussion is limited to mammals.

Selenocysteine (sec) codon reassignment involves several molecules: a specific tRNA ($tRNA^{[ser]sec}$), a seryl-tRNA synthetase with sec-generating ability, an mRNA-binding protein (SBP2), and a specialized elongation factor (EFsec). The seryltRNA synthetase binds $tRNA^{sec}$, and serine is attached to its acceptor arm. The seryl moiety is then converted to selenocysteine by a pyridoxal phosphate–containing enzyme called selenocysteine synthase to form sec-$tRNA^{sec}$. The coding of selenocysteine requires a **SECIS (*sel*enocysteine *i*nsertion *s*equence) element** in the 3′ UTR of the mRNAs for all selenoproteins. The SECIS element rearranges into a stem-loop-bubble structure that recruits SBP2 to form a SECIS/SBP2 complex (**Figure 19C**). An AGU/AG sequence in the stem forms a nucleotide quartet that is the primary conserved sequence within the SECIS. One suggested function of the sequence is stop codon suppression. Once formed, the SECIS/SBP2 complex binds to EFsec, which has previously bound sec-$tRNA^{sec}$. When the ribosomal A site becomes vacant (i.e., the UGA codon has moved into position), the SECIS element complex donates the sec-$tRNA^{sec}$, after which a peptide bond forms. The selenocysteine residue is now incorporated into the polypeptide.



Selenocysteine

FIGURE 19B

The 21st Amino Acid

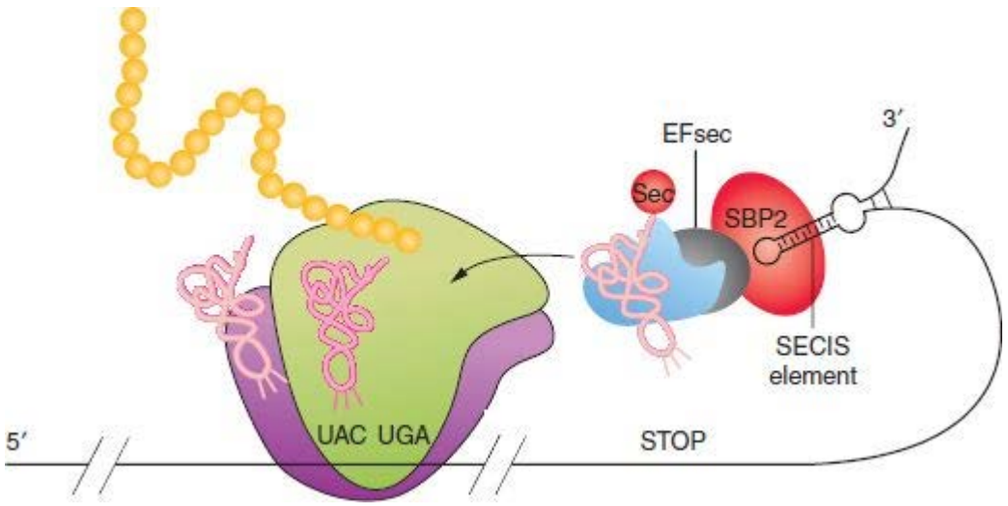


FIGURE 19C

Mechanism of Selenocysteine Incorporation into Eukaryotic Selenoproteins

The stop codon UGA is reassigned for selenocysteine incorporation by the interaction of a complex composed of a SECIS element bound to SBP2 (red) and EFsec (blue and gray) that is charged with sec-tRNA^{sec} with the ribosome. The sec-tRNA^{sec} complex is shown approaching the A site of the 80S ribosome. Once the sec-tRNA^{sec} has been donated to the A site, a peptide bond forms between it and the nascent polypeptide.

SUMMARY In context-dependent codon reassignment, a specific tRNA, a tRNA synthetase, and other molecules are used to transform a stop codon into one that codes for the incorporation of a nonstandard amino acid.

POSTTRANSLATIONAL MODIFICATIONS IN EUKARYOTES Most nascent polypeptides undergo one or more types of covalent modification. These alterations, which may occur either during ongoing polypeptide synthesis or afterward, consist of reactions that modify the side chains of specific amino acid residues or break specific bonds. In general, posttranslational modifications prepare each molecule for its functional role and/or for folding into its native (i.e., biologically active) conformation. More than 200 different types of posttranslational processing reaction have been identified in the human proteome. Most of them occur in one of the following classes.

Proteolytic Cleavage. The proteolytic processing of proteins is a common regulatory mechanism

in eukaryotic cells. Typical examples of proteolytic cleavage (hydrolysis by proteases) include removal of the N-terminal methionine and signal peptides (see p. 780). Proteolytic cleavage is also used to convert inactive precursor proteins, called **proproteins**, to their active forms. Recall, for example, that certain enzymes, referred to as proenzymes or zymogens, are transformed into their active forms by cleavage of specific peptide bonds. The proteolytic processing of insulin (**Figure 19.16**) provides a well-researched example of the conversion of a polypeptide hormone into its active form. The inactive insulin precursor produced by removing the signal peptide is referred to as proinsulin. Inactive precursor proteins with removable signal peptides are called **preproproteins**. The insulin precursor containing a signal peptide is referred to as preproinsulin.

Glycosylation. A wide variety of eukaryotic proteins are glycosylated for structural and informational purposes (p. 270). Glycosylation reactions begin in the ER where N-linked oligosaccharides are synthesized in association with phosphorylated dolichol (**Figure 19.17**). The product of this process, $Glc_3Man_9GlcNAc_2$, is transferred to a nascent polypeptide at an asparagine residue in either of the tripeptide sequences Asn-X-Ser or Asn-X-Thr by oligosaccharyltransferase. (X is any amino acid residue except proline.) Two of the three terminal glucoses are then removed, generating $Glc_1Man_9GlcNAc_2$, the binding site for the Ca^{2+}-requiring molecular chaperones calnexin and calreticulin. Calnexin (a membrane-bound protein) and calreticulin (a lumenal protein) are lectin-like molecules that prevent misfolded proteins from proceeding to the Golgi. (Lectins, described on p. 271, are carbohydrate-binding proteins.) When the glycoprotein is properly folded, the terminal glucose residue is removed, which causes release by the molecular chaperones. The now properly folded glycoprotein is then transferred to the Golgi apparatus where the *N*-oligosaccharide is modified further to yield a high-mannose, complex, or hybrid product (p. 265).

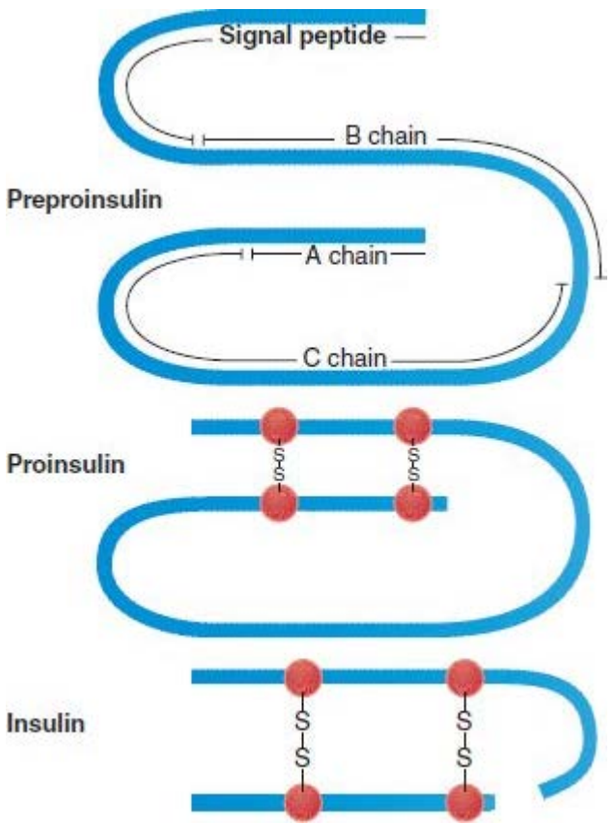


FIGURE 19.16

Proteolytic Processing of Insulin

After the removal of the signal peptide, the peptide segment referred to as the C chain is removed by a specific proteolytic enzyme. Two disulfide bonds are also formed during insulin's posttranslational processing.

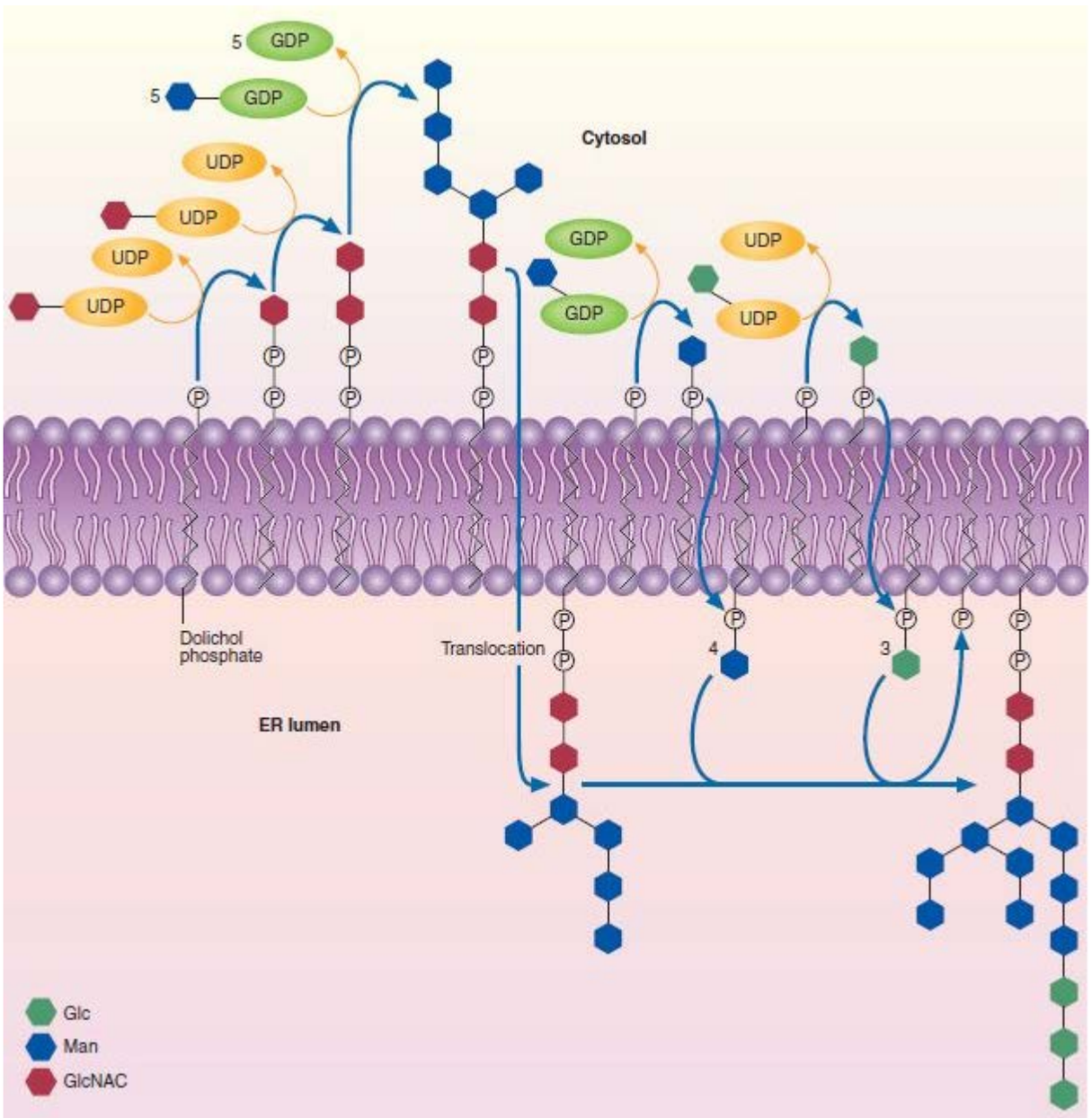


FIGURE 19.17

Synthesis of N-Linked Oligosaccharide

In the first step, GlcNAc-1-P is transferred from UDP-GlcNAc to dolichol phosphate (Dol-P). (Dolichol is a polyisoprenoid found within all cell membranes. Phosphorylated dolichol is found predominantly in ER membrane.) The next GlcNAc and the following five mannose residues are then transferred from nucleotide-activated forms linked to UDP and GDP, respectively. After the entire structure has flipped to the lumenal side of the membrane, each of the remaining sugars (four mannoses and three glucoses) is transferred first to Dol-P and then to the growing oligosaccharide. N-glycosylation of protein takes place in the ER in a one-step reaction catalyzed by a membrane-bound enzyme called glycosyl transferase.

Hydroxylation. Hydroxylation of the amino acids proline and lysine is required for the structural integrity of the connective tissue proteins collagen (Section 5.3) and elastin. Additionally, 4-

hydroxyproline is found in acetylcholinesterase (the enzyme that degrades the neurotransmitter acetylcholine) and complement (a series of plasma proteins involved in the immune response). Three RER mixed-function oxygenases (prolyl-4-hydroxylase, prolyl-3-hydroxylase, and lysyl hydroxylase) are responsible for hydroxylating certain proline and lysine residues. Substrate requirements are highly specific. For example, prolyl-4-hydroxylase hydroxylates only proline residues in the Y position of peptides containing Gly-X-Y sequences, whereas prolyl-3-hydroxylase requires Gly-Pro-4-Hyp sequences (Hyp stands for hydroxyproline; X and Y represent other amino acids). Hydroxylation of lysine occurs only when the sequence Gly-X-Lys is present. (Polypeptide hydroxylation by prolyl-3-hydroxylase and lysyl hydroxylase occurs only before helical structure forms.) **Figure 19.18** illustrates the synthesis of 4-Hyp. Ascorbic acid (vitamin C) is required to hydroxylate proline and lysine residues in collagen. Inadequate dietary intake of vitamin C can result in scurvy. The symptoms of scurvy (e.g., blood vessel fragility and poor wound healing) are the effects of weak collagen fiber structure.

Visit the companion website at **www.oup.com/us/mckee** to read the related Biochemistry in Perspective essay on scurvy for Chapter 7.

Phosphorylation. Examples of the roles of protein phosphorylation in metabolic control and signal transduction have already been discussed. Protein phosphorylation may also play a critical (and interrelated) role in protein–protein interactions. For example, the autophosphorylation of tyrosine residues in platelet-derived growth factor (PDGF) receptors precedes the binding of cytoplasmic target proteins.

Lipophilic Modifications. Lipidation, the covalent attachment of lipid moieties to proteins, improves membrane-binding capacity and/or certain protein–protein interactions. Among the most common lipophilic modifications are acylation (the attachment of fatty acids) and prenylation (Section 11.1). Although the fatty acid myristate (14:0) is relatively rare in eukaryotic cells, myristoylation is one of the most common forms of acylation. N-Myristoylation (the formation of an amide bond between the myristoyl group of myristic acid and the α-amino group of a polypeptide's N-terminal glycine residue) has been shown to increase the affinity of the α subunit of certain G proteins (p. 606) for membrane-bound β and γ subunits.

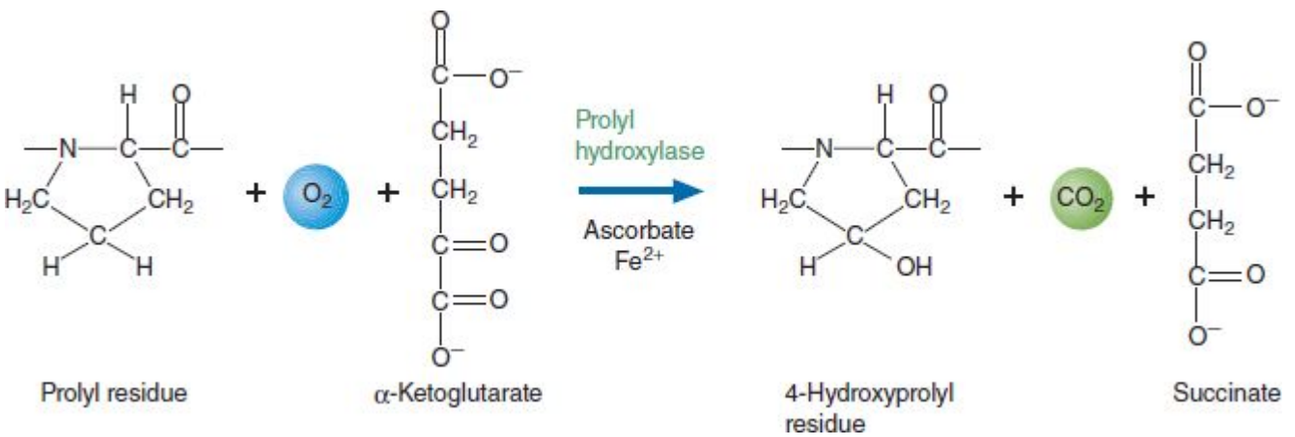


FIGURE 19.18

Hydroxylation of Proline

Prolyl-4-hydroxylase, the enzyme that catalyzes the hydroxylation of the C-4 position of certain prolyl residues in nascent polypeptides, is an example of a dioxygenase. Both oxygen atoms of O_2 are incorporated into the two substrates, α-ketoglutarate and the prolyl residue, to form the two products, succinate and the 4-hydroxyproline residue. In the reaction mechanism, Fe(II) and O_2 form a cyclic peroxide with α-ketoglutarate that facilitates the decarboxylation of α-ketoglutarate to form succinate and Fe(IV)=O, which serves as a substrate for proline hydroxylation. Ascorbic acid, acting as a reducing agent, restores cofactor iron to the ferrous state.

Methylation. Protein methylation serves several purposes in eukaryotes. The methylation of altered aspartate residues by a specific type of methyltransferase promotes either the repair or the degradation of damaged proteins. Other methyltransferases catalyze reactions that alter the cellular roles of certain proteins. For example, methylated lysine residues have been found in such disparate proteins as ribulose-2,3-bisphosphate carboxylase, calmodulin, histones, certain ribosomal proteins, and cytochrome c. Other amino acid residues that may be methylated include histidine (e.g., histones, rhodopsin, and eEF2) and arginine (e.g., hsps and ribosomal proteins).

Carboxylation. Vitamin K-dependent carboxylation of glutamyl residues to form γ-carboxyglutamyl residues increases a protein's sensitivity to Ca^{2+}-dependent modulation. The carboxylation process requires an NADPH-dependent reductase to convert phylloquinone, a vitamin K quinone (**Figure 11.16**), to its hydroquinone form, a carboxylase that adds a carboxyl group to the γ-carbon of a glutamyl residue, and an epoxide reductase that converts the vitamin K-2,3-epoxide product of the carboxylation reaction to the original vitamin K quinone. The target protein must contain an appropriate amino acid sequence that binds to the carboxylase enzyme and a $(\text{Glu}XXX)_n$ repeat (n = 3−12, X = other amino acids). Many of the known target proteins are involved in blood clotting (factors VII, IX, and X, and prothrombin). The anticoagulant coumadin (warfarin), used to prevent blood clot formation, acts by inhibiting the two reductases required in protein carboxylation.

Disulfide Bond Formation. Disulfide bonds are generally found only in secretory proteins (e.g., insulin) and certain membrane proteins. (Recall that disulfide bridges are favored in the oxidizing environment outside the cell and confer considerable structural stability on the molecules that contain them.) As described earlier (Section 5.3), cytoplasmic proteins generally do not possess disulfide bonds because of the reducing conditions within cytoplasm. The ER has a nonreducing environment, so disulfide bonds form spontaneously in the RER as nascent polypeptides emerge into the lumen. Although some proteins have disulfide bridges that form sequentially as the polypeptide enters the lumen (i.e., the first cysteine pairs with the second, the third residue pairs with the fourth, etc.), this is not true for many other molecules. Proper disulfide bond formation for these latter proteins is facilitated by **disulfide exchange**, a process in which disulfide bonds rapidly migrate from one position to another until the most stable structure is achieved. An ER enzymatic activity, referred to as protein disulfide isomerase (PDI), is a thioredoxin-like enzyme that catalyzes this process. Protein disulfide isomerase also acts as a chaperone by rescuing wrongly folded polypeptides.

TARGETING Despite the vast complexities of eukaryotic cell structure and function, each newly synthesized polypeptide is normally directed to its proper destination. There appear to be two principal mechanisms by which polypeptides are directed to their correct locations: transcript localization and signal peptides.

Transcript localization, directing mRNAs to discrete cellular locations, begins in the nucleus with the assembly of mRNPs, in which diverse sets of RNA-binding proteins bind to specific

3′UTR sequences. For example, mRNP cargo moves along cytoskeletal filaments to discrete cell locations via attachment to motor proteins such as kinesin, dynein, or myosin. It is the structural properties of the proteins bound to specific mRNA localization signals (or “zip codes”) that dictates this activity. Once an mRNP leaves the nucleus, its proteins determine not only its destination but also the degree to which its mRNA will be translated. Bicoid mRNP is a well-researched example.

Bicoid is a *morphogen* (a molecule whose nonuniform distribution directs pattern formation during development) within *Drosophila* oocytes (immature egg cells). Mature *Drosophila* eggs contain a gradient of bicoid. A high concentration of bicoid in the anterior portion of the egg is required for the normal development of anterior body parts (i.e., head segments), whereas the low bicoid concentration in the posterior portion of the egg cytoplasm promotes the development of posterior body parts. If posterior cytoplasm is removed from one egg and substituted for anterior cytoplasm in a second egg, two sets of posterior body parts appear in the larva that develops from the recipient egg. Originally produced by nearby nurse cells, bicoid mRNP is transported into oocytes, where it is carried along microtubules toward the anterior end of the cell via the interaction between dynein (p. 63) and Egl, a protein component of bicoid mRNP. After the mature egg has been fertilized, translation of bicoid mRNA, coupled with protein diffusion, gives rise to the concentration gradient.

The **signal hypothesis** explains how polypeptides destined for secretion or for use in the plasma membrane or any of the membranous organelles are specifically targeted to their proper location by sorting signals referred to as **signal peptides**. Each signal peptide sequence facilitates the insertion of the polypeptide that contains it into an appropriate membrane. Signal peptides generally consist of a positively charged region followed by a central hydrophobic region and a more polar region. Although numerous signal peptides occur at the amino terminal, they may also occur elsewhere along the polypeptide.

The transfer of a polypeptide across the RER membrane (**Figure 19.19**) begins as soon as about 70 amino acids have been incorporated into the polypeptide that emerges from a ribosome. A rod-like ribonucleoprotein complex called the **signal recognition particle** (SRP) then binds to the ribosome. The SRP, which consists of six proteins and a 7S RNA, is a GTPase that recognizes and transiently binds short RER signal sequences with about eight nonpolar amino acid residues. As a result of SRP-ribosome binding, the EF2 binding site is blocked and translation is arrested. The SRP mediates binding of the ribosome to the ER via *docking protein*, a heterodimer composed of two GTPases, also referred to as **SRP receptor protein**. Once ribosome binding to the RER has occurred, both docking protein GTPs are hydrolyzed. Subsequent SRP release allows translation to restart. The growing polypeptide then inserts into the **translocon**, a protein complex composed of a transmembrane pore and several associated proteins that facilitate translocation and processing. When polypeptide synthesis and translocation occur simultaneously, the process is referred to as **cotranslational transfer**. In **posttranslational translocation**, previously synthesized polypeptides are transported across the RER membrane by ATP-binding translocon-associated molecular chaperones hsp40 and hsp70.

The fate of a targeted polypeptide depends on the location of the signal peptide and any other signal sequences. As illustrated in **Figure 19.19**, soluble secretory protein transmembrane transfer is usually followed by removal of an N-terminal signal peptide by signal peptidase, a process that releases the protein into the ER lumen. Such molecules usually undergo further posttranslational processing. The initial phase of the translocation of transmembrane proteins is similar to that of secretory proteins. For these molecules, the amino-terminal signal peptide serves as a *start signal* that remains bound in the membrane as the remaining polypeptide sequence is threaded through the membrane. So-called single-pass transmembrane proteins possess a *stop transfer signal* (or

stop signal), which prevents further transfer across the membrane (**Figure 19.20a**). Membrane proteins with multiple membrane-spanning segments (multipass) possess a series of alternating start and stop signals (**Figure 19.20b**).

Most proteins that are translocated into the RER are directed to other destinations. After undergoing initial posttranslational modifications, both soluble and membrane-bound proteins are transferred to the Golgi complex via transport vesicles that bud off from the ER and fuse with the cis face of the Golgi membrane (**Figure 19.21**). Proteins that ultimately reside in the ER possess retention signals. In most vertebrate cells, this signal consists of the carboxy-terminal tetrapeptide Lys-Asp-Glu-Leu (KDEL) sequence.

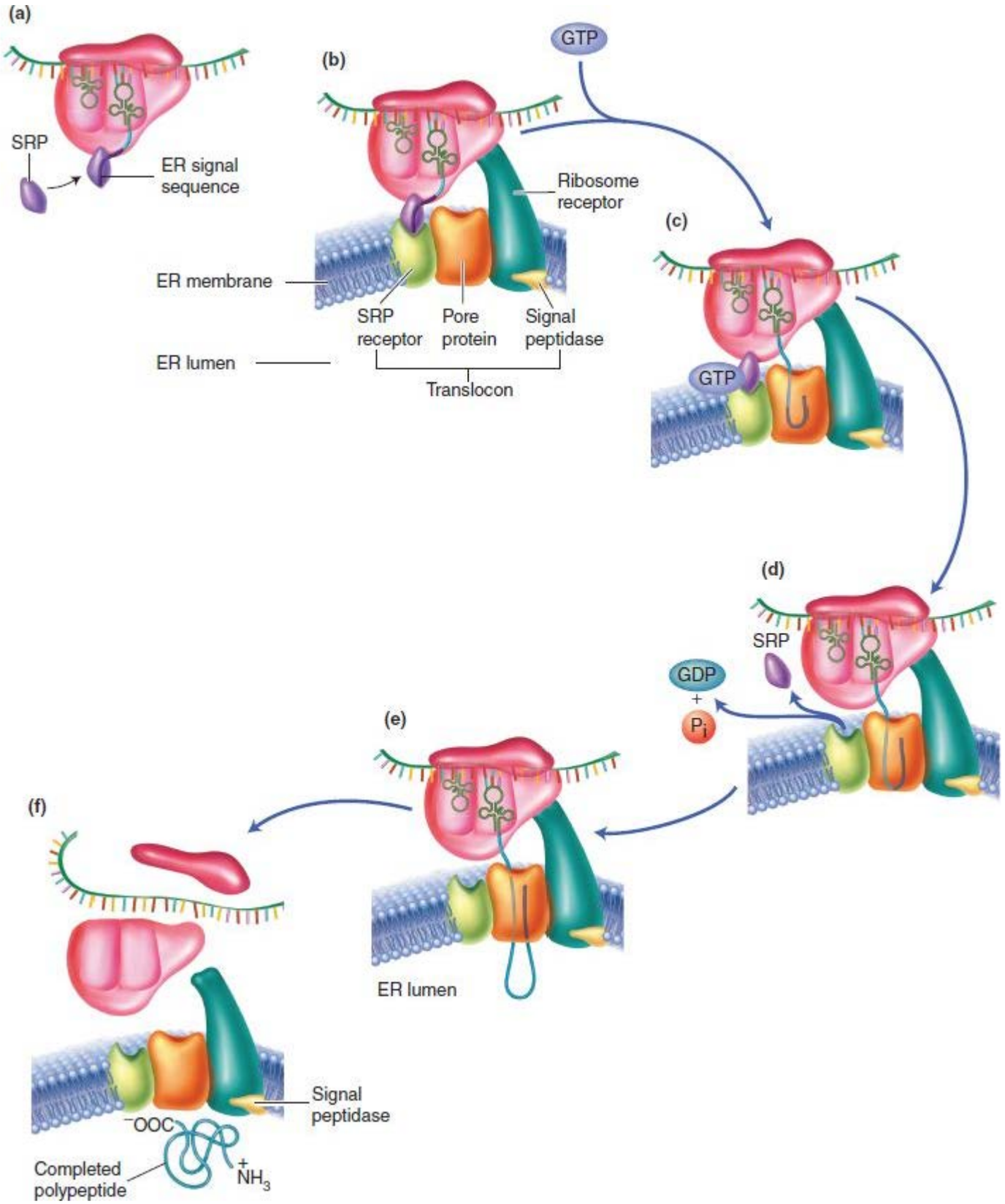


FIGURE 19.19

Cotranslational Transfer across the RER Membrane

(a) When the nascent polypeptide is long enough to protrude from the ribosome, the SRP binds to the signal sequence, causing a transient cessation of translation. (b) The subsequent binding of SRP to the SRP receptor results in the binding of the ribosome to the translocon complex in the RER membrane. (c) Polypeptide synthesis begins again as GTP binds to the SRP–SRP receptor complex. GTP hydrolysis accompanies the binding of the signal sequence to the translocon and (d) the dissociation of SRP from its receptor. (e) The

polypeptide continues to elongate until (f) translation is completed. The signal peptide is removed by signal peptidase in the RER lumen. The polypeptide is released into the lumen.

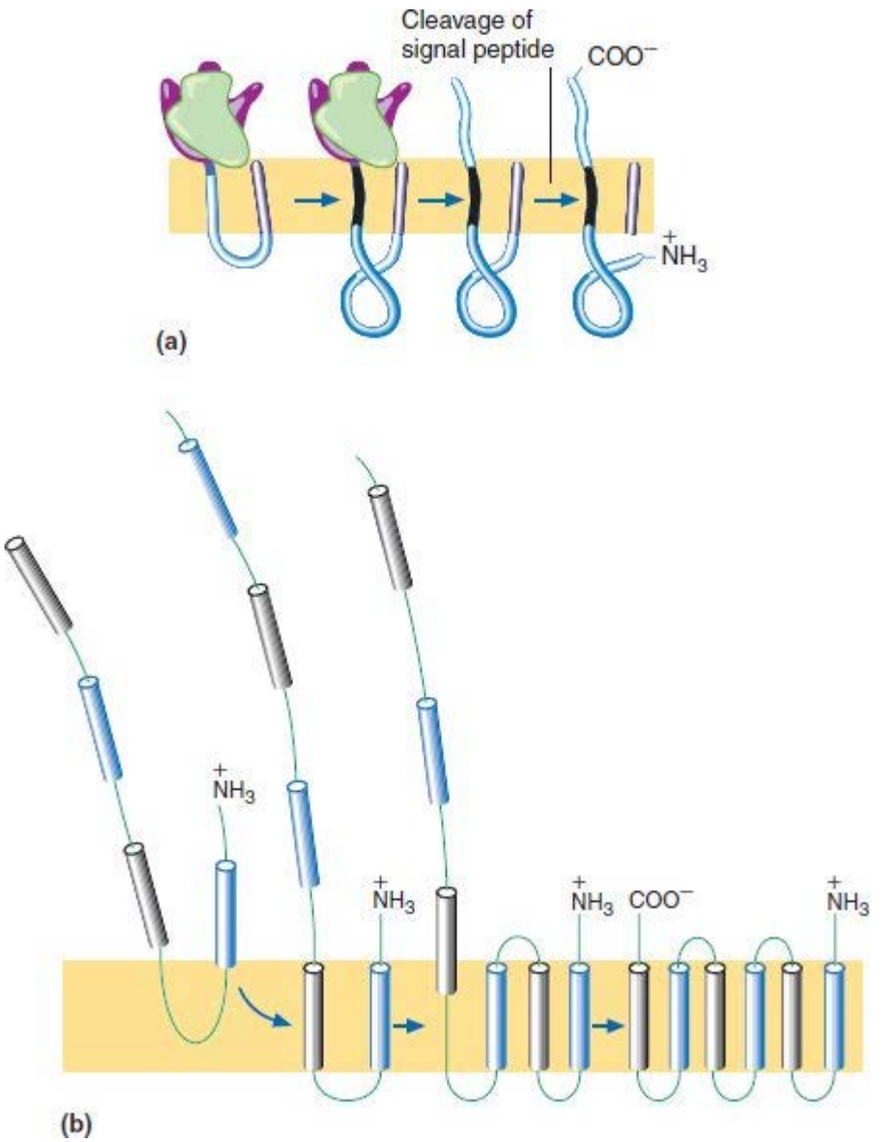


FIGURE 19.20

Cotranslational Transfer of Integral Membrane Proteins

(a) Transfer of a single-pass transmembrane protein. (b) Transfer of a multipass membrane protein. For the sake of clarity, the transfer apparatus has been omitted from the diagrams. In addition, the ribosome has been omitted from (b). The shaded segment is a signal peptide. The black segment is a stop transfer signal.

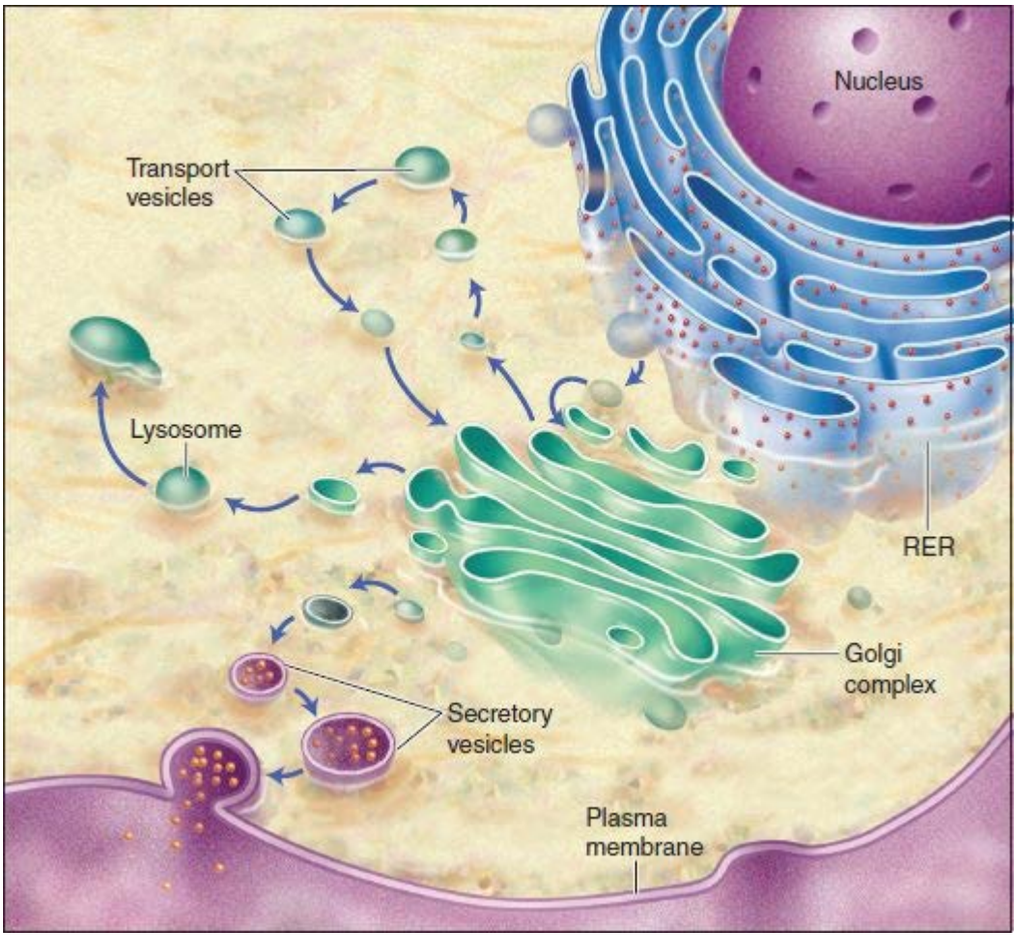


FIGURE 19.21

The ER, Golgi, and Plasma Membrane

Transport vesicles transfer new membrane components (protein and lipids) and secretory products from the ER to the Golgi complex, from one Golgi cisterna to another, and from the *trans*-Golgi network to other organelles (e.g., lysosomes) or to the plasma membrane.

Within the Golgi complex, proteins undergo further modifications. For example, N-linked oligosaccharides are processed further, and O-linked glycosylation of certain serine and threonine residues occurs. Lysosomal proteins are targeted to the lysosomes by adding a mannose-6-phosphate residue. It is still unclear what signals direct secretory proteins to the cell surface (via exocytosis) or promote the delivery of plasma membrane proteins to their destination, although a "default mechanism" has been proposed. (In default mechanisms, the absence of a signal results in a specific sequence of events.) When protein modification is complete, transport vesicles exit from the trans face of the Golgi and move to their target locations.

Proteins targeted to mitochondria are synthesized on cytoplasmic ribosomes as preproteins and then bind to a multichaperone complex composed of the ATPases hsp70 and hsp90. Transport into a mitochondrion begins with the docking of the molecular chaperone-bound preprotein to receptors of the TOM complex (*t*ranslocase of the *m*itochondrial *o*uter *m*embrane). The delivery of preprotein to the TOM complex protein-conducting channel is an ATP-dependent process. About half of imported mitochondrial proteins are translocated into the mitochondrial matrix. The translocation of a matrix protein (**Figure 19.22**) begins with the transfer of the preprotein through the TOM channel. As the N-terminal signal sequence emerges in the intermembrane space, it binds to a receptor of the TIM (*t*ranslocase of the mitochondrial *i*nner *m*embrane) complex. The preprotein is then transported through the TIM channel in a process driven by the electrochemical

proton gradient of the inner membrane. Once it is in the matrix, the preprotein's signal sequence is cleaved by a signal peptidase. ATP-dependent mitochondrial matrix chaperones, mthsp70 and mthsp60, assist in the folding of the protein into its biologically active conformation.

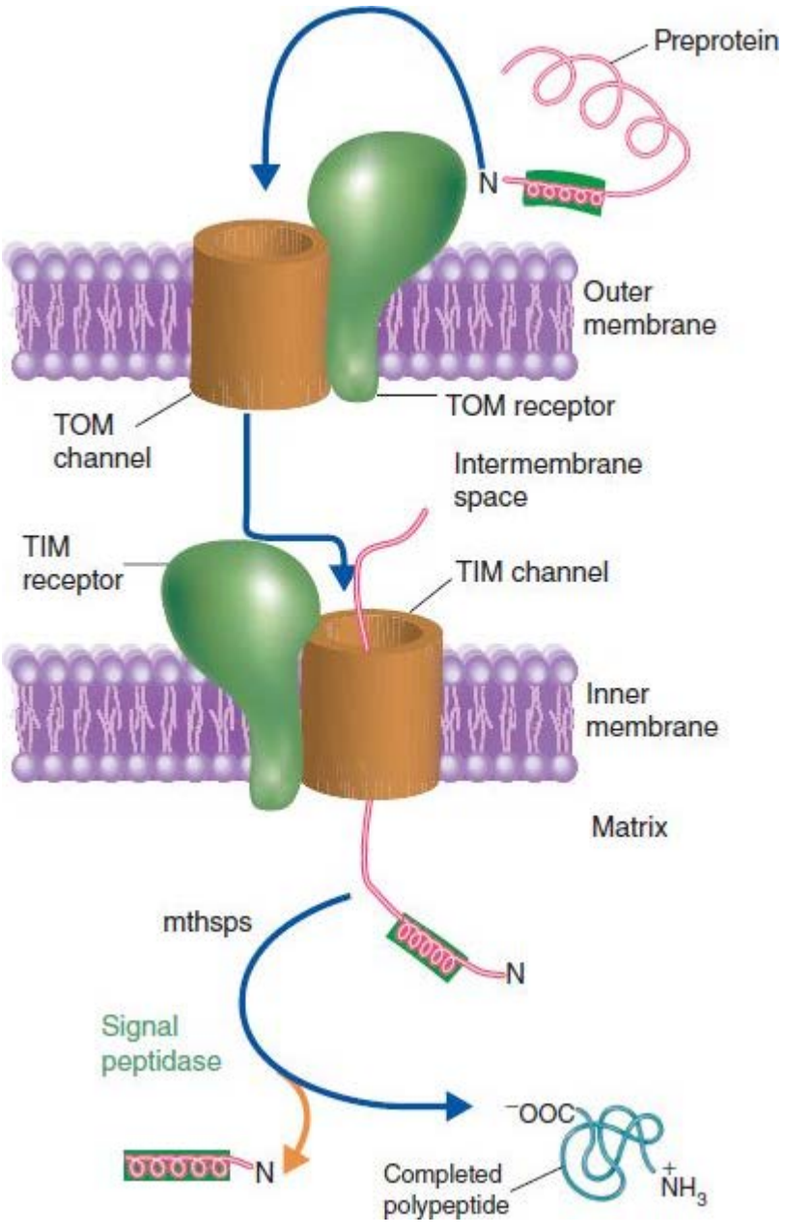


FIGURE 19.22

Import of a Mitochondrial Matrix Protein

The preprotein, which is complexed with a multichaperone complex (not shown), is recognized by a TOM receptor protein via its N-terminal signal sequence (an α-helix in which there are hydrophobic residues on one side and positively changed residues on the other). Preprotein translocation through the TOM transmembrane channel is assisted by ATP hydrolysis catalyzed by the hsp70 and hsp90 components of the multichaperone complex. Once the N-terminal signal sequence enters the intermembrane space and is recognized by a TIM receptor, the preprotein enters the TIM channel. Translocation of the preprotein across the inner membrane is driven by the electrochemical proton gradient created by the electron transport process. Once the preprotein enters the matrix, it is processed by a signal peptidase and mitochondrial chaperones to yield the final biologically active protein.

TRANSLATION CONTROL MECHANISMS Although many aspects of eukaryotic translational control are currently unresolved, the following features are believed to be important.

mTOR-Mediated Translational Control. Recall that the mTORC1 signaling pathway (p. 612) integrates nutrient availability, energy levels, and hormone and growth factor signals. Because amino acid polymerization consumes a substantial portion of cellular resources, it is not surprising that mTORC1 has a significant effect on the rate of protein synthesis. mTORC1, a nutrient, energy, and redox sensor, activates translation in a process that involves three proteins: eIF4E, ribosomal protein S6, and eEF2. The availability of eIF4E (cap-binding protein) for cap-binding complex formation is regulated by eIF4E-binding protein 1 (eIF4E-BP1). In its inactive state, eIF4E is bound to hypophosphorylated eIF4E-BP1, thereby preventing an interaction with eIF4G. eIF4E is activated when mTORC1 phosphorylates eI4E-BP1, allowing it to separate from eIF4E-BP1. As a result, the eIF4F complex forms, allowing cap-dependent initiation to proceed. mTORC1 also promotes the activating S6 kinase 1 (SK1)–catalyzed phosphorylation of the ribosomal protein S6. SK1 also upregulates the activity of the mRNA helicase eIF4B. Finally, mTORC1 signaling activates eEF2 by stimulating the phosphorylation of the eEF2-inhibiting enzyme eEF2 kinase.

mRNA EXPORT The spatial separation of transcription and translation afforded by the nuclear membrane appears to provide eukaryotes with significant opportunities for gene expression regulation. mRNP export through the nuclear pore complex (p. 739) is known to be a carefully controlled, energy-driven process whose minimum requirements include the presence of a 5′-cap and a 3′ poly(A) tail. As described (p. 739) mRNAs are exported as mRNPs. In addition to exon junction complexes (p. 729) and export proteins (p. 739), an elaborate group of targeting proteins are associated with mRNAs.

mRNA STABILITY Transcription of a gene is not a guarantee that it will be translated, since there is an intricate network of RNA-binding proteins and ncRNAs (lncRNAs and miRNAs) that regulate moment-by-moment cellular metabolic and signaling processes. Changes in environmental cues or stress conditions can rapidly alter nuclear regulatory processes (e.g., lncRNA and miRNA transcription). If an mRNP reaches the cytoplasm in a manner conducive to translation, its most important property is stability, that, is, how long it can survive without being degraded by nucleases. In general, the translation rate of any mRNA species is related to its abundance, which is in turn dependent on its rates of both synthesis and degradation. mRNA half-lives range from about 20 minutes to more than 24 hours.

Several features of mRNA structure are known to affect its stability. The presence of certain sequences may confer resistance to nuclease action (e.g., palindromes that create hairpins), whereas other sequences may increase the likelihood of nuclease action, particularly if present in multiple copies. The binding of specific RNA-binding proteins to certain sequences can also affect mRNA stability. For example, AU-rich elements (AREs), found in the 3’-UTR in numerous mammalian mRNAs (about 7% of human mRNAs), allow cells to rapidly respond to changing environmental conditions (e.g., signal transduction, nutrient transport, or stress conditions). Depending on the type of ARE–RNA-binding protein interactions and the type of alternate splicing 3’UTR variant, an mRNA may be stabilized or destabilized. Decay of destabilized mRNAs begins with excision of the poly(A) tail and the release of PABPs, followed by the actions of 3’, 5’-exonucleases.

KEY CONCEPTS

- Eukaryotic protein synthesis is slower and more complex than that of its prokaryotic counterpart. In addition

to requiring a larger number of translation factors and a more complex initiation mechanism, the eukaryotic process involves vastly more complicated posttranslational processing and targeting mechanisms.

- Eukaryotes use a wide spectrum of translational control mechanisms.

QUESTION 19.5

The mechanism involved in the posttranslational transport of proteins into chloroplasts has so far received only limited attention. However, the import of plastocyanin into the thylakoid lumen has been determined to require two import signals near the N-terminal of the newly synthesized protein. Assuming that chloroplast protein import resembles the import process for mitochondria, suggest a reasonable hypothesis to explain how plastocyanin (a lumen protein associated with the inner surface of the thylakoid membrane) is transported and processed. What enzymatic activities and transport structures do you expect are involved in this process?

19.3 THE PROTEOSTASIS NETWORK

Within the highly crowded and dynamic interior of living cells, millions of proteins perform a vast array of functions such as DNA replication and transcription, cell signal transduction, immune responses, cell cycle control, and molecular transport. Life depends on the proper function of proteins, which in turn requires that these linear macromolecules fold into their “native states” and yet retain some degree of conformational flexibility. As a result, many proteins, especially those that are composed of 100 or more amino acids or are completely or even partially unstructured, are marginally stable and therefore prone to misfolding. Misfolded or partially misfolded proteins often have exposed hydrophobic patches that may interact with other molecules to form amorphous aggregates. In addition, some misfolded molecules may rearrange to form the β-strands of amyloid fibrillar aggregates. The proteome is also challenged by a constant barrage of metabolic and environmental stresses (e.g., heat or heavy metals exposure, amino acid side chain oxidation, hypoxia, and toxins) that can damage them. When combined with the incidence of random errors in protein synthesis, proteotoxic stress-related protein misfolding and other types of damage are a severe threat to cell function.

Healthy young cells maintain proteostasis (p. 39) with a robust and highly conserved interconnected network of pathways, referred to as the **proteostasis network** (PN) (**Figure 19.23**). Using stress-responsive signaling pathways, the PN monitors proteins from their synthesis by ribosomes, through folding, refolding, and transport or degradation when their useful life is over or they are damaged. PN processes are accomplished with the aid of molecular chaperones (p. 170), stress-response transcription factors, detoxifying enzymes, and degradation processes such as the ubiquitin-proteosomal system (p. 572) and autophagy (p. 574). The resources that are devoted to proteome protection indicate the importance of the PN. For example, the human PN involves about 2000 genes. Under stressful conditions, PN processes can be activated throughout the cell, that is, cytoplasm and the nucleus and other organelles (e.g., the unfolded protein response in the ER [erUPR] (p. 49) and mitochondria [mtUPR]).

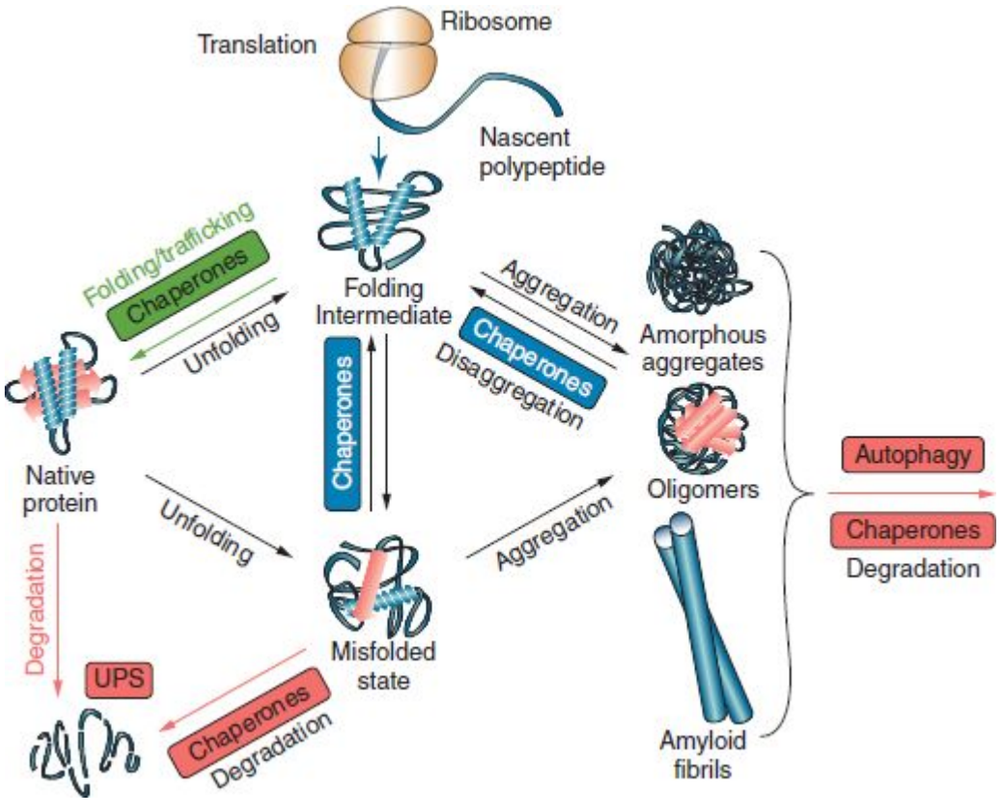


FIGURE 19.23

The Proteostasis Network

The proteostasis network consists of molecular chaperones that assist proteins in de novo folding and in maintaining them in their native states. The network also includes enzymes and protein complexes that degrade misfolded, damaged, and obsolete proteins. As each nascent polypeptide emerges from the exit tunnel, it encounters ribosome-associated chaperones. If necessary, additional folding assistance is provided by downstream molecular chaperones such as the hsp70s and hsp90s and their associated proteins. Misfolded proteins are degraded by a combination of chaperones and E3 ubiquitin ligases that together recognize and target them for destruction by the UPS (ubiquitin proteosomal system). Aggregated proteins that resist digestion by proteasomes are removed by autophagy.

The Heat Shock Response

The best understood cellular stress response is the *heat shock response* (HSR). As with other stress responses, the HSR works to protect an affected cell and its proteome from heat-induced damage (as well as oxidative stress and heavy metals). It does so by rapid and global changes in gene expression that inhibit nonessential protein synthesis on ribosomes and increase the concentration of PN components. In *E. coli*, HSR is mediated by σ^{32}, the bacterial transcription initiation factor that directs RNA polymerase to the promoters of HSR target genes. Synthesis of σ^{32} is initiated by heat-induced melting of the secondary structure of the 5′ UTR of the σ^{32}-coding mRNA, revealing the molecule's Shine–Dalgarno sequence (p. 763). In eukaryotes, HSR (heat-induced cytoplasmic protein damage repair) is initiated by heat shock factors (HSFs), most notably heat shock factor 1 (HSF1). In unstressed cells, HSF1 exists in cytoplasm as a monomer bound to hsp90. Heat stress triggers the assembly of an HSF1 trimer that relocates to the nucleus, where it stimulates the activity of chromatin remodeling proteins and the subsequent transcription of HSR genes. HSF1 activation is the result of hsp90 displacement as the molecular chaperone attempts to refold heat-damaged proteins. HSF1 trimer formation occurs as a result of the binding of the lncRNA molecule heat shock RNA1 (HSR1) to the translation elongation factor eEF1A.

The Proteostasis Network and Human Disease

Defective protein folding is responsible for a large number of human diseases, which are referred to as protein-folding diseases or protein conformational disorders. Some protein-folding diseases involve a single genetic mutation that causes a change in a polypeptide's sequence, which in turn results in its improper folding. Examples of such loss-of-function disorders include cystic fibrosis (pp. 431–33) and CBS deficiency (a disorder of methionine metabolism caused by defective cystathionine β-synthase; refer to **Figure 15.12**). Mutations in the gene coding for p53, a major tumor suppressor protein, occur in more than 50% of all cancers. Normal p53 regulates the cell cycle, monitors the genome for damage, and can initiate apoptosis if DNA damage is irreparable. A mutated and improperly folded p53 results in inadequately repaired DNA damage, a circumstance that eventually allows cells to grow in an uncontrolled manner.

In a large number of protein-folding disorders, there is chronic proteostasis dysfunction that arises from adverse interactions between aggregated proteins and proteosomal components. Examples include Alzheimer's, Huntington's, and Parkinson's diseases and type 2 diabetes. A major risk factor for these and other protein-folding diseases is aging, since the PN progressively declines in effectiveness as individuals grow older. A common mechanism in these disorders is the misfolding of specific proteins, such as β-amyloid (Alzheimer's disease), α-synuclein (Parkinson's disease), and amylin (type 2 diabetes). When aging cells fail to degrade these molecules, a small number of them form oligomers. Oligomers eventually associate to form insoluble aggregates that may entangle other denatured proteins and components of the now seriously compromised PN. The aggregated proteins are described as having a toxic gain of function.

Biochemistry IN THE LAB

Proteomics

The technology of proteomics is being developed to investigate the proteome, the functional output of the genome. The proteome of each organism or cell type includes not only the entire set of the protein products of mRNA translations, but also all of their covalent modifications, few of which can be predicted from mRNA sequences. The goals of proteomics are primarily twofold: to study the global changes in the expression of cellular proteins and to determine the identity and functions of all the proteins in the proteomes of organisms. The potential applications for proteomic research are many and varied. In addition to providing opportunities to resolve basic biological problems (e.g., ascertaining the precise mechanisms by which cellular processes such as neuronal transport or mRNA splicing occur), proteomics-based technology has obvious uses in biomedical research. Examples of the latter include investigations of the causes and diagnosis of genetic and infectious diseases and the development of effective drugs.

Proteomic Tools

The investigation of proteomes is currently focused primarily on developing fast, automated methods for identifying and characterizing the proteins produced by normal and diseased cells.

Laser capture dissection microscopy (LCDM), which uses a laser beam to harvest single cells from a tissue without altering or damaging the samples collected, is providing scientists with a powerful new tool. Combined with increasingly sophisticated proteomic technologies and bioinformatic analysis, LCDM is having a significant impact on the pace of discovery. Among the most useful proteomic technologies is **mass spectrometry** (MS).

As described previously (p. 187), MS is a technique in which molecules are vaporized and then bombarded by a high-energy electron beam, which causes them to fragment as cations. As the ionized fragments enter the spectrometer, they pass through a strong magnetic field that separates them according to their mass-to-charge (*m/z*) ratio. Each type of molecule is identified by the pattern of fragments that is generated, each pattern or "fingerprint" being unique. Because proteins do not vaporize, they are instead digested and then dissolved in a volatile solvent and sprayed into the vacuum chamber of the mass spectrometer. The electron beam ionizes these peptide fragments, and the positively charged peptides are passed through the magnetic field. The peptide mass fingerprint that results is then compared to fragmentation information in protein databases. Although MS is highly accurate and automated, it is usually insufficient for identifying all the proteins in a sample. To improve protein identification, tandem MS (MS/MS), a method in which two mass spectrometers are linked in tandem, has been developed. In this technique, the oligopeptide fragments produced in the first MS are transferred to the second MS, where they are further fragmented and analyzed. MS/MS is used to rapidly sequence proteins.

Quantitative analysis of complex protein mixtures (e.g., blood serum and plasma) can be analyzed efficiently by *shotgun proteomics* using *HPLC* (*high-performance liquid chromatography*, the analysis of sample mixtures that uses pumps to pass a pressurized liquid solvent through a solid matrix-containing column) and tandem mass spectrometry. Using small amounts of blood, for example, shotgun proteomics can identify large numbers of proteins in a digested sample, separate them by HPLC, and then identify peptides by tandem mass spectroscopy, all within several hours.

Protein Microarrays

Protein microarrays, also called protein chips, are a high-throughput proteomic method. Protein chips are manufactured arrays of specific molecules, called *capture molecules*, on a solid support (e.g., a glass slide, bead, or microwell plate). Capture molecules can include antibodies, receptors, or enzymes, among others. The most commonly used detection method involves the use of fluorescent tags, although other detection methods are available (e.g., carbon nanowire sensors that measure differences in conductance). An extracted protein mixture (labeled with fluorescent dyes) is first applied to a slide containing a library of immobilized antibodies. Binding between specific antibodies and proteins in the mixture is detected by a fluorescence scan. Protein microarrays are replacing older, slower proteomic methods such as 2-D gel electrophoresis.

Yeast Two-Hybrid Screening

Two-hybrid screening is a technique that uses genetically modified yeast cells to detect protein–protein and protein–DNA interactions that may indicate a functional relationship. A productive binding of two proteins is identified as follows. One of the proteins, called the "bait protein," is fused, using DNA recombinant technology, to the DNA-binding domain fragment of a specific transcription factor. The other protein is fused to the "prey protein," an activation domain fragment of the same transcription factor. Plasmids containing the DNA sequences coding for the two fused proteins are then simultaneously introduced into mutant yeast cells. The binding of the bait and

prey segments of the two fused proteins results in the formation of a functional transcription factor. This in turn causes a change in the cell's phenotype. Thus, the transcription of a reporter gene results in the synthesis of an easily detectable protein.

Despite the recent advances in proteomic research techniques, several problems are posing a serious barrier to accomplishing the enormous task of characterizing entire proteomes. Among the most important of these problems are the difficulties in handling proteins (susceptibility to denaturation) so that they remain properly folded and functional and the lack of a technique equivalent to PCR for amplification of proteins found in very small amounts.

Chapter Summary

1. Protein synthesis is a complex process in which information encoded in nucleic acids is translated into the primary sequence of proteins. During the translation phase of protein synthesis, the incorporation of each amino acid is specified by one or more triplet nucleotide base sequences, referred to as codons.
2. The genetic code consists of 64 codons: 61 codons that specify the amino acids and 3 stop codons. One of the stop codons (UGA) can be used to code for the nonstandard amino acid selenocysteine in the synthesis of selenoproteins such as glutathione peroxidase.
3. Translation involves the tRNAs, a set of molecules that act as carriers of the amino acids. The base-pairing interactions between mRNA codons and the anticodon base sequence of tRNAs result in the accurate translation of genetic messages.
4. Translation consists of three phases: initiation, elongation, and termination. Each phase requires several types of protein factor. Although prokaryotic and eukaryotic translational mechanisms bear a striking resemblance to each other, they differ in several respects. One of the most notable differences is the identity, quantity, and function of the translation factors.
5. A unique set of posttranslational modifications can occur that prepare the polypeptide for its functional role, assist in folding, or target it to a specific destination. These covalent alterations include proteolytic processing, modification of certain amino acid side chains, and insertion of cofactors.
6. Prokaryotes and eukaryotes differ in their usage of translational control mechanisms. Prokaryotes use variations in Shine–Dalgarno sequences and negative translational control (the repression of the translation of a polycistronic mRNA by one of its products). In contrast, a wide variety of eukaryotic translational controls have been observed. These mechanisms range from global controls in which the translation rate of a large number of mRNAs is altered to controls in which the translation of a specific mRNA or small group of mRNAs is altered.
7. The proteostasis network (PN) is a system of interconnected pathways that maintain the structural integrity of cellular proteins. The PN monitors proteins from their synthesis by ribosomes, through folding, refolding, transport, and degradation (via the ubiquitin proteasomal system or autophagy). The heat shock response protects cells from stress-induced damage (e.g., heat, oxidative stress, and heavy metals) by inducing rapid and global changes in gene expression that can repair damage if it is not too severe.
8. Proteomics is a technology that is used to investigate the proteome, the complete set of proteins produced from an organism's genome. The goals of proteomics are to study the global changes in the expression of cellular proteins over time and to determine the identity and functions of all the proteins produced by organisms.

Take your learning further by visiting the **companion website** for Biochemistry at **www.oup.com/us/mckee** where you can complete a multiple-choice quiz on protein synthesis to help you prepare for exams.

Suggested Readings

Diebel KW, et al. 2016. Beyond the ribosome: extra-translational functions of tRNA fragments. Biomarker Insights 11(S1):1–8.

Garcia-Maurino SM, et al. 2017. RNA binding protein regulation and cross-talk in the control of AU-rich mRNA fate. Frontiers Mol Biosci doi:10.3389/fmolb.00071.

Hinnebusch AG, et al. 2016. Translational control by 5'-untranslated regions of eukaryotic mRNAs. Science 352:1413–16.

Kurosaki T, Maquat LE. 2016. Nonsense-mediated mRNA decay in humans at a glance. J Cell Sci 129:461–7.

Lane N. 2009. Life ascending. New York (NY): W. W. Norton.

Lazzaewtti D, Bono F. 2017. mRNA localization in metazoans: a structural perspective. RNA Biol 14(11):1473–84.

Leppek K, et al. 2018. Functional 5'UTR mRNA structures in eukaryotic translation regulation. Nat Rev Mol Cell Biol 19(3):158–74.

Mills EW, Green R. 2017. Ribosomopathies: there's strength in numbers. Science 358:608.

Schuller AP, Green R. 2018. Roadblocks and resolutions in eukaryotic translation. Nat Rev Mol Cell Biol 19(8):526–41.

Winata CL, Korzh V. 2018. The translational regulation of maternal mRNAs in time and space. FEBS Lett doi:10.1002/1873-3468.13183.

Yamashita A, Takeuchi O. 2017. Translational control of mRNAs by 3'-untranslated region binding proteins. BMB Reports 50(4):194–200.

Key Words

aminoacyl-tRNA synthetase, *757*

anhydride, *758*

anticodon, *756*

AU-rich element, *784*

cap-binding complex, *768*

codon, *754*

context-dependent codon reassignment, *775*

cotranslational transfer, *780*

decoding center *762*

disulfide exchange, *779*

elongation, *759*

48S initiation complex, *769*

43S preinitiation complex, *768*

genetic code, *754*

GTPase associated region, *762*
initiation, *759*
Kozak sequence, *768*
mass spectrometry, *787*
mixed anhydride, *758*
nascent, *765*
open reading frame, *755*
peptidyl transferase center, *762*
poly(A)-binding protein, *769*
polysome, *759*
posttranslational modification, *761*
posttranslational translocation, *780*
preproprotein, *776*
proprotein, *776*
proteomics, *753*
proteostasis network, *785*
releasing factor, *765*
ribosome recycling factor, *766*
SECIS element, *775*
Shine–Dalgarno sequence, *763*
signal hypothesis, *780*
signal peptide, *780*
signal recognition particle, *780*
SRP receptor protein, *780*
targeting, *761*
termination, *759*
transcript localization, *779*
translocation, *759*
translocon, *780*
wobble hypothesis, *756*

Review Questions

SECTION 19.1

Comprehension Questions

1. Define the following terms:
    a. proteomics
    b. translation
    c. genetic code
    d. antibiotic resistance

e. β-lactam

2. Define the following terms:

   a. open reading frame
   b. degenerate coding system
   c. nonoverlapping coding sequence
   d. codon
   e. anticodon

3. Define the following terms:

   a. wobble hypothesis
   b. aminoacyl-tRNA synthetase
   c. tRNA
   d. AUG sequence
   e. synonymous codon

4. Define the following terms:

   a. anhydride
   b. mixed anhydride
   c. tRNA linkage
   d. proofreading site
   e. translation termination

Fill in the Blanks

5. Four basic properties of the genetic code are specificity, degenerate, almost universal, and _____.
6. The chemical bond that links amino acids in a polypeptide is an _____.
7. The chemical bond involved between the nucleotides in a polynucleotide strand is a _____ .

Short-Answer Questions

8. What two observations prompted the wobble hypothesis?
9. Describe the two sequential reactions that occur in the active site of aminoacyl-tRNA synthetases.
10. Provide a DNA base sequence that could code for the following peptide:

    Ala-Ser-Phe-Tyr-Ser-Lys-Lys-Leu-Ala-Asp-Val-Ile

11. What is the mRNA base sequence for the peptide in Question 10?
12. What would be the effect of a single-base deletion in the codon for the second Ser residue in the DNA sequence that codes for the peptide in Question 10?
13. Determine the codon sequence for the peptide sequence glycylserylcysteinylarginylalanine. How many possibilities are there?
14. Why are tRNAs described as adaptor molecules?

Critical-Thinking Questions

15. In the first step of aminoacyl group linkage to a tRNA, an amino acid reacts with ATP to yield an aminoacyl-AMP and pyrophosphate. The bond between the amino acyl group and the AMP is described as a mixed anhydride. Explain the differences between an anhydride bond and a mixed anhydride bond. Why is this reaction irreversible?
16. Given an amino acid sequence for a polypeptide, can the base sequence for the mRNA that codes it be predicted?
17. Explain the significance of the following statement: The functioning of the aminoacyl-tRNA synthetases is referred to as the second genetic code.
18. Although aminoacyl-tRNA synthetases make few errors, occasionally an error does occur. How can these errors be detected and corrected?
19. Because of the structural similarity between isoleucine and valine, the aminoacyl-tRNA synthetases that link them to their respective tRNAs possess proofreading sites. Examine the structures of other α-amino acids and determine other sets of amino acids whose structural similarities might also require proofreading.
20. What factors ensure accuracy in protein synthesis? How does the level of accuracy usually attained in protein synthesis compare with that of replication or transcription?
21. Selenocysteine differs in the way that it is converted to its cognate charged tRNA. Explain.
22. The following pattern has been observed in the genetic code. For many codons, the first base specifies a biosynthetic precursor: U = pyruvate, C = α-ketoglutarate, A = oxaloacetate, G = any of a number of simple precursors. The second base of the codon tends to be associated with water solubility: water-soluble amino acids have a G, A, or C as the middle position, whereas five of seven of the most hydrophobic amino acids have U as the middle base. The third base in a codon is often information free; that is, many of the codons for the same amino acid differ only with the third base. Review amino acid biosynthesis (see Figures 5.2, 14.7, and 14.8) and determine which amino acids obey these rules. What are the exceptions?

SECTION 19.2

Comprehension Questions

23. Define the following terms:

    a. translation
    b. initiation
    c. initiator tRNA
    d. polysome
    e. elongation

24. Define the following terms:

    a. posttranslational modification
    b. transpeptidation
    c. translocation
    d. termination
    e. targeting

25. Define the following terms:

    a. peptidyl transferase center
    b. GTPase associated region

c. 70S bacterial ribosome
d. decoding center
e. Shine–Dalgarno sequence

26. Define the following terms:

a. guanine nucleotide exchange factor (GEF)
b. pretranslocation state
c. posttranslocation state
d. nascent polypeptide
e. releasing factors

27. Define the following terms:

a. ribosome recycling factor
b. trigger factor
c. signal peptide
d. stop transfer signal
e. ternary complex

28. Define the following terms:

a. chemotaxis
b. lac operon
c. 80S ribosome
d. mRNA secondary structure
e. mRNA scanning

29. Define the following terms:

a. cap-binding complex
b. 43S preinitiation complex
c. poly(A)-binding protein
d. 48S initiation complex
e. glycosylation

30. Define the following terms:

a. proteolytic cleavage
b. proproteins
c. preproproteins
d. hydroxylation
e. phosphorylation

31. Define the following terms:

a. PABP
b. lipophilic modification
c. methylation
d. Kozak sequence
e. carboxylation

32. Define the following terms:

a. disulfide bond formation
b. transcript localization

c. disulfide exchange
d. signal recognition particle
e. trapped ribosome

33. Define the following terms:

a. cotranslational transfer
b. SRP receptor protein
c. posttranslational translocation
d. SECIS
e. dolichol

34. Define the following terms:

a. pseudomembrane
b. CBC
c. context-dependent codon reassignment
d. KDEL
e. TOM complex

Fill in the Blanks

35. The three steps in protein synthesis are initiation, elongation, and _____.
36. Posttranslational modifications serve two purposes: to prepare a polypeptide for a specific function and _____.
37. A posttranslational modification that directs a polypeptide to a specific location is called a _____ modification.
38. Inactive precursor proteins with removable signal peptides are called _____.
39. The major differences between eukaryotic and prokaryotic translation are speed, location, complexity and variety of _____.
40. The _____ is an explanation for why there are fewer tRNAs than expected in living cells.
41. A _____ is a purine-rich sequence in close proximity to AUG on a prokaryotic mRNA that binds to a complementary sequence on the 30S ribosome subunits, thereby promoting the formation of the correct preinitiation complex.

Short-Answer Questions

42. What are the major differences between eukaryotic and prokaryotic translation?
43. What are the major prokaryotic translation control mechanisms?
44. Describe the steps in the elongation cycle.
45. Explain the differences among preproproteins, proproteins, and proteins.
46. Describe the structure and function of the signal recognition particle.
47. Describe the function of the translocon in cotranslational transfer.
48. Describe how eukaryotic mRNA structure can affect translational control.
49. In general terms, describe the intracellular processing of a typical glycoprotein that is destined for secretion from a cell.
50. What steps in the elongation cycle of protein synthesis require GTP hydrolysis? What role does it play in each step?
51. Name and explain the roles of the protein factors that participate in the initiation phase of

prokaryotic protein synthesis.

52. Compare the RNA and protein components of prokaryotic and eukaryotic ribosomes.
53. Explain the roles of large and small subunits of ribosomes.
54. The term *translation* refers to which of the following?

    a. DNA → RNA
    b. RNA → DNA
    c. proteins → RNA
    d. RNA → proteins

55. List and describe the major classes of eukaryotic posttranslational modifications.
56. Explain the importance of the proper targeting of nascent polypeptides.
57. Explain the critically important role of aminoacyl-tRNA synthetases in protein synthesis.
58. A peptide sequence is composed of 10 serine residues. Determine how many different mRNA sequences could code for the peptide.
59. Describe each step in the eukaryotic elongation process.
60. Describe the process whereby proteins are directed to their final destinations.
61. List the major translational control mechanisms of eukaryotes.
62. Determine the total amount of nucleotide bond energy that is required in the synthesis of the following tetrapeptide: Lys-Ala-Ser-Val.
63. Indicate the phase of protein synthesis during which each of the following processes occurs:

    a. A ribosomal subunit binds to a messenger RNA.
    b. The polypeptide is actually synthesized.
    c. The ribosome moves along the codon sequence.
    d. The ribosome dissociates into its subunits.

Critical-Thinking Questions

64. Eukaryotic protein synthesis is considerably slower than synthesis of prokaryotes. Explain.
65. Discuss the role of GTP in the functioning of translational factors.
66. What feature of eukaryotic translation is especially responsible for its efficiency?
67. What is the signal hypothesis?
68. What function do guanine nucleotide exchange factors serve in translation?
69. How do proteins targeted to the mitochondrial matrix reach their destination?
70. The three-dimensional structures of ribosomal RNA and ribosomal protein are remarkably similar among species. Suggest reasons for these similarities.
71. What advantages are there for synthesizing an inactive protein that must be subsequently activated by posttranslational modifications?
72. Posttranslational modifications serve several purposes. Discuss and give examples.
73. Describe how the base pairing between the Shine–Dalgarno sequence and the 30S subunit provides a mechanism for distinguishing a start codon from a methionine codon. What is the eukaryotic version of this mechanism?
74. What specific roles do translation factors play in both prokaryotic and eukaryotic translation processes?
75. Can you suggest a reason why ribosomes in all living organisms consist of two subunits and not one supramolecular complex?

SECTION 19.3

Comprehension Questions

76. Define the following terms:

    a. proteostasis network
    b. molecular chaperones
    c. heat shock response
    d. heat shock factors
    e. protein chip

Fill in the Blanks

77. The _____ is a series of pathways that protect cells from misfolded proteins.
78. The _____ protects cells from excessive heat.
79. _____ a proteomic technique that uses genetically modified yeast cells to detect protein-protein and protein-DNA interactions.

Short-Answer Questions

80. Name the misfolded protein associated with each of the following diseases: Alzheimer's, Parkinson's, and type 2 diabetes.
81. Describe the essential features of the heat shock response.
82. Describe the role of the proteostasis network.

Critical-Thinking Questions

83. A prokaryotic species is facing a new environmental stress that can be ameliorated by a catalytic activity that requires the side chain of a unique amino acid derivative called pyrolysine. How would such an organism develop a mechanism for incorporation of this nonstandard amino acid into an enzyme molecule? What would be the properties of the molecules required to solve this problem?

MCAT Study Questions

84. Protein synthesis requires sufficient energy to proceed. Which of the following is a major energy sensor and regulator of energy-requiring processes such as protein synthesis?

    a. CBC
    b. PABP
    c. mTORC1
    d. FOXP2

85. Which of the following steps in protein synthesis does **not** require a direct supply of energy?

    a. proofreading step by certain aminoacyl-tRNA synthetases
    b. translocation of mRNA in a ribosome
    c. linkage of an amino acid to its cognate tRNA
    d. alignment of a tRNA anticodon with an mRNA codon

86. What type of linkage is catalyzed by peptidyl transferase activity?

a. anhydride
b. mixed anhydride
c. amide
d. ester

87. A eukaryotic mRNA has a mutation that creates a premature STOP signal. Which of the following is involved in detecting this error?

a. CBC20/80
b. eIF1
c. 5S RNA
d. peptidyl transferase

88. Which of the following processes assists cells in maintaining proteostasis (high-protein folding quality control)?

a. preinitiation complex assembly
b. autophagy
c. nonstop-mediated polypeptide decay
d. polypeptide proofreading

Appendix

Chapter 1: End-of-Chapter Questions

3

metabolism

6

biochemical pathway

9

A few examples of life science fields that require a solid understanding of biochemistry include agronomy, forensics, marine biology, plant biology, pharmacology, plant or animal genetics, environmental science, and wildlife biology.

12

a. hydrocarbon: hydrophobic molecules composed of carbon and hydrogen atoms
b. hydrophilic: capable of hydrogen bonding with water and polar biomolecules
c. hydrophobic: incapable of hydrogen bonding with water and polar biomolecules
d. functional group: group of atoms within an organic molecule with distinct chemical properties
e. R group: group of atoms that make up a side chain in amino acids

15

a. sugar: a polyhydroxy aldehyde or ketone; the basic unit of carbohydrates
b. glucose: an aldohexose sugar
c. monosaccharide: a carbohydrate that consists of a single sugar molecule
d. polysaccharide: polymer of sugar molecules containing more than 20 monosaccharide units
e. cellulose: a glucose polymer with β (1,4) glycosidic bonds

18

a. DNA: deoxyribonucleic acid; a polynucleotide that is an organism’s genetic information
b. RNA: ribonucleic acid; a polynucleotide involved in protein synthesis or in diverse roles such as gene regulation
c. genome: an organism’s entire set of DNA sequences
d. transcription: process where RNA molecules are synthesized from a DNA template
e. fructose: a ketohexose sugar

21

coenzyme A

24

The functional groups in each molecule are:

a. aldehyde
b. carboxylic and amino groups
c. sulfhydryl groups
d. ester group
e. ester group
f. amide group
g. ketone group
h. alcohol group

See Student Study Guide, p. 10, Question 20.

27

Amino acids are components of peptides and proteins. Sugars occur in oligosaccharides and polysaccharides. Nucleotides are the components of nucleic acids. Fatty acids are components of several types of lipid molecules, e.g., triacylglycerols and phospholipids

30

a. The functions of fatty acids include energy storage and components of lipid molecules within cell membranes. Some fatty acids are also precursors to hormones.
b. Sugars function as energy sources and as components of polysaccharides (such as starch, cellulose, glycogen, and chitin). Nucleotides contain the sugars ribose or deoxyribose. Glycoproteins and glycolipids located on cell membranes also contain sugars and play critical roles in many cellular interactions.
c. Nucleotides are involved in energy transformations. They are also components of DNA and RNA.
d. Most amino acids are building blocks of proteins. Some have special functions as neurotransmitters or as precursors of other molecules (e.g., b-alanine is a precursor of pantothenic acid).

33

There are 20 standard amino acids. Using the equation X^n, where X = 20, and *n* (chain length) = 10, the number of possible decapeptides is 20^{10} or 1.024×10^{13}. To draw all of the possibilities at the rate of 1 every 5 minutes would require about 97,300,000 years.

36

a. nucleophile: an atom or group with an unshared pair of electrons that is involved in a displacement (nucleophilic substitution) reaction
b. electrophile: an electron deficient species
c. leaving group: the outgoing nucleophile that leaves with its electron pair
d. addition reaction: two molecules combine to form a single product
e. anhydride: a molecule containing two carbonyl groups linked through an oxygen atom

39

a. coenzyme: small molecules that function in association with enzymes as carriers of small molecular groups or electrons
b. anabolic pathway: small precursors are used to generate large complex molecules
c. catabolic pathway: large complex molecules that are degraded into smaller products
d. signal transduction pathway: a pathway that permits a cell to receive and respond to signals from its environment
e. glycolysis: a 10 reaction pathway that degrades glucose to two pyruvates to generate energy

42

autopoiesis

45

Both human-designed complex systems (such as machines or factories) and living systems require raw materials (nutrients) and energy to manufacture components; systems of both types also produce waste products and heat. Machines are designed to self-regulate upon receiving feedback from the environment (e.g., by monitoring temperature to determine heating or cooling needs). In contrast, living systems are self-sustaining; that is, they produce and repair all of their own structural and functional components and, via nucleic acids, they build machines (enzymes) that make the components. Even the nucleic acids themselves are reproduced by living systems. This level of self-sustainability is not present in human-designed complex systems.

48

Important ions found in living organisms are Na^+, K^+, Ca^{2+}, and Cl^-. Many polyatomic ions are also common, such as NH_4^+, PO_4^{3-}, and CO_3^{2-}.

51

NADH is the reducing agent and propionic acid is the oxidizing agent.

54

The smaller the pK_a value of a chemical group, the better it functions as a leaving group. In other words, its anion is more stable.

57

a. network: a group of interconnected molecules that perform one or more functions; a metabolic network consists of interconnected biochemical pathways that synthesize and degrade biomolecules
b. metabolic network: a network which consists of interconnected biochemical reaction pathways that synthesize and degrade biomolecules; reactant and product molecules connect these pathways to each other
c. signaling network: a network which is composed of receptor proteins that receive information and signaling pathways
d. module: a component or subsystem that performs specific functions
e. motif: a recurring regulatory circuit

60

system

63

The growth of cells is regulated by elaborate feedback mechanisms that regulate cell growth. Mutations in just a few of the control genes can disable this elaborate process and result in out-of-control cell growth.

66

The statement “the whole is more than the sum of its parts” refers to how many small parts collaborate to function as a whole. Understanding how the heart contracts to pump blood does not explain how all the cardiac cells work together to produce the contraction.

69

The ability of biological systems to have multiple functions is a strength of biological systems. In the case of the genetic code, 61 codons (triple base sequence) code for 20 different amino acids. Most amino acids have more than one codon. This example of degeneracy helps to minimize the effect of base substitution mutations in DNA.

72

c. i and ii

75

b. i, iii, and iv

Chapter 2: In-Chapter Questions

2.1

The volume of a prokaryotic cell is calculated as follows:

$$\pi r^2 h = 3.14 \times (0.5\ \mu\text{m})^2 \times 2\ \mu\text{m} = 1.57\ \mu\text{m}^3$$

The volume of a eukaryotic cell is calculated as follows:

$$4/3\ \pi r^3 = 4 \times (3.14 \times 10^3)/3 = 4200\ \mu\text{m}^3$$

By dividing the volume of the hepatocyte by the volume of the prokaryotic cell (4200 μm^3/1.57 μm^3) the number of prokaryotic cells that would fit within the heptocyte is obtained: 2700.

2.2

Without a means of disposal, the lipid molecules will accumulate in the cells. Cell function is eventually compromised and the cells die.

2.3

Ten percent of 70 kg is 7 kg, which can be expressed as 7×10^6 mg. To calculate the weight of a single “average” mitochondrion, divide the estimated total weight of mitochondria by the estimated number of mitochondria (1×10^{16}):

$$\frac{7 \times 10^6 \text{ mg}}{1 \times 10^{16}}$$

The answer (i.e., the weight of an average mitochondrion) is approximately

$$7 \times 10^{-10} \text{ mg, or } 7 \times 10^{-7} \text{ mg. } (1 \text{ mg} = 1 \times 10^{-6} \text{ g})$$

2.4

Cell division involves the highly organized restructuring of the microtubules that form the mitotic spindle during the phases of mitosis. Microtubule function depends on dynamic instability, that is, the capacity to rapidly shorten and lengthen via polymerization/depolymermerization reactions. Cell division, which is unregulated in cancerous cells, is suppressed by taxol because this drug stabilizes microtubule structure.

Chapter 2: End-of-Chapter Questions

3

a. polar head group: a hydrophilic charged or uncharged polar group in a phospholipid
b. hydrophobic tail: hydrophobic group such as the acyl group in the fatty acids of phospholipids
c. ligand: a molecule that binds to a protein or receptor
d. motor protein: proteins that utilize nucleotide hydrolysis to do work
e. GTP: guanosine triphosphate

6

a. proteotoxic stress: a potentially lethal condition in which there is an accumulation of misfolded proteins caused by genetic variations or environmental insults such as oxidative stress
b. proteosome: a large protein complexes that degrades proteins
c. proteome: the characteristic set of proteins produced by a cell
d. ubiquitin-proteosome system: a mechanism in which the multiprotein proteasome complex destroys proteins that are covalently bound to ubiquitin
e. autophagy: a process by which lysosomes degrade cellular debris

9

integral

12

The kinesins are motor proteins that move particles along the outer pair of microtubules of cilia and flagella toward the cell periphery. Dyneins move molecules in the opposite direction.

15

Examples of factors that promote protein misfolding include metabolic stress (e.g. illness), oxidative stress (oxygen radical formation), inflammatory signaling processes, and genetic factors (e.g. a gene mutation that results in the synthesis of a defective protein).

18

Multiple uses of antibiotics result in the loss of colonization resistance, the capacity of the digestive tract to resist being colonized by pathogenic organisms. Colonization resistance is difficult to

reestablish because dysbiosis has been established, i.e. some beneficial organisms have been replaced by pathogens in a process that creates a new less healthy form of colonization resistance.

21

Neurotransmitters, hormones, and cytokines are examples of eukaryotic signal molecules that allow organisms to process information. Neurotransmitters are products of neurons, hormones are products of glandular cells, and cytokines are products of white blood cells.

24

Frequent antibiotic use can result in dysbiosis, a condition in which allows the overgrowth of pathogens at the expense of beneficial gut bacteria. Some of these beneficial microorganisms protect the cells in the gut lining. The loss of beneficial gut flora promotes low-level chronic systemic inflammation, such as inflammatory bowel disease (IBS).

27

a. endotoxin: toxic bacterial molecule that leaks across the cell wall
b. periplasmic space: a bacterial structural layer that lies between the outer membrane and the plasma membrane
c. slime layer: disorganized accumulation of polysaccharides formed when microorganisms adhere to surfaces and grow
d. lipopolysaccharide: a toxic bacterial molecule that leaks across the gut wall; also called an endotoxin
e. pilus: a fine hair-like structure that allows cells to attach to food sources and host tissue

30

photosynthesis

33

peptidoglycan

36

Referring to Figure 2; draw a simple diagram of a bacterial cell. The nucleoid is the centrally located chromosome compartment. Plasmids are small circular DNA molecules separate from the chromosome. The cell wall in Figure 2.7 is composed of an outer membrane and a plasma membrane separated by the periplasmic space. Pili are fine hair-like structures that allow cells to attach to food sources and host tissues. Flagella are long flexible corkscrew-shaped filaments used for locomotion.

39

Bacterial DNA is contained in both the nucleoid (a spacious, irregularly shaped, centrally located region that contains long, circular DNA called a chromosome) and small circular DNA molecules called plasmids. These plasmids can replicate independently of the chromosome and be transferred between bacterial cells via conjugation (through pili). Plasmid DNA sequences that code for antibiotic resistance (via synthesis of a protein that inactivates the antibiotic) are readily transferred between bacteria. Bacterial cells that possess the plasmid grow and reproduce, while those cells that are susceptible to the antibiotic die.

42

Biofilms, also referred to as slime layers, are polysacchraride layers that accumulate with time and provide microorganisms with a protective barrier against antibiotic therapies and human immune system defenses.

45

Since the diameter of a spherical mycoplasma cell is 0.3 μm, the radius is 0.15 μm. The volume calculation of a spherical mycoplasma cell:

$$V = 4\pi r^3/3 = (4)(3.14)(0.15\ \mu\text{m})^3/3 = 0.014\ \mu\text{m}^3$$

Assuming that *E.coli* is cylindrical, with dimensions of 1 μm × 2 μm (i.e., a typical rod-shaped bacterium), its volume is

$$V = \pi r^2 h = (3.14)(0.5\ \mu\text{m})^2(2\mu\text{m}) = 1.6\ \mu\text{m}^3$$

At 0.014 μm^3, the volume of a typical mycoplasma is significantly smaller than that of the cylindrical *E.coli* cell. To be more specific, the mycoplasma is 0.9% of the size of an *E.coli* (i.e., 0.014 μm^3/1.6 μm^3). Alternatively, the *E.coli* is about 114 times larger than the mycoplasma (i.e., 1.6 μm^3/0.014 μm^3).

48

a. ER lumen: space enclosed by the ER membrane
b. rough ER: endoplasmic reticulum that has ribosomes on its cytoplasmic surface
c. smooth ER: endoplasmic reticulum that lacks ribosomes; involved in synthesis of lipid molecules and biotransformation reactions
d. clathrin: a protein complex that binds to membrane adaptor proteins to form a basketlike latticework that forces membrane into the shape of a bud
e. Golgi apparatus: stacks of membranous vesicles involved in the processing, packaging, and delivery of cell products to internal and external destinations

51

a. acid hydrolase: a class of digestive enzymes located in lysosomes
b. mitochondrion: organelle that is the site of aerobic respiration
c. lysosome: a vesicle that contain granules or aggregates of digestive enzymes
d. autophagy: a mechanism that destroys unnecessary or dysfunctional cell components
e. peroxisome: a small spherical membranous organelles that contain oxidative enzymes

54

ER tubule, ERMES (ER-mitochondria encounter structure)

57

nuclear lamina

60

The principal function of the rough endoplasmic reticulum is the synthesis of membrane proteins and protein for export from the cell. Smooth endoplasmic reticulum, so named because it lacks attached ribosomes, is involved in lipid synthesis and biotransformation processes.

63

Peroxisomes are small spherical membranous organelles that contain oxidative enzymes. Primary functions of peroxisomes are the generation and degradation of peroxides and oxidation of toxic molecules. Additional functions include the synthesis of certain membrane lipids and the degradation of fatty acids and purine bases. Peroxisome formation involves nuclear genes that code for the enzymes and membrane proteins. These molecules are synthesized on cytoplasmic ribosomes and then imported into preperoxisomes. The ER provides the peroxisomal membrane, and peroxins, a group of proteins required for peroxisome assembly. [Peroxisomes are also involved in photorespiration in plants (Chapter 13 of your text.)]

66

Eukaryotic cells can generate significantly more energy than can prokaryotic cells; hence they have a much more complex structure, as indicated by the compartmentation made possible by organelles. This vast increase in complexity requires a larger cell volume.

69

The components of the endomembrane system are the plasma membrane, endoplasmic reticulum, Golgi apparatus, nucleus, and lysosomes. All of these control transport of ions and molecules across its membrane. Each membrane encloses an internal space that requires such control to function properly, i.e., for key biochemical reactions to take place. The compartments of the endomembrane system are connected via membranous vesicles that bud off from a donor membrane in one component in the system and fuse with the membrane of another component. For example, proteins synthesized in the RER are transferred via vesicles to the Golgi apparatus for further processing reactions.

72

This defect could lie anywhere in the synthesis scheme for the LDL receptor. The high cholesterol levels result because cholesterol is not being transferred into the cell and is building up in the blood, thus contributing to severe atherosclerosis. The most common defects are improper insertion of the receptor into the plasma membrane or a defective receptor incapable of binding or internalizing LDLs.

75

Primary cilia contain the sensory components of the motile cilia but lack several of the motility component (i.e., the central microtubule pair, the dynein arms and the radial spokes within the axoneme). Also, a large number of receptors are present in the cilia membrane, thereby allowing the integration of numerous signaling processes.

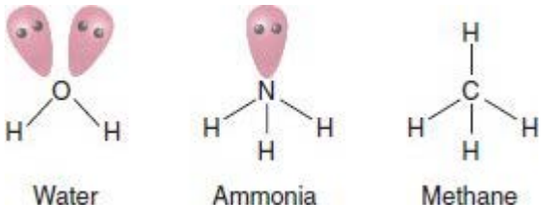
78

a. The nucleolus is a membrane-bound structure.

Chapter 3: In-Chapter Questions

3.1

The tetrahedral structures of the three molecules are as follows:



In the solid state of water, the oxygen atom has two electron pairs that form hydrogen bonds with neighboring water molecules. The nitrogen atom in ammonia has one unshared electron pair that can form a hydrogen bond with a neighboring ammonia molecule and methane has none. Note that the heats of fusion for these substances parallel the number of unshared electron pairs. Because of the ability of each ammonia molecule to form a hydrogen bond with a neighboring molecule, ammonia “ice” would be expected to be less dense than liquid ammonia.

3.2

From left to right in the illustration, the noncovalent interactions are ionic, hydrogen bonding, and van der Waals interactions.

3.3

Tendons and ligaments contain large amounts of collagen and other molecules that bind substantial amounts of structured water molecules. Water is an incompressible substance; that is, it cannot be forced to occupy a smaller space. As a result, structures containing large amounts of water can absorb relatively large amounts of force without damage.

3.4

The equilibrium shifts to the right to replace lost bicarbonate, and the acid concentration increases. The resulting condition is called acidosis.

Chapter 3: End-of-Chapter Questions

3

oxygen

6

104.5°

9

Water is an indispensable component for biological processes such as protein folding, biomolecular recognition, self-assembly of supramolecular structures, and gene expression.

12

a. salt bridges: form as a result of attraction between positively and negatively charged amino acid side chains in polypeptides
b. ionic interactions: occur between charged atoms or groups
c. van der Waals forces: relatively weak electrostatic interactions that arise when neutral permanent dipoles approach each other or an inducible dipole (such as a π cloud)
d. dipole–dipole interaction: occur between molecules containing electronegative atoms, causing molecules to orient themselves so that the positive end of one polar group is directed toward

the negative end of another

e. London dispersion forces: induced dipole-induced dipole interactions; a transient dipole in one molecule polarizes the electrons in a neighboring molecule

15

permanent

18

electrons

21

For hydrogen bonding to occur, the bond must be very polar (i.e., the atoms involved must have different electronegativity values). In the first three compounds, the electronegativity difference between hydrogen and the other atom (fluorine, oxygen, or nitrogen, respectively) is very different. The electronegativity of carbon and hydrogen is about the same; hence, the bond is nonpolar and the hydrogens are incapable of hydrogen bonding.

24

a. heat of fusion: the energy required to melt a solid
b. heat of vaporization: the energy required to vaporize one mole of a liquid at a pressure of one atmosphere
c. heat capacity: the energy that must be added or removed to change the temperature of a substance by one degree Celsius
d. boiling point of water: 100°C
e. freezing point of water: 0°C

27

heat of vaporization

30

The regular crystal lattice of the ice crystal is more open than the tightly hydrogen-bound liquid water. If ice were more dense than water, ice formed in lakes and oceans would sink to the bottom. Eventually, only a narrow layer at the surface would be liquid. This environmental condition is incompatible with life. Most aquatic life would not be able to survive.

33

a. solvation sphere: shells of water molecules that cluster around both positive and negative ions
b. amphipathic: a molecule that contains both polar and nonpolar groups
c. micelle: formed when amphipathic molecules are mixed with water with the polar surface exposed to water and the nonpolar surface internalized
d. hydrophobic effect: results because nonpolar molecules are attracted to each other by van der Waals forces and are unable to hydrogen bond with water, resulting in a water cage surrounding the nonpolar molecules
e. clathrate: a water-caged structure

36

a. hydrocarbon tail: a linear nonpolar molecular group composed of carbon and hydrogen atoms

b. polar head: the charged species in an amphiphilic molecule
c. dialyzing membrane: refers to a semipermeable that permits molecules other than water to cross it
d. van't Hoff factor: represents the degree of ionization of a solute in an osmolarity solution
e. osmometer: a device that measures osmotic pressure

39

solvation shells

42

colloid

45

The calculation of the molecular weight of the protein (g/mol) is as follows:

$$\pi = iMRT; \text{ where } \pi = 0.01 atm,$$
$$i = 1, R = 0.0821 L \cdot atm/mol \cdot K,$$
$$T = 298 K$$

Solving for M:

$$0.01 atm = (1)(0.0821 L \cdot atm/mol \cdot K)(298 K)(M)$$
$$M = 4.09 \times 10^{-4} mol/L$$

Solving for molecular weight of the protein:

$$0.056 g/0.030 L = 1.867 g/L$$
$$1.867 g/L = 4.09 \times 10^{-4} mol/L$$
$$1 mol \text{ of protein} = 4565.79 g = 4600 g$$

48

Lithium ions have a larger hydration sphere than sodium. Sodium does not easily diffuse into the cell because of its large hydration sphere. Lithium has an even larger hydration sphere and would have an even greater tendency to remain outside of the cell.

51

Sodium ions would have the highest concentration in the well because the hydrated volume of K^+ is much smaller than that of Na^+. The smaller hydrated ion will diffuse through the gelatin more readily than Na^+, thereby lowering the amount of K^+ remaining in the well.

54

The structure of cells is based on the phase separation of hydrophobic and hydrophilic substances. The function of the cell membrane is possible only because lipids are insoluble in water. If water dissolved every molecule, living organisms would not be able to create a barrier (membranes!) between themselves and their surroundings, and living organisms would not be possible.

57

The highly concentrated sugar solution pulls water out of any bacterial cells present, which kills them, thereby preserving the fruit.

60

a. hydronium ion: H_3O^+, a species formed in water when a proton combines with a water molecule
b. acid: a species that donates protons
c. base: a species that accepts protons
d. K_a: the acid dissociation constant; a quantitative measure of the strength of an acid
e. pK_a: the negative log of the dissociation constant, K_a

63

Le Chatelier's principle

66

hydronium ion, H_3O^+

69

Carbon dioxide is present in the blood in sufficient quantities to make it effective as a buffer and CO_2 levels can be regulated by physiological processes. Phosphate concentration in blood is too low for this compound to be an effective buffer. Within cells, the phosphate concentration is much higher, and it can therefore act as an effective buffer.

72

When 1 mL of 1 M HCl is added to 1 L of water, the new $[H^+]$ becomes:

$$(0.001\ \text{L}) \times (1\ \text{M}\ H^+) = 0.001\ \text{mol}\ H^+$$

$$(0.001\ \text{mol}\ H^+)/(1.001\ \text{L total}) = 9.99 \times 10^{-4}\ \text{M}\ H^+$$

$$\text{pH} = -\log\ [H^+] = -\log\ (9.99 \times 10^{-4}\ \text{M}\ H^+)$$

$$\text{pH} = 3$$

75

Magnesium, with a double positive charge, forms a strong hydration sphere. For Mg^{2+} to move into the structured water of macromolecules, its hydration sphere must be removed in a process that requires a large amount of energy. Chloride ion, on the other hand, has only a single negative charge and is a larger ion. Its hydration sphere is not as tightly held and it would take less energy to remove. As a result, chloride would more easily associate with macromolecules.

78

The blood is so highly buffered by the bicarbonate buffer and the large amounts of blood proteins that under normal physiological conditions the transport of weak acids in the blood does not appreciably change its pH. For example, in the presence of bicarbonate, any acid that ionizes produces carbon dioxide (which is exhaled). The pH of the blood then remains virtually unchanged.

81

c. 6.1×10^2 g/mol

Chapter 4: In-Chapter Questions

4.1

$$\Delta G' = \Delta G^{\circ\prime} + RT \text{ in } [ADP][P_i]/[ATP]$$

where $R = 8.315 \times 10^{-3}$ kJ/mol • K

$T = 310$ K

$[ADP] = 0.00135$ M, $[ATP] = 0.004$ M,

$[P_i] = 0.00465$ M

$\Delta G^{\circ\prime} = -30.5$ kJ/mol

$\Delta G' = -30.5$ kJ/mol + (8.315 J/mol • K)(310)

ln(0.00135 M)(0.00465 M)/(0.004 M)

$\Delta G' = -30.5 + 2.577$ (ln 0.00157)

$= -30.5 - 16.64$

$= -47.14$ kJ/mol $= -47.1$ kJ/mol

4.2

Amount of ATP required to walk a mile

= (100 kcal/mi)/7.3 kcal/mol

= 13.7 mol/mi × 507 g/mol = 6945.2 g/mi = 6950 g/mi

Amount of glucose required to produce 100 kcal through ATP

= (100 kcal)/(.04)(686 kcal/mol)

= 100 kcal/274.4 kcal/mol

= 0.36 mol

= 0.36 mol × 180 g/mol = 65.6 g of glucose

Chapter 4: End-of-Chapter Questions

3

a. endothermic reaction: a reaction that requires an energy input; $\Delta H > 0$
b. exothermic reaction: a reaction that releases heat; $\Delta H < 0$
c. isothermic reaction: a reaction that has no heat exchanged with the surroundings; $\Delta H = 0$
d. spontaneous process: reactions that occur with a release of energy
e. nonspontaneous process: reactions that do not occur without a constant input of energy

6

disorder

9

free energy

12

Although it is thermodynamically favorable for methane autoignition, this will not occur unless the temperature is very high (in this case about 600°C). Below this temperature the process is not kinetically favorable.

15

The first law of thermodynamics concerns the conservation of energy where energy cannot be created or destroyed. Energy can be transformed from one form to another.

$$\Delta E = q + w; \text{ where } \Delta E = \text{the change in energy of the system}$$
$$q = \text{the heat absorbed or released by the system}$$
$$w = \text{the work done by or to the system}$$

The second law of thermodynamics concerns the spontaneity of reactions where spontaneous reactions occur in the direction that increases the total disorder of the universe, $\Delta S_{univ} > 0$, where $\Delta S_{univ} = \Delta S_{surr} - \Delta S_{sys}$; S is entropy in the universe (univ), surroundings (surr), and system (sys).

18

Work is defined as a change in energy that produces a physical change. Physiological examples include the active transport across membranes of ions and molecules, biomolecule synthesis, and muscle contraction.

21

In the case of endothermic solutions, the enthalpy may be negative but the entropy is sufficiently positive to make the overall $\Delta G°$ favorable.

24

The balanced equation for the reaction is:

$$C_{17}H_{35}COOH + 26O_2 \rightarrow 18CO_2 + 18H_2O$$

The calculation of ΔH for this reaction is:

$$\begin{aligned}\Delta H &= \Delta H_{products} - \Delta H_{reactants} \\ &= 18 \text{ mol}(-94 \text{ kcal/mol}) + 18 \text{ mol}(-68.4 \text{ kcal/mol}) \\ &\quad - 1 \text{ mol}(-211.4 \text{ kcal/mol}) \\ &= -1692 \text{ kcal} - 1232 \text{ kcal} + 211.4 \text{ kcal} \\ &= -2711.8 \text{ kcal} = -2712 \text{ kcal}\end{aligned}$$

27

standard

30

Given that the ionization constant for acetic acid is 1.8×10^{-5}, the $\Delta G°$ for the reaction would be calculated as follows:

$$\Delta G = -RT\ln K_{eq}$$
$$= -(8.315 \text{ J/mol} \cdot \text{K})(298 \text{ K})(\ln(1.8 \times 10^{-5}))$$
$$= -27{,}071 \text{ J/mol} = -27.1 \text{ kJ/mol}$$

33

Under standard conditions, the following statements are true: b, d, and e.

36

At equilibrium: $\Delta G°' = -RT \ln K_{eq}$

$$13{,}800 \text{ J/mol} = -(8.314 \text{ J/mol} \cdot \text{K})(298 \text{ K}) \ln K_{eq}$$
$$5.57 = \ln K_{eq}$$
$$K_{eq} = 262$$
$$K_{eq} = [\text{glucose}][P_i]/[\text{glucose-6-phosphate}]$$
$$[\text{glucose-6-phosphate}] = 4 \text{ mM} = 4 \times 10^{-3} \text{ M}$$
$$262 = [\text{glucose}][P_i]/[4 \times 10^{-3} \text{ M}]$$
$$1.05 = [\text{glucose}][P_i]$$

Assuming that $[\text{glucose}] = [P_i]$, then $[P_i] = 1.02$ M

39

ΔG is the most useful criterion of spontaneity because it reflects the change in entropy, which must increase for a reaction to be spontaneous.

42

In the absence of Mg^{2+}, increased repulsion between adjacent negative charges of ATP would cause it to have less stability than ATP with Mg^{2+} present.

45

a. resonance stabilization: improvement in molecular stability due to presence of two or more alternative structures that differ only in the position of electrons
b. electrostatic repulsion: unfavorable interactions between species of like charge
c. phosphoanhydride bond: a high energy present in ADP and ATP; free energy is released when a phosphoanhydride bond is hydrolyzed
d. phosphodiester bond: a bond formed when two hydroxyl groups of phosphoric acid react with hydroxyl groups of other molecules to form two ester linkages

e. orthophosphate: an inorganic phosphate group; PO_4 or P_i

48

ATP

51

AMP hydrolysis involves cleavage of an ester bond and therefore releases the least energy. Hydrolysis of the other phosphate linkages involves the hydrolysis of either anhydride or an enol bond.

54

With an intermediate phosphoryl group transfer potential, ATP can accept a phosphate group from compounds that have a higher phosphoryl group transfer potential and transfer it to lower energy compounds. In other words, the number of compounds that can transfer a phosphate to or from ATP is maximized.

57

The energy liberated by the hydrolysis of 12.5 mol of ATP is

$$(12.5\ \text{mol})(-30.5\text{kJ/mol}) = -381.3\ \text{kJ}$$

The energy required to produce 12.5 mol of ATP is 1142.2 kJ. The apparent efficiency of the process is

$$(381.3/1142.2) \times 100\% = 33.4\ \%$$

60

Glucose-1-phosphate has a higher phosphoryl transfer potential than glucose-6-phosphate because it contains an anhydride bond, which is more reactive than the ester bond in glucose-6-phosphate.

63

b. the free energy value for the hydrolysis of ATP is greater than that for ADP

Chapter 5: In-Chapter Questions

5.1

Amino acids a and b are neutral, nonpolar, c is basic, and d is an acidic amino acid.

5.2

Bacteria with surface polypeptides composed of d-amino acids are resistant to degradation because the proteases, the enzymes that immune system cells use to degrade protein in foreign cells, can only catalyze the hydrolysis of peptide bonds between l-amino acids. In other words, the active sites of proteases are stereospecific; that is; they can only effectively bind peptides composed of l-amino acids.

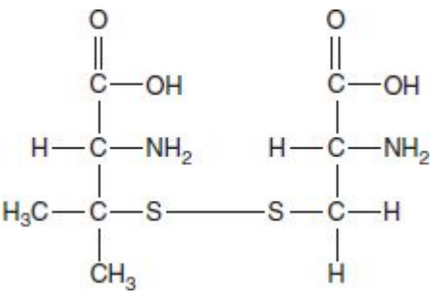
5.3

The isoelectric point for the tripeptide valylmethionyltryptophan is calculated as follows: The pK_a

values are (1) valyl amino group = 2.32; (2) tryptophan carboxyl group = 9.39. Since the methionine side chain cannot ionize, it has no pK_a value. The electrically neutral species of this tripeptide would contain a positively charged amino group and a negatively charged carboxylate group. The pI is calculated by adding the pKa values for the amino and carboxylate groups and dividing by two, i.e., 2.32 + 9.39/2 = 5.86. Since the pK_R for the cysteinyl side chain is more than two pH units away from the pI, it remains largely uncharged.

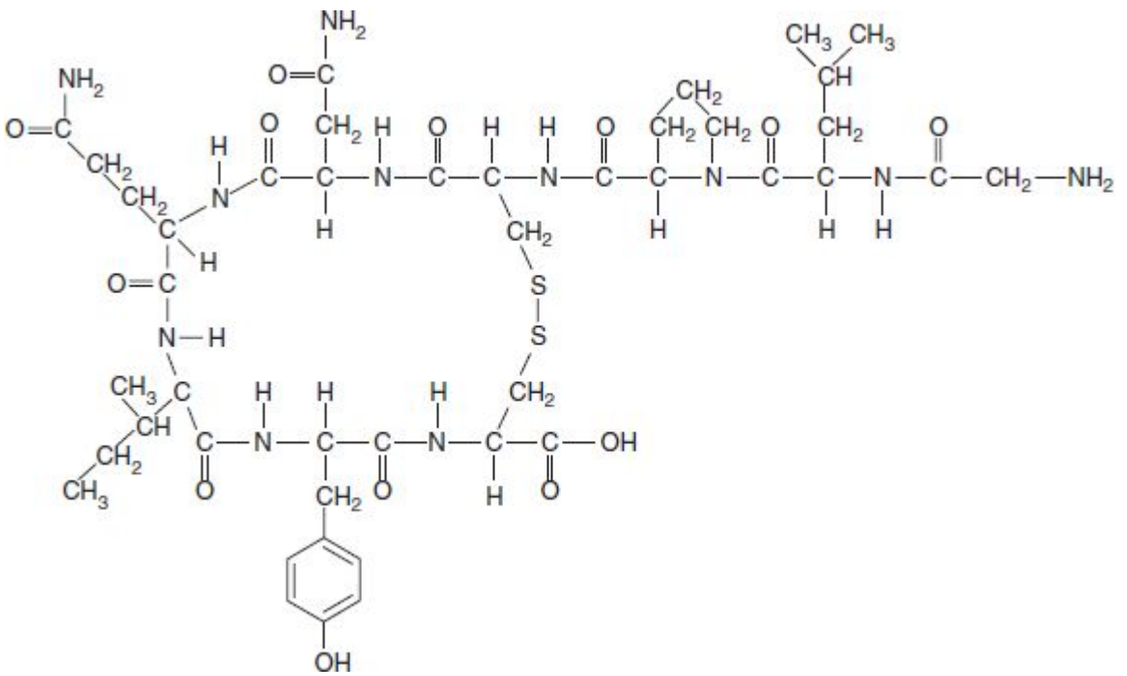
5.4

The structure of the penicillamine-cysteine disulfide is



5.5

The complete structure of oxytocin is



At pH 4 the terminal amino group of the glycine would be protonated to give the molecule a +1 charge. The isoelectric point of oxytocin is 5.6. Therefore at pH 9 the molecule will have a net negative charge. Atoms in oxytocin that can potentially form hydrogen bonds with water molecules include the carbonyl oxygens and the hydrogens in the peptide bonds. Other atoms include the amide side chains of glutamine and asparagine, the hydroxyl hydrogen of tyrosine, the N-terminal amino group and the C-terminal carboxyl group.

5.6

The partial overlap in the biological properties of vasopressin and oxytocin can be explained in part

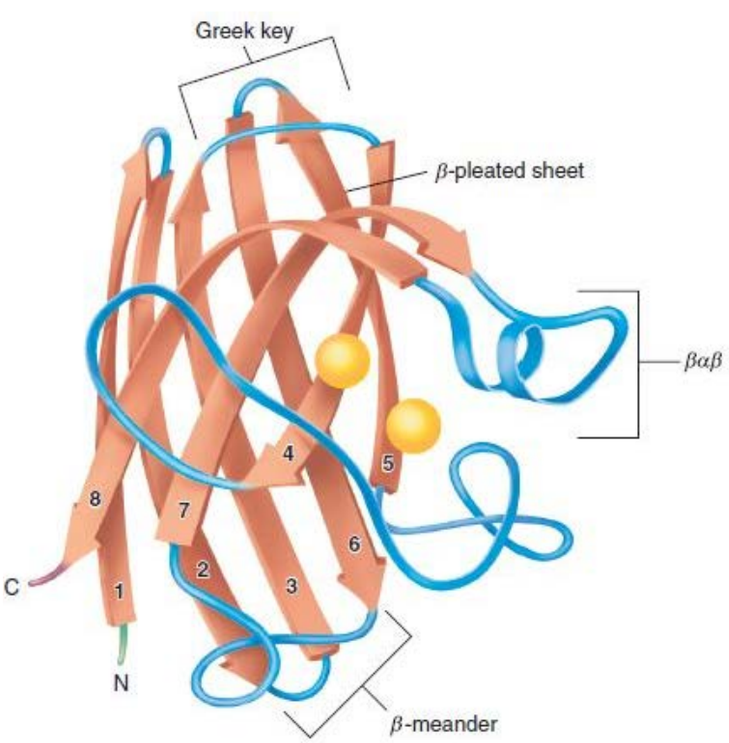
by the Ile residue at position 3 of oxytocin. Presumably, given the overall similarities in size and the identical nature of six of the eight amino acid residues in the two molecules, the hydrophobic side chain of this Ile can reach and partially fit into the hydrophobic pocket of the vasopressin receptor. The degree of functional overlap between the two peptides is reduced by the leucine residue at position 8 of oxytocin because the side chain of this residue is not only neutral but smaller than the positively charged arginine residue of vasopressin. If the arginine residue of vasopressin is replaced by a lysine residue, it is expected that there will be a decrease in the molecule's binding properties because of the structural differences between the two side chains. This decrease would probably not be large, since the side chains are similar in length and are both positively charged.

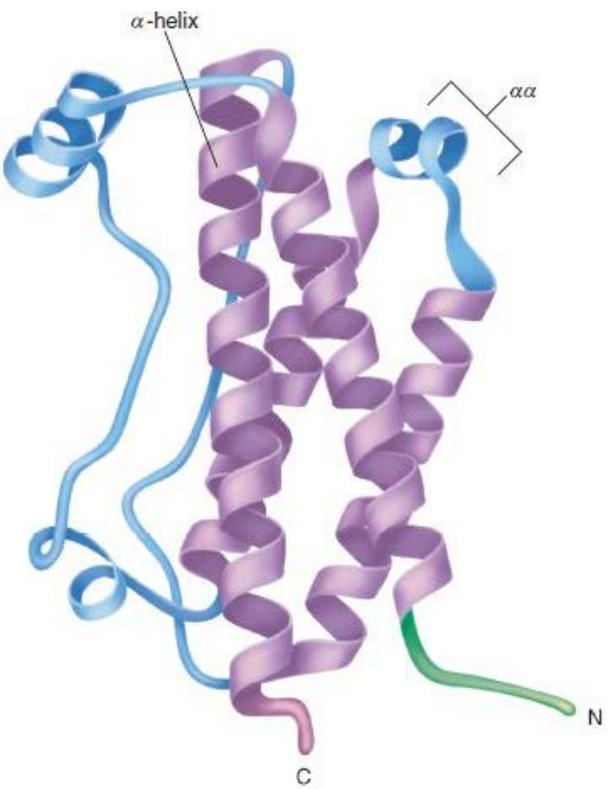
5.7

The trait is X-linked recessive. Since two copies of the aberrant gene are required for full expression of the disease in females, most symptomatic patients are males. Primaquine induces the production of excess amounts of the strong oxidizing agent hydrogen peroxide. In the absence of sufficient amounts of the reducing agent NADPH, the peroxide molecules cause extensive damage to the cell. No, a higher than normal peroxide level in blood cells is damaging to the malarial parasite and is selected for in geographical regions where malaria occurs.

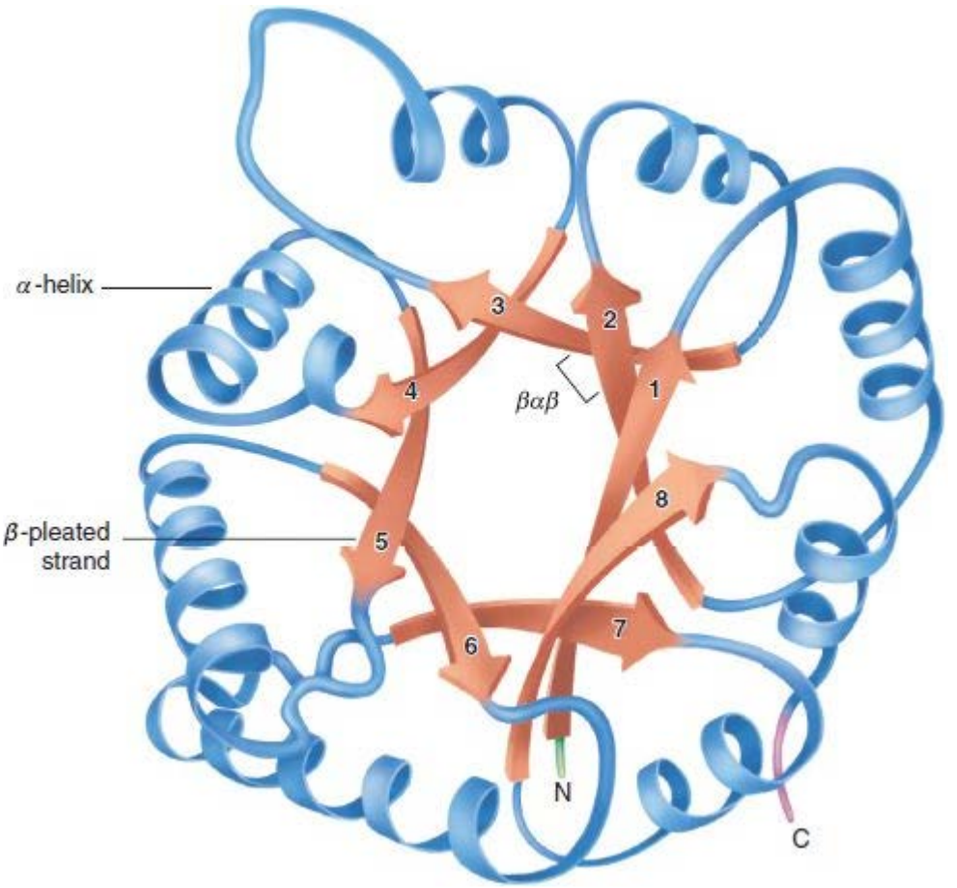
5.8

The solution to the first two illustrations is



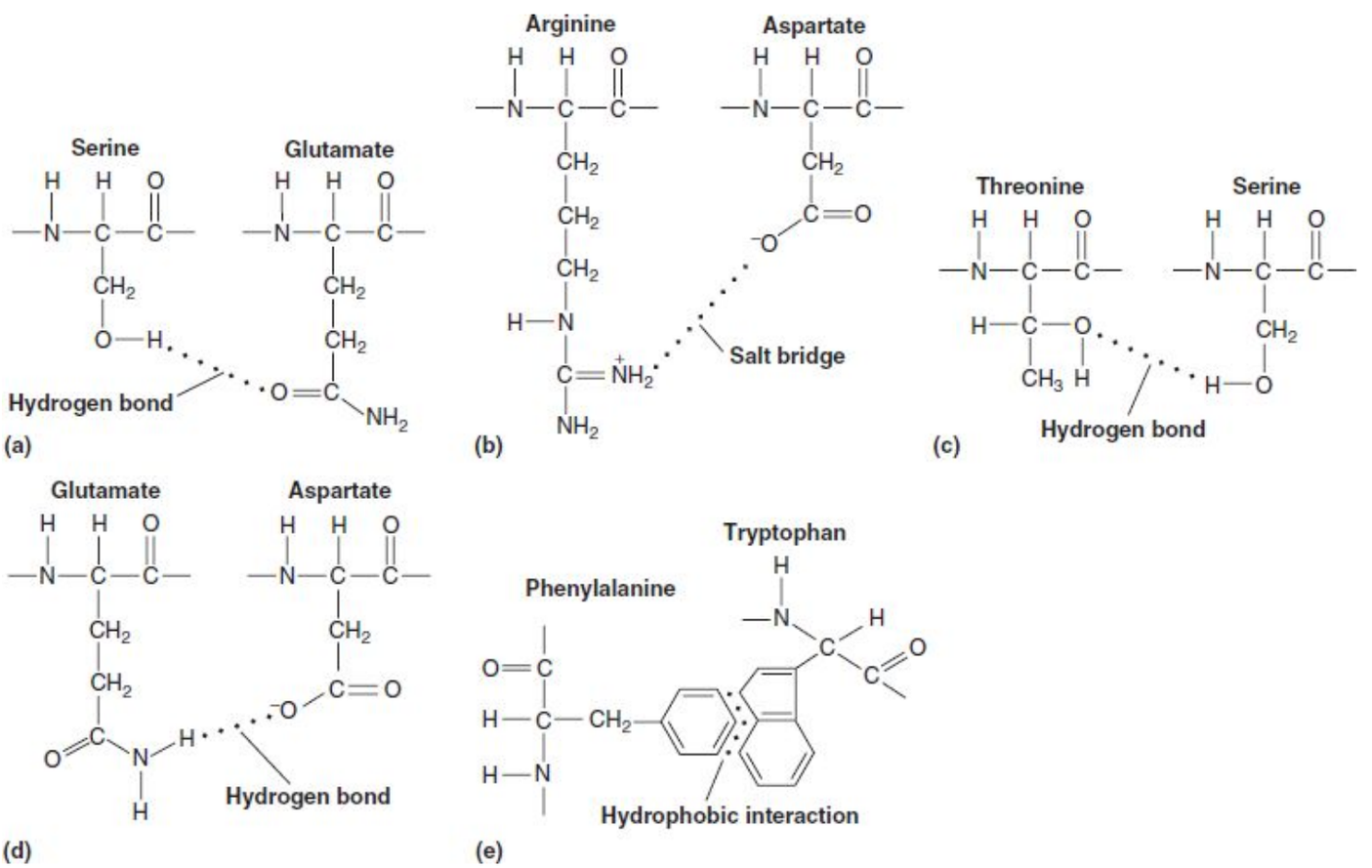


The solution to the third illustration is



This protein is an example of an α/β-barrel structure.

5.9



5.10

Collagen is a major structural protein found in connective tissues. Consequently, the failure of collagen molecules to form properly weakens these tissues causing diverse symptoms. Examples include cataracts, easily deformed bones, torn tendons and ligaments, and ruptured blood vessels.

5.11

BPG stabilizes deoxyhemoglobin. In the absence of BPG, oxyhemoglobin forms more easily. Fetal hemoglobin binds BPG poorly and, therefore, has a greater affinity for oxygen.

5.12

Myoglobin, composed of a single polypeptide, binds oxygen in a simple pattern—it binds the molecule tightly and releases it only when the cells' oxygen concentration is very low. The binding of oxygen by hemoglobin, a tetramer, has a more complicated pattern of oxygen-binding and release that is allosterically regulated. This mechanism is made possible by the noncovalent interactions among its four subunits.

Chapter 5: End-of-Chapter Questions

3

a. peptide bond: the amide bond formed between amino acids when the α-carboxyl group of one amino acid bonds with the amino group of another
b. Schiff base: imine products of amine groups that react with carbonyl groups
c. disulfide bridge: occurs when two cysteine amino acids form a disulfide bond
d. aldimine: Schiff bases formed by an amino group reacting with an aldehyde group
e. amphoteric: a molecule that can behave as an acid or a base

6

Tyrosine, cysteine, and asparagine are the amino acids with polar sidechains. The acidic amino acid glutamic acid and the basic amino acids lysine and histidine also possess side chains with electronegative atoms.

9

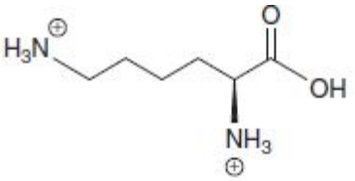
polypeptide

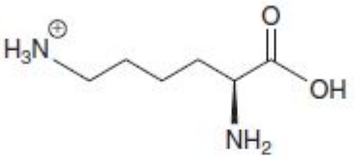
12

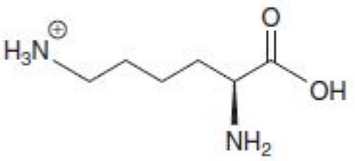
amphoteric

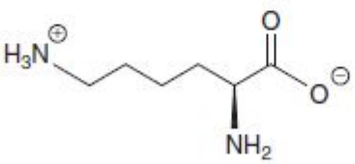
15

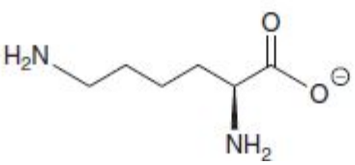
The structures and net charges of lysine as various pH values are as follows:

pH = 1  Net charge = +2

pH = 4  Net charge = +1

pH = 7  Net charge = +1

pH = 10  Net charge = 0

pH = 12  Net charge = −1

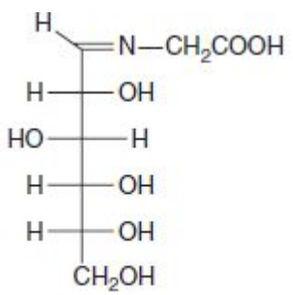
18

A polypeptide is a polymer containing more than 50 amino acid residues. A protein is composed of one or more polypeptide chains. A peptide is a polymer with fewer than 50 amino acid residues.

21

Cystinuria is a genetic disorder where excessive secretion of cystine into the urine results in crystallization/formation of calculi (kidney stones). Penicillamine is believed to be an effective treatment because of the formation of a penicillamine-cysteine disulfide bond, which allows the solubilization of excess cystine.

24



Adduct of Glucose and Glycine

27

a. peroxide: a molecule containing two oxygen atoms bonded together; R-O-O-R

b. GSSG: two oxidized glutathione molecules linked by disulfide bond
c. aquaporin: integral membrane protein that facilitates the passage of water through cell membranes; water channels
d. hypothalamus: a small structure located at the base of the brain near the pituitary gland that regulates a wide variety of functions including water balance, appetite, body temperature, and sleep.

30

glutathione

33

Blood pressure is influenced by both vasopressin (antidiuretic hormone) and atrial natriuretic factor (ANF). Vasopressin stimulates water reabsorption in the kidneys, via aquaporins, and blood pressure rises as more water flows into the bloodstream. ANF is produced in response to the stretching of the atrial wall due to high blood volume and/or high Na^+ concentrations. ANF increases Na^+ excretion, which promotes an increase in the excretion of water, which lowers blood pressure. ANF also inhibits the release of molecules that promote an increase in blood pressure (e.g., renin, aldosterone, and vasopressin). These agents act together, in addition to other peptides/biomolecules, to promote homeostasis and regulate systemic blood pressure.

36

Vasopressin's principle functions include antidiuretic activity and promoting parenting behavior and social bonding. Oxytocin stimulates uterine contractions and milk ejection from mammary glands, during and after birth, respectively. It also has modulating roles in social behaviors such as between mating pairs and mother and child. When a signaling molecule binds to its receptor it triggers a signaling cascade that causes a specific set of cell responses. Living organisms often use the same signaling molecules to cause different responses in different cells because it is more efficient than increasing the number of signal molecule-coding genes. Instead, living organisms alter the signaling cascades within different cells triggered by signal molecule-receptor binding. For example, the binding of insulin to its receptor in the liver and adipose tissue cells results in the synthesis of glycogen and triacylglycerol (fat), respectively.

39

a. metalloprotein: protein with a metal cofactor
b. primary structure: the amino acid sequence of a polypeptide as specified by genetic information
c. secondary structure: formation of certain localized arrangement of adjacent (but not necessarily contiguous) amino acids
d. tertiary structure: the overall three-dimensional shape assumed by a polypeptide
e. quaternary structure: the association of two or more folded polypeptides to form a functional protein

42

a. salt bridge: a strong electrostatic interaction that occurs in proteins between ionic groups of opposite charge
b. allosteric transition: ligand induced conformational changes in proteins
c. molecular disease: an illness caused by a deleterious mutation in a specific gene
d. protein denaturation: unregulated and irreversible disorganization of secondary, tertiary, or

quaternary structure

e. α-helix: a rigid, rod-like structure forming when a polypeptide twists into a right-handed helical conformation due to a repeating pattern of hydrogen bonds that form between the NH group and the carbonyl oxygen of the amino acid four residues away

45

transport

48

storage

51

a. heat: hydrogen bonding (secondary, tertiary, and quaternary structure)
b. strong acid: hydrogen bonding (secondary and tertiary structure) and salt bridges (secondary, tertiary, and quaternary structure)
c. saturated salt solution: salt bridges (tertiary structure)
d. organic solvents: hydrophobic interactions (tertiary and quaternary structure)

54

The primary driving force in protein folding is the requirement to achieve a low energy state despite the decrease in entropy that occurs as the protein's three-dimensional structure becomes more ordered. Key considerations include the energy associated with different bond angles and bond rotation, the chemical properties of the amino acid side chains (e.g., whether or not the side chain will be charged at cellular pH), and interactions between side chains. Of the noncovalent interactions that are possible, hydrophobic interactions are particularly important. (Recall that hydrophobic interactions are driven in part by the increase in entropy in the surrounding water molecules.)

57

In addition to mediating protein folding, molecular chaperones help protect newly synthesized proteins from inappropriate, premature protein-protein interactions by binding to and stabilizing proteins during early stages of folding. Molecular chaperones also direct the refolding of proteins that have partially unfolded; if refolding is not possible, they promote protein degradation. Ribosome-associated chaperones bind to a nascent polypeptide as it emerges from a ribosome. They prevent folding until an entire domain or polypeptide has exited. Hsp70s bind to, stabilize and promote the folding of nascent polypeptides and rescue misfolded and aggregated polypeptides. When a hydrophobic peptide segment binds within a hsp70 binding pocket, it is refolded as a result of ATP hydrolysis. Hsp90s finalize the folding of a limited, but diverse set of client proteins, in collaboration with hsp70. Hsp90s also facilitate the assembly of protein complexes such as RNA polymerase II. The chaperonins are large double-ring complexes that promote fast and efficient polypeptide refolding.

60

The folding of intrinsically disordered proteins (IDPs) into stable three-dimensional conformations is prevented by amino acid sequences that contain high percentages of polar and charged amino acids (e.g., Ser, Lys, and Glu). Protein folding is assisted by hydrophobic amino acids (e.g., Leu, Val, Phe, and Trp) as they strongly promote three-dimensional structure; hydrophobic residues are sequestered into the protein interior to minimize contact with water. Polar and charged amino

acids are solvated by water molecules and therefore are more randomly oriented with respect to the overall protein structure.

63

Living cells possess complex molecular chaperone mediated mechanisms that assist the proper folding of nascent polypeptides. These mechanisms are poorly understood and cannot yet be duplicated in the laboratory.

66

The large size of enzymes is required to stabilize the shape and functional properties of the active site and to shield it from extraneous molecules. In addition, structural features of the protein may function in recognition processes in signaling or binding to cellular structures.

69

During hyperventilation, CO_2 levels in the blood decrease, resulting in alkalosis (higher than normal pH). Since hemoglobin's release of O_2 is facilitated by H^+, less oxygen will be released by oxyhemoglobin.

72

c. 16 (256)

75

d. use of β-mercaptoethanol and 8 M urea at pH 8

Chapter 6: In-Chapter Questions

6.1

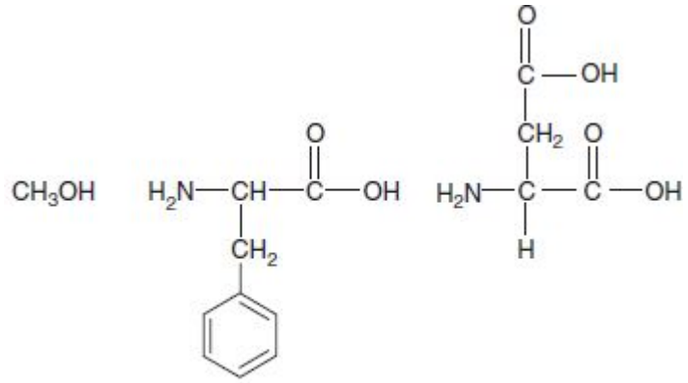
The amino acid residues forming the three-dimensional structure of the active site are chiral. As a result, the active site is chiral. It can bind only one isomeric form of a hexose sugar, in this case the d-isomer.

6.2

a. isomerase
b. transferase
c. lyase
d. oxidoreductase
e. ligase
f. hydrolase

6.3

The products of the degradation are the following compounds:



Cleavage of the ester bond is catalyzed by an esterase; the amide bond is cleaved by a peptidase.

6.4

Menkes' syndrome—injections of copper salts into the blood would avoid intestinal malabsorption and provide the copper necessary to form adequate levels of ceruloplasmin and offset the symptoms of the disease.

6.5

Wilson's disease—zinc induces the synthesis of metallothionein, which has a high affinity for copper. Some organ damage can be averted because metallothionein sequesters copper and prevents this toxic metal from binding to and inactivating susceptible proteins and enzymes. Penicillamine forms a complex with copper in the blood. This complex is transported to the kidneys, where it is excreted.

6.6

a. cofactor
b. holoenzyme
c. apoenzyme
d. coenzyme
e. coenzyme

6.7

6.8

Dialysis removes the formaldehyde, formic acid, and methanol that build up in the bloodstream. The bicarbonate neutralizes the acid produced and helps offset the resultant acidosis. The ethanol competitively binds with the alcohol dehydrogenase. This slows the dehydrogenation of the methanol and allows time for the kidneys to excrete it.

6.9

The patient that failed to show improvement probably had a higher level of acetylating enzymes. The patient's dosage should be based on capacity to process the drug and not on body weight.

Chapter 6: End-of-Chapter Questions

3

a. PTGS: prostaglandin endoperoxide synthase; the enzyme that is inhibited by aspirin
b. COX: cyclooxygenase; the enzymatic activity in PTGS that converts arachidonic acid to yield a precursor of the eicosanoids
c. nociceptors: receptors in the bare nerve endings in primary sensory neurons
d. lock-and-key model: enzyme model where active site and substrate have complementary structures
e. induced-fit model: enzyme model where noncovalent interactions between active site and substrate change the shape of active site so it can bind substrate and initiate reaction

6

transition state

9

hexokinases

12

Organisms may control enzyme activities directly, principally through the binding of activators or inhibitors, and the covalent modification of enzyme molecule.

15

Enzymes are composed of chiral amino-acid residues, a circumstance that creates a unique spatial arrangement within active sites. One stereoisomer has a complementary spatial arrangement within the enzyme's active site and therefore can enter the site, where it undergoes the reaction. Its enantiomer will not fit properly into the active site. As a result, there is no reaction.

18

a. oxidoreductase: an enzyme that catalyzes an oxidation-reduction reaction
b. lyase: an enzyme that catalyzes an addition or elimination reaction, adding a molecule across a double bond or removing a molecule to form a double bond
c. ligase: an enzyme that catalyzes the joining of two molecules, often using ATP for energy

21

transferases

24

a. cytochrome $P_{450:}$ one of a large enzyme class that processes toxic substances in animals
b. enzyme cofactor: the nonprotein component of an enzyme (either an inorganic ion or a coenzyme) required for catalysis
c. coenzyme: a small organic molecule required in the catalytic mechanisms of certain enzymes
d. Lewis acid: an electron pair acceptor
e. pH optimum: the pH value at which an enzyme's activity is maximal

27

coenzymes

30

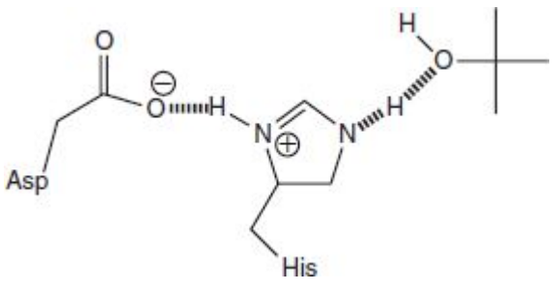
Transition metals can act as Lewis acids because they can accept electrons; they are effective electrophiles in chemical reactions.

33

Unprotonated side chains of Asp, Glu. Tyr, Cys and Lys can act as general bases. If the side chains if these amino acids are protonated they can act as general acids. Histidine is one of the most important amino acids in acid/base catalysis because the pKa of the imidazole group of histidine is approximately 6. Therefore, the histidine side chain ionizes within the physiological pH range. Its protonated form is a general acid and the unprotonated form is a general base.

36

The histidine protonates the OH of the alcohol making it a better leaving group.



39

The functional properties of aspartate–histidine dyads result from a polarized imidazolium ring caused by the close proximity of the aspartate's negatively charged carboxylate group. The polarization of histidine by the aspartate residue allows the imidazolium ring to become a stronger base.

42

a. Lineweaver–Burk plot: a graph is which K_m and V_{max} values for an enzyme are determined using the reciprocals of initial velocities and substrate concentrations
b. allosteric enzyme: an enzyme whose activity is affected by the binding of effector molecules
c. macromolecular crowding: the dense packing of an enormous variety of macromolecules and other molecules within the interior of cells
d. zymogen: an inactive form of a proteolytic enzyme
e. metabolon: a complex of enzymes that share intermediates of a metabolic pathway so that the product of one enzyme is in close proximity to the active site of the next enzyme in the pathway

45

inhibitors

48

competitive

51

Negative feedback inhibition is a process in which the product of a metabolic pathway inhibits the activity of an enzyme at/near the start of the pathway or a branch point. This type of mechanism allows pathways to be carefully controlled and the concentrations of metabolites to be highly regulated.

54

The two major types of enzyme inhibitors are reversible and irreversible. Reversible inhibition does not destroy the enzyme. Adding substrate or removing the inhibitor can remove the inhibitory effect. Malonate, whose structure resembles that of succinate, is a reversible competitor of succinate dehydrogenase. An irreversible inhibitor binds permanently (usually by covalent bond) to the enzyme and destroys its catalytic activity. Removing irreversible inhibitors will not restore enzyme activity. Iodoacetate is an irreversible inhibitor that alkylates a side chain with the active site of glyceraldehyde-3-phosphate dehydrogenase.

57

Since the enzyme is unchanged during the course of the reaction it does not appear in the rate equation and hence does not affect the order of the reaction.

60

The data indicates that the reaction is first-order in pyruvate and ADP and second order in P_i as the reaction rate = $k[\text{Pyruvate}]^1[\text{ADP}]^1[P_i]^2$. Therefore, the overall reaction is fourth order.

63

Since the K_m values are essentially the same and the *V*max values change between the two reactions, the addition of the unknown compound acts as a noncompetitive inhibitor. K_m values are close enough.

66

Inhibitor 1 exhibits competitive inhibition (K_m increases while V_{max} is unaffected) while Inhibitor 2 exhibits uncompetitive inhibition (both K_m and V_{max} are reduced). Inhibitor 1 demonstrates an affinity for the enzyme (E) while Inhibitor 2, being uncompetitive, binds to the enzyme-substrate complex (ES) and E.

69

a. No
b. If α is less than 1, the reaction would be faster than the uninhibited reaction.

72

Alcohol dehydrogenase selectively metabolizes ethyl alcohol in preference to ethylene glycol. Treatment of the patient with ethyl alcohol will block conversion to ethylene glycol to toxic metabolites. The ethylene glycol can then be excreted.

75

a. allosteric regulation: the regulation of enzymatic activity through binding of effector molecules
b. homotropic effects: allosteric regulation where the modulating ligand is the same as the substrate
c. heterotropic effects: allosteric regulation where the modulating ligand is different than the

substrate

d. concerted model: enzymatic model in which enzyme exists in only two states: taut and relaxed. Activators stabilize the relaxed conformation, while inhibitors stabilize the taut conformation
e. sequential model: enzymatic model in which binding of ligand to one subunit triggers sequential conformational change in all other subunits

78

heterotropic

81

Compartmentation within eukaryotic cells is the physical separation of enzymes by a membrane (i.e. by containing certain enzymes within an organelle), or by attachment of enzymes to membranes or cytoskeletal filaments. Compartmentation (1) prevents competing reactions from occurring simultaneously and allows them to be regulated separately ("divide" and "control"), (2) reduces or removes diffusion barriers by locating enzymes and metabolites close to each other, (3) provides specialized reaction conditions (e.g., low pH) that would not be possible otherwise, and (4) protects other cellular components from potentially toxic reaction products ("damage control").

84

Alcohol dehydrogenase (ADH), a zinc-containing tetramer converts acetaldehyade into ethanol in *S. cerevisiae*. Unlike most fermenting organisms, yeasts are protected by 250 genes that code for molecules that protect against ethanol toxicity. Humans can consume moderate amounts of toxic ethanol molecules because of the detoxifying reaction catalyzed by the liver ADH isoenzymes.

87

Prominent examples of substrates for snake venom enzymes include various proteins and membrane and other lipid components. Snakes are protected from the ill effects of their own venom enzymes by specific inhibitors circulating in their blood.

91

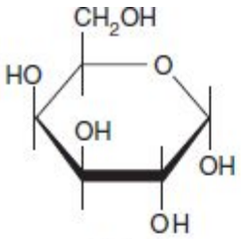
a. coenzymes

Chapter 7: In-Chapter Questions

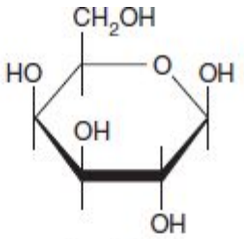
7.1

a. Aldotetrose b. Ketopentose c. Ketohexose

7.2

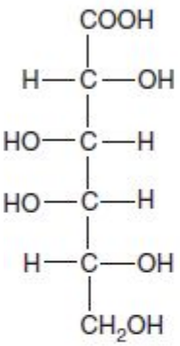


α-D-Galactose

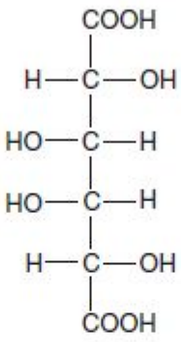


β-D-Galactose

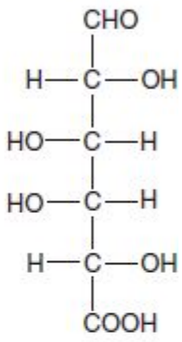
(a)



Aldonic acid

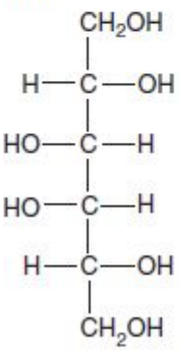


Aldaric acid



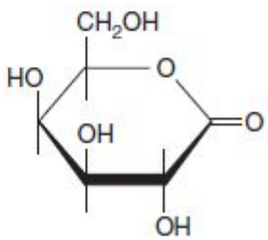
Uronic acid

(b)



Galactitol

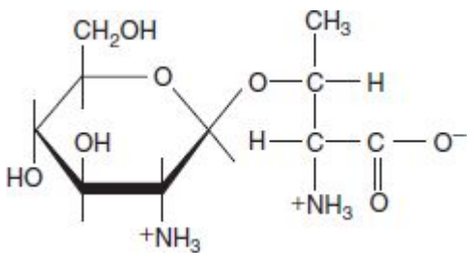
(c)



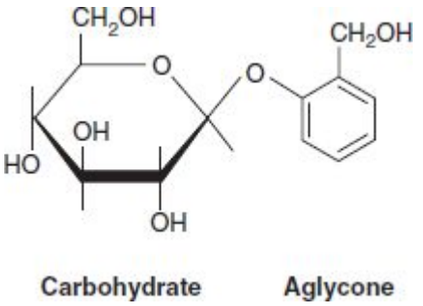
δ-Lactone of galactonic acid

(d)

7.3



7.4



7.5

a. maltose-reducing sugar
b. fructose-reducing sugar
c. α-methyl-D-glucoside: nonreducing
d. sucrose: nonreducing

Sugars a and b are capable of mutarotation.

7.6

The larger, insoluble glycogen molecule makes a negligible contribution to the osmotic pressure of the cell. In contrast, each molecule of an equivalent number of glucose molecules contributes to osmotic pressure. If the glucose molecules were not linked to form glycogen, the cell would burst.

7.7

In an analog system, information is encoded as a continuous signal. For example, in old-fashioned clocks the hands move continuously and not in small steps. In analog systems, small fluctuations in information processing can be meaningful. The sugar code is an analog system because of micro-heterogeneity, the continuous spectrum of carbohydrate structures that cells can synthesize to encode biologically relevant signaling information. In a digital system, information is represented by discrete values ("digits") of a physical quantity. Digital clocks display time as a progression of discrete numbers with no intermediate values. The DNA code is a digital system composed of four digits (bases). During protein synthesis, each DNA base triplet that ultimately codes for an amino acid has a specific meaning.

Chapter 7: End-of-Chapter Questions

3

a. reducing sugar: a sugar that can be oxidized by weak oxidizing agents
b. alditol: a sugar alcohol; the product of the reduction of the aldehyde or ketone group of a monosaccharide
c. enediol: the intermediate formed during the isomerization reactions of monosaccharides; contains a double bond with a hydroxyl group on each carbon of the double bond
d. acetal: one of a family of organic compounds with the general formula $RCH(OR')_2$; formed from the reaction of a hemiacetal with an alcohol
e. ketal: one of a family of organic compounds with the general formula $RRC(OR')_2$; formed from the reaction of a hemiketal with an alcohol

6

a. epimer: diastereomers that differ in the configuration at a single asymmetric carbon atom
b. acetal linkage: an arrangement where two organic molecules are connected through a –O–C(H)–O– (acetal) functional group
c. reducing sugar: a sugar that can be oxidized by weak oxidizing agents
d. monosaccharide: a polyhydroxy aldehyde or ketone containing at least three carbon atoms
e. anomer: an isomer of a cyclic sugar that differs from another in its configuration about the hemiacetal or acetal carbon
f. diastereomer: a stereoisomer that is not an enantiomer (mirror-image isomer)

9

epimers

12

alditol

15

sorbitol

18

The following sugar pairs are classified as enantiomers, diastereomers, epimers, or an aldose–ketose pair.

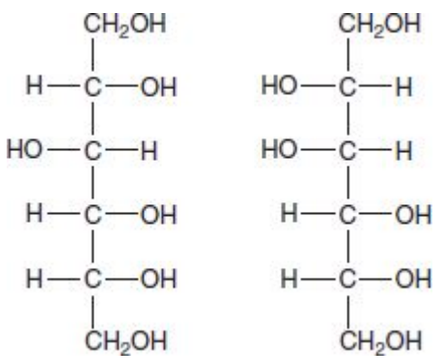
a. D-erythrose and D-threose are epimers
b. D-glucose and D-mannose are epimers
c. D-ribose and L-ribose are enantiomers
d. D-allose and D-galactose are diastereomers
e. D-glyceraldehyde and dihydroxyacetone are an aldose-ketose pair

21

Because the test for reducing sugars requires an aldehyde group, all aldoses are reducing sugars because they contain the required aldehyde. Ketoses give a positive test because they are converted to aldoses in the alkaline-reducing sugar test solution.

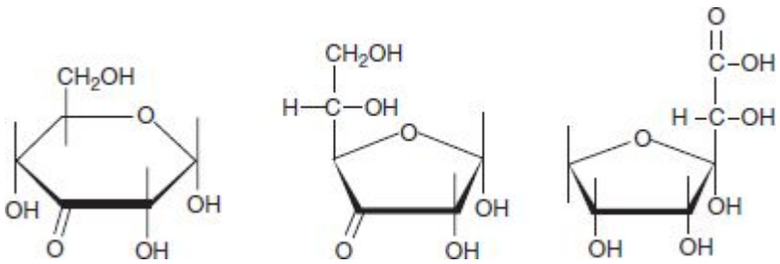
24

The two diastereomeric products of the reduction of fructose are:



27

Three α-anomeric ring forms of 3-ketoglucose are shown. The pyranose form is the most stable because the bond angles are less strained.



30

a. enterocyte: one of the numerous cells that line the small intestine; absorbs digested nutrients such as monosaccharides
b. N-glycan: an oligosaccharide linked to a protein via a *β*-glycosidic bond between the core N-acetylglucosamine anomeric carbon and a side chain amide nitrogen of an asparagine residue
c. O-glycan: an oligosaccharide linked to a protein via an α-glycosidic bond to the hydroxyl oxygen of serine or threonine residues; a mucin-type polysaccharide
d. glycogen: energy storage molecule in animals composed of glucose residues linked by *α*(1,4) and *α*(1,6) linkages
e. glycosaminoglycan: a long unbranched heteropolysaccharide chain composed of disaccharide repeating units

33

heteroglycans

36

polysaccharides

39

Starch and glycogen are both homoglycans containing glucose monomers linked by α- (1,4) glycosidic bonds with branch points connected by *α*-(1,6) glycosidic bonds. Glycogen, however, is much more highly branched than starch. Cellulose is a linear polymer of glucose linked by *β*-(1,4) glycosidic bonds.

42

Glycosaminoglycans contain amino and sulfate groups, both of which have ionic charges attract water molecules. In addition, the large numbers of polar hydroxyl groups of glycoaminoglycans participate in hydrogen bonding with water molecules.

45

Glucuronic acid possesses both a carboxylic acid group and several alcohol functional groups, both of which improve water solubility via hydrogen bonding.

48

In order for the sugar to undergo mutarotation there must be a hemiacetal or hemiketal as part of the structure. Sucrose’s anomeric carbons are linked in a full acetal.

51

a. asparagine-linked oligosaccharide: an N-linked heteroglycan; oligosaccharides linked to

proteins through asparagine residues

b. mucin-type oligosaccharide: an O-linked heteroglycan

c. Na^+-K^+-ATPase: an integral membrane glycoprotein; ion pump found in the plasma membrane of animal cells

d. major histocompatibility antigens: integral membrane glycoproteins; cell-surface markers used to cross-match organ donors and recipients

e. Hurler's syndrome: genetic disease associated with proteoglycan metabolism; deficiency of a specific lysosomal enzyme causes excessive accumulation of dermatan sulfate

54

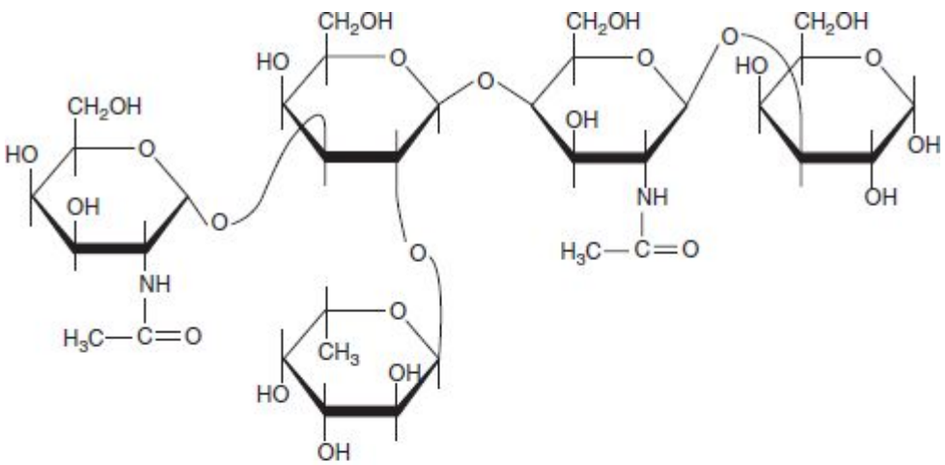
syndecans, glypians

57

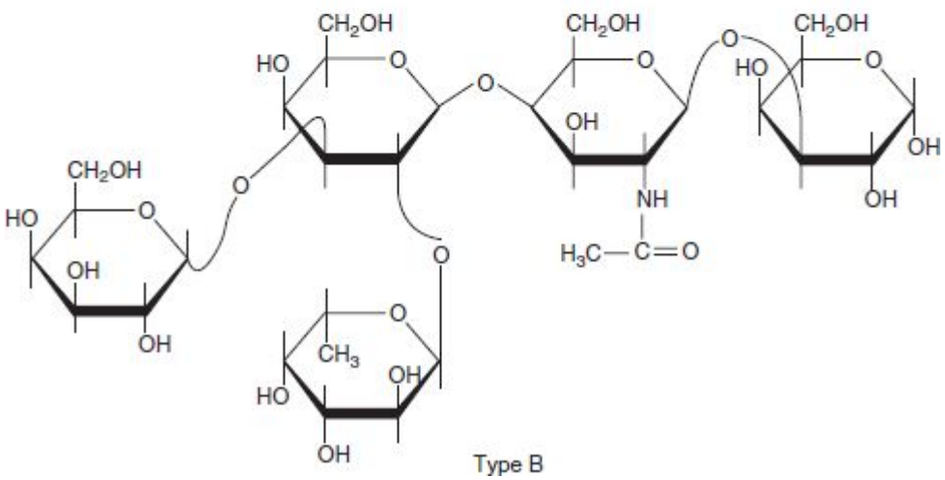
Chondroitin sulfate and proteoglycans are extensively negatively charged at physiological pH and as such are spread out, binding large amounts of water. The interwoven chains block the passage of large molecules. Smaller molecules can pass between the chains.

60

The complete structures of the Type A and Type B antigens are as follows:



Type A



Type B

63

a. sugar code: the coding capacity of proteins is significantly enhanced by the covalent attachment of carbohydrate groups
b. lectin: a carbohydrate binding protein
c. glycoform: one of several slightly different forms of a glycan component of a glycoprotein
d. glycome: the total set of sugars and glycans that a cell or organism produces
e. microheterogeneity: variations in the glycan components of each type of glycoprotein that cells use to generate cell- or tissue-specific signal transduction ligands

66

20,000 (scientists expected 100,000 protein coding genes. See Chapter 17.)

69

The damaging effects of the cholera toxin occurs only after endocytosis into the host cell, a process that is initiated by lectin-ligand binding. The binding of the B subunit of cholera toxin to a glycolipid on the surface of intestinal cells results in the uptake of the toxic A subunit. Once internalized, the A subunit proceeds to disrupt the mechanism that regulates chloride transport, a process that results in a life-threatening diarrhea.

72

During embryonic development, complex genetic programming changes in which genes are turned on and off as totipotent cells undergo a large number of cell divisions which eventually give rise to the differentiated cells of organs and tissues. Cell surface glycoconjugate pattern modifications caused by alterations in the transcription of genes such as those coding for glycosylation enzymes play a vital and integral role in the developmental processes. The cell-cell and cell-signal molecule interactions mediated by glycoconjugates direct cells into the different pathways leading to the cell differentiation and specialization of a multicellular organism.

75

a. sucrose

Chapter 8: In-Chapter Questions

8.1

The large excess of NADH that is produced by these reactions drives the conversion of pyruvate to lactate.

8.2

Chromium is acting as a cofactor.

8.3

In the absence of O_2, energy is produced only through glycolysis, an anaerobic process. Glycolysis produces less energy per glucose molecule than does aerobic respiration. Consequently, more glucose molecules must be metabolized to meet the energy needs of the cell. When O_2 is present, the flux of glucose through glycolysis is reduced.

8.4

At three strategic points, glycolytic and gluconeogenic reactions are catalyzed by different

enzymes. For example, phosphofructokinase and fructose-l,6-diphosphatase catalyze opposing reactions. If both reactions occur simultaneously (i.e., in a futile cycle) to a significant extent, ATP hydrolysis in the reaction catalyzed by phosphofructokinase releases large amounts of heat. If the heat is not quickly dissipated, an affected individual could die of hyperthermia.

8.5

In gluconeogenesis pyruvate is converted to oxaloacetate. NADH and H^+ are required to reduce glycerate 1,3-bisphosphate to glyceraldehyde-3-phosphate. NAD^+ is the oxidized form of NADH also produced in this reaction. ATP is needed to provide the energy to carboxylate pyruvate to oxaloacetate and phosphorylate glyceraldehyde-3-phosphate to glycerate-1,3-bisphosphate. Both of these reactions also produce ADP and P_i. GTP is hydrolyzed in the conversion of oxaloacetate to phosphoenolpyruvate to yield GDP and P_i. Water is involved in the hydrolysis reactions of ATP to ADP and P_i, the conversion of phosphoenolpyruvate to 2-phosphoglycerate, and the hydrolysis of glucose-6-phosphate to glucose. Six protons are formed when four molecules of ATP and two molecules of GTP are hydrolyzed.

8.6

Without glucose-6-phosphatase activity, the individual cannot release glucose into the blood. Blood glucose levels must be maintained by frequent consumption of carbohydrate. Excess glucose-6-phosphate is converted to pyruvate, which is then reduced by NADH to form lactate.

8.7

The enzyme deficiencies prevent the breakdown of glycogen. Because the synthetic enzymes are active, some glycogen continues to be produced and causes liver enlargement. Because of the liver's strategic role in maintaining blood glucose, defective debranching enzyme causes hypoglycemia (low blood sugar).

Chapter 8: End-of-Chapter Questions

3

a. aldolase: the reversible conversion of fructose-1, 6-bisphosphate to DHAP and glyceraldehyde-3phosphate
b. GADPH: an important intracellular reducing agent; the reduced form of $NADP^+$, required for reductive processes (e.g. lipid biosynthesis) and antioxidant mechanisms (NADPH is a powerful antioxidant)
c. substrate-level phosphorylation: the synthesis of ATP from ADP by phosphorylation coupled with the exergonic breakdown of a high-energy organic substrate molecule
d. tautomerization: chemical reaction by which two tautomers are interconverted by the movement of a hydrogen atom and a double bond
e. tautomers: isomers that differ from each other in the location of a hydrogen atom and a double bond (e.g., keto-enol tautomers)

6

a. fructose-2,6-bisphosphate: an effector molecule that activates PFK-1 and stimulates glycolysis
b. glucagon: a hormone released in response to low blood glucose levels; increases the level of glucose in blood via the breakdown of liver glycogen

c. insulin: a hormone released in response to high blood glucose levels; promotes glucose uptake into the cells of certain target organs (muscle and adipose tissue)
d. lipogenesis: the biosynthesis of fatty acids and triacylglycerol (body fat)
e. AMPK: aMP-activated protein kinase; a metabolic regulatory enzyme that plays central role in energy metabolism

9

tautomers

12

The glycolytic reaction in which oxidation occurs is the conversion of glyceraldehyde-3-phosphate to glycerate-1,3-bisphosphate, catalyzed by glyceraldehyde-3-phosphate dehydrogenase.

15

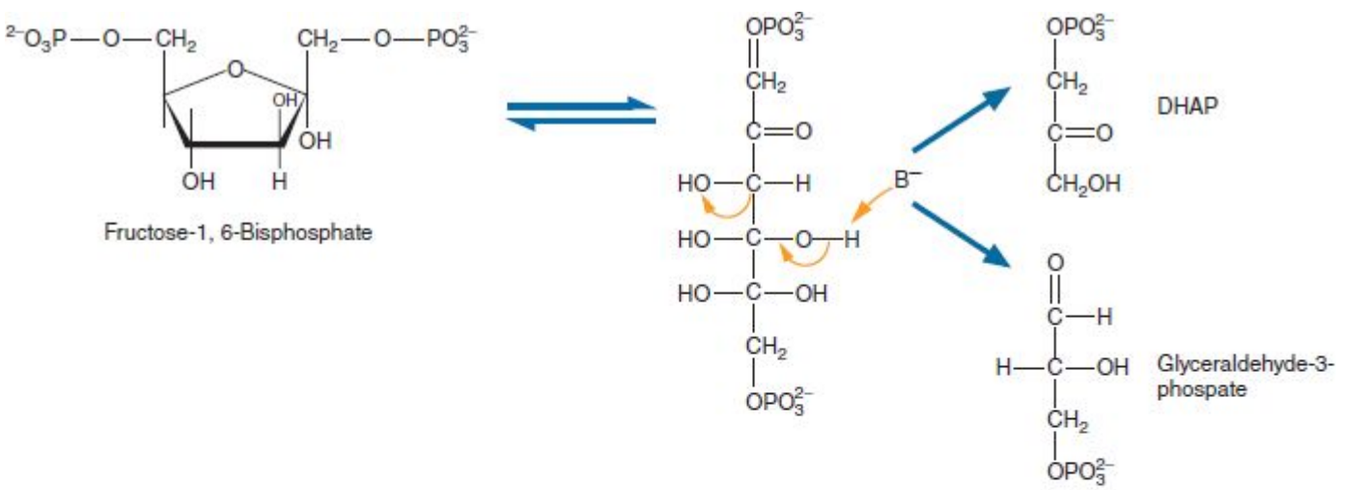
Severe hypoglycemia so dangerous because brain cells (and red blood cells) rely solely on glucose for their energy needs. Severe hypoglycemia causes fainting and requires immediate medical attention. When the condition is prolonged and/or recurrent, it may also result in long-term brain impairment or dysfunction.

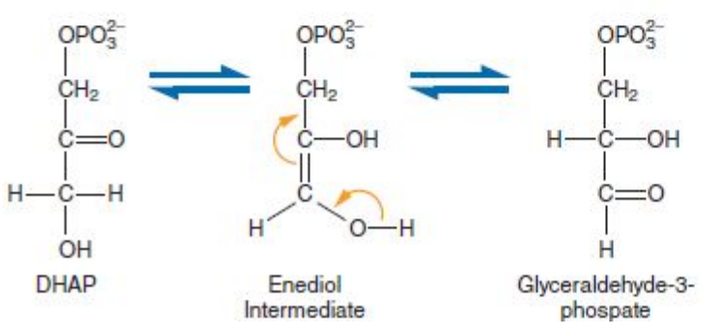
18

(i) ATP consumption: reaction 1(glucose-6-phosphate synthesis) and reaction 3 (fructose-1,6-bisphosphate synthesis); (ii) ATP synthesis – reactions 7 (glycerate-3-phosphate synthesis) and 10 (pyruvate synthesis); (iii) NADH synthesis – reaction 6 (glycerate-1,3-bisphosphate synthesis).

21

The mechanism by which fructose-1,6-bisphosphate is converted to two molecules of glyceraldehyde is as follows:





24

Two common oxidizing agents in anaerobic metabolism are NAD^+ and $NADP^+$.

27

a. turbo design: refers to a metabolic mechanism in which highly exergonic reactions at the beginning of a pathway allow subsequent energy-producing reactions to run thermodynamically downhill
b. malate shuttle: oxaloacetate is transferred from the mitrochondrial matrix to the cytoplasm by reversible conversion to malate; allows gluconeogenesis to continue because it produces NADH
c. hepatocyte: the major type of a liver cell
d. enterocyte: an absorptive cell of the intestinal lining
e. Cori cycle: a metabolic process in which lactate produced in skeletal muscle is transferred to liver where it becomes a substrate for gluconeogenesis

30

Cori

33

Assuming that pyruvate (a three-carbon compound) is used solely in the resynthesis of glucose (a six-carbon compound), 6 ATPs would be required. The conversion of two molecules of pyruvate and 2 CO_2 to oxaloacetate requires 2 ATP (1 ATP for each molecule of pyruvate). The subsequent decarboxylation of oxaloacetate to yield two molecules of phosphoenolpyruvate and CO_2 requires 2 ATP equivalents in the form of 2 GTP. The last ATP-requiring reaction in gluconeogenesis is the phosphorylation of two molecules of glycerate-3-phosphate at the expense of 2 ATP to yield two molecules of glycerate-1,3-bisphosphate.

36

Only AMP acts as an inhibitor of gluconeogenesis by depressing the activity of fructose-1,6-bisphosphate. Lactate, ATP, pyruvate, glycerol, and acetyl- CoA all stimulate gluconeogenesis. The increased availability of pyruvate (a product of lactate reduction via the Cori cycle); ATP and glycerol promote gluconeogenesis because enzymes of non-reversible reactions are activated when substrate levels are high. Acetyl-CoA is an allosteric activator of pyruvate carboxylase, one of two enzymes that by-pass the irreversible conversion of PEP to pyruvate.

39

The unique reactions in gluconeogenesis are (1) synthesis of phosphoenolpyruvate, catalyzed by pyruvate carboxylase and PEP carboxylase, (2) conversion of fructose 1,6-bisphosphate to fructose-6-phosphate, catalyzed by fructose-1,6-bisphosphatase, and (3) the formation of glucose from glucose-6-phosphate, catalyzed by glucose-6-phosphatase. Each of these reactions is necessary for bypassing irreversible glycolytic reactions. Each of these reactions is necessary for bypassing irreversible glycolytic reactions.

42

When the cell membrane is compromised, the contents of the glycosome and the general cytoplasm mix and glycolysis proceeds unchecked. All the ATP in the cell is then used up and the cell dies.

45

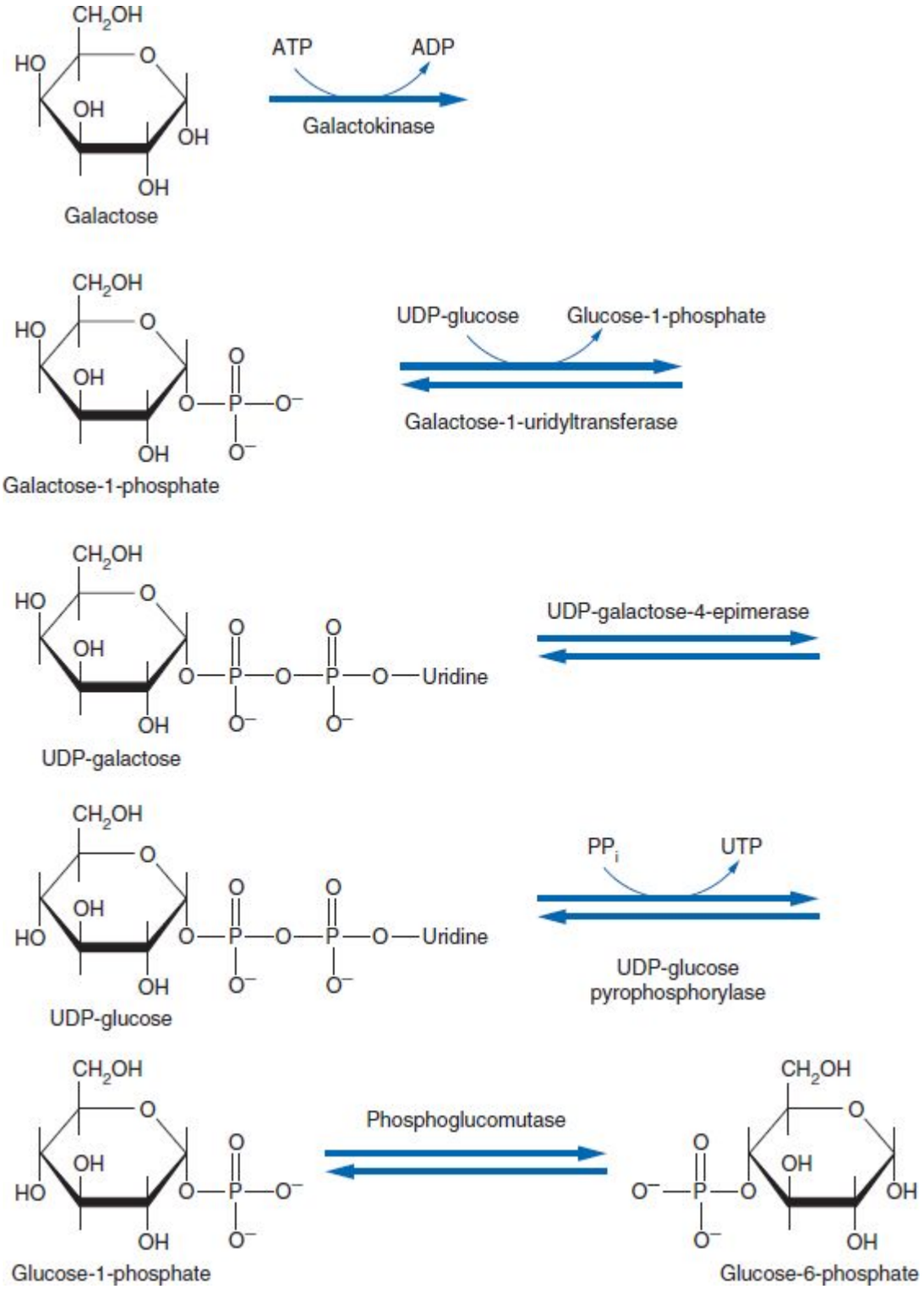
a. lactone: a cyclic ester
b. antioxidants: a substance that prevents the oxidation of other molecules
c. transketolase: a thiamine pyrophosphate (TPP)–requiring enzyme that transfers two-carbon units from a ketose to an aldose
d. transaldolase: the enzyme that transfers a three-carbon unit from a ketose to an aldose
e. pentose phosphate pathway: a biochemical pathway that produces NADPH, ribose and several other sugars

48

oxidative, pentose phosphate

51

The reactions that convert galactose to glucose are as follows:



54

In liver, fructose is metabolized more rapidly than glucose because its metabolism bypasses two regulatory steps in the glycolytic pathway: the conversion of glucose to glucose-6-phosphate and fructose-6-phosphate to fructose-1,6-bisphosphate. Recall that fructose-1-phosphate is split into glyceraldehyde and DHAP, both of which are subsequently converted to glyceraldehyde-3-phosphate, a glycolytic intermediate. The subsequent reactions in glycolysis are regulated largely by substrate availability.

57

Refer to Figures **8.14a** and **8.14b**, which illustrate the pentose phosphate pathway. Note that the ^{14}C label at the C-2 of glucose-6-phosphate is C-1 of ribulose-5-phosphate as the result of a decarboxylation reaction. When ribulose-5-phosphate enters the nonoxidative phase of the pentose phosphate pathway the radioactive label will appear as C-1 of ribose-5-phosphate and xylulose-5-phosphate. Because of seemingly random transfer of carbon groups by transketolase and transaldolase. Further ^{14}C-tracing through the nonoxidative pathway are very different to interpret with some molecules containing two ^{14}C-labeled carbon atoms.

60

UDP-glucose

63

In muscle, glycogen is synthesized to store glucose for energy, or metabolized to yield glucose when energy is needed. In the liver, glycogenesis and glycogenolysis are regulated to maintain blood glucose levels.

66

Glucokinase plays a vital role in the regulation of blood glucose. In an individual with glucokinase deficiency, following a carbohydrate meal blood glucose levels would be higher than normal. Recall that the kinetic properties of glucokinase allow the liver to remove excess glucose from blood. Since glucokinase acts as a glucose sensor in pancreatic β-cells, glucokinase deficiency results in reduced insulin release which also contributes to blood glucose levels.

69

b. ATP

72

a. pyruvate kinase

Chapter 9: In-Chapter Questions

9.1

With a $\Delta E_0'$ value of $-0.345\ V$, the oxidation of NO_2^- is not spontaneous as written. The oxidation of ethanol is spontaneous as written because its $\Delta E_0'$ value is positive ($+0.275\ V$).

9.2

Reactions 3, 4, and 5 are redox reactions. In reaction 3, lactate is the reducing agent and NAD^+ is the oxidizing agent. In reaction 4, cyt b (Fe^{2+}) is the reducing agent and NO_2^- the oxidizing agent. In reaction 5, NADH is the reducing agent and CH_3CHO is the oxidizing agent.

9.3

The oxidation states of the functional group carbon (indicated in bold) are:

$CH_3\mathbf{C}H_2OH$	$0 - 1 - 1 + 1 = -1$
$CH_3\mathbf{C}HO$	$0 - 1 + 2 = +1$
$CH_3\mathbf{C}OOH$	$0 + 1 + 2 = +3$

9.4

As it is incorporated into an organic molecule, the carbon atom in CO_2 is reduced.

9.5

Figure 9.8 illustrates that the incoming labeled carbons in the acetyl group are not released as CO_2 in the first turn of the cycle. During this first turn, labeled carbons are evenly distributed in the methylene and carbonyl carbons of symmetric succinate molecules and their derivatives. In the second turn of the cycle, all of the labeled carbons originating as the carbonyl groups of acetyl-CoA are released as CO_2 by the reactions catalyzed by isocitrate dehydrogenase and a-ketoglutarate dehydrogenase. The release of labeled methyl group-derived carbons as CO_2 begins in the third cycle because these carbons are now carbonyl carbons in intermediates such as oxaloacetate. During the third turn and each subsequent turn, one-half of the labeled carbon will be released as CO_2.

9.6

Review the structures of isocitrate and a-ketoglutarate. In the conversion of isocitrate to 2-hydroxyglutarate, the mutated enzyme IDH1 causes the release of CO_2 yielding a new methylene group but fails to convert the 2-hydoxy group of isocitrate to the carbonyl group of a-ketoglutarate.

9.7

Pyruvate carboxylase converts pyruvate to oxaloacetate. If the enzyme is inactive, concentrations of pyruvate in the system rise and pyruvate is converted by NADH to lactate. Excess lactate is then excreted in the urine.

Chapter 9: End-of-Chapter Questions

3

a. aerobic metabolism: the mechanism by which the chemical bond energy of food molecules is captured and used to drive the oxygen-dependent synthesis of ATP
b. aerobic respiration: the metabolic process in which oxygen is used to generate energy from food molecules
c. $\Delta G^{o\prime}$: free energy change at biochemical standard state conditions; ΔG^o at pH 7
d. $\Delta E^{\circ\prime}$: the difference in reduction potential between an electron donor and an electron acceptor under standard conditions: pH = 7, temperature = 25^{o}C, and pressure = 1 atm
e. F: Faraday constant; 96,485 J/V·mol, the magnitude of electric charge per mol of electrons

6

VDAC

9

Ancient earth possessed an atmosphere that contained methane, ammonia and was devoid of oxygen. With the development of photosynthesis, oxygen was released into the atmosphere. Subsequently, this oxygen reacted with methane to form carbon dioxide and with ammonia to form molecular nitrogen. The continued release of oxygen produced an oxidizing atmosphere

consisting of primarily oxygen, nitrogen, and carbon dioxide.

12

The difference in energy yield is

$$\Delta G^{o\prime} \text{ (O2 reaction)} - \Delta G^{o\prime} \text{ (S reaction)}$$
$$= 219{,}985 \text{ J} - 17{,}367 \text{ J}$$
$$= 202{,}618 \text{ J} = 200 \text{ kJ}$$

15

The half-cells for the reaction are

$$1/2\ 0_2 + 2\ H^+ + 2\ e^- \rightarrow H_2O \qquad E^{o\prime} = +0.82 \text{ V}$$
$$FADH_2 \rightarrow FAD + 2H^+ + 2\ e^- \qquad E^{o\prime} = +0.\ 22 \text{ V}$$

For the formation of water the reaction would be

$$1/2\ O_2 + FADH_2 \rightarrow H_2O + FAD$$
$$\Delta E^{o\prime} = +0.82 \text{ V} + 0.22 \text{ V} = 1.04 \text{ V}$$
$$\Delta G^{o\prime} = 2(96485 \text{ J/mol.V})\ (1.04) = 200688 \text{ J} = 201 \text{ kJ}$$

Note that the free energy for 1/2 0_2 + NADH + H+ → H_2O + NAD+ is 220 kJ, which is a higher value than the free energy for the reaction involving $FADH_2$.

18

a. HETPP: hydroxyethyl-thiamine pyrophosphate – an intermediate in the oxidative decarboxylation of pyruvate to yield acetyl-CoA
b. nucleoside diphosphate kinase: an enzyme that catalyzes a reversible reaction in which phosphoryl group of a nucleoside triphosphate is donated of to a nucleoside diphosphate to yield a new nucleoside triphosphate
c. amphibolic pathway: a metabolic pathway that functions in both anabolism and catabolism
d. anaplerotic reaction: a reaction that replenishes a substrate needed for a biochemical pathway
e. citric acid cycle: a biochemical pathway that degrades the acetyl group of acetyl-CoA to yield CO_2 and H_2O as three molecules of NAD^+ and one molecule of FAD are reduced

21

mitochondrial matrix

24

anaplerotic

27

PKB

30

The enzyme pyruvate carboxylase converts pyruvate to oxaloacetate. Its deficiency causes a

buildup of pyruvate in cytoplasm, where it reacts with NADH to yield lactate. When blood levels of lactate are higher than normal, the molecule is excreted in urine.

33

Pyruvate dehydrogenase complex is a large multienzyme complex that contains three enzyme activities, each of which is present in multiple copies. E1 is pyruvate dehydrogenase (or pyruvate decarboxylase) with TPP, E2 is dihydrolipoyl transacetylase with lipoic acid and CoASH, and E3 is dihydrolipoyl dehydrogenase, with NAD+ or FAD. The pyruvate dehydrogenase complex in mammalian cells contains 60 copies of E2 and 20–30 copies each of E1 and E3.

36

The net equation for the citric acid cycle is:

$$\text{Acetyl-CoA} + 3\text{NAD}^+ + \text{FAD} + \text{GDP (or ADP)} + \text{Pi} + 2\text{H2O}$$
$$2\text{CO2} + 3\text{NADH} + \text{FADH2} + \text{CoASH} + \text{GTP (or ATP)} + 2\text{H}^+$$

39

Isocitrate dehydrogenase requires NAD^+ (or $NADP^+$) and the cofactor Mn^{2+}. The α-ketoglutarate dehydrogenase complex requires TPP (thiamine pyrophosphate), lipoic acid, CoASH (coenzyme A), and NAD^+. Succinate dehydrogenase requires FAD. Malate dehydrogenase requires NAD^+.

42

The steps in the citric acid cycle that are regulated are those catalyzed by citrate synthase, which converts acetyl-CoA and oxaloacetate into citrate; isocitrate dehydrogenase, which converts isocitrate to α-ketoglutarate; and α-ketoglutarate dehydrogenase, which forms succinyl-CoA. Citrate synthase is regulated by substrate availability, especially oxaloacetate. Isocitrate dehydrogenase is inhibited by NADH and ATP and activated by NAD^+ and ADP. α-Ketoglutarate is inhibited by NADH and its product succinyl-CoA, and activated by AMP. These reactions represent important metabolic branch points. Because citrate can penetrate the mitochondrial membrane and be cleaved to form acetyl-CoA and oxaloacetate, it therefore can be used to transport acetyl-CoA out of the mitochondrion for fatty acid synthesis. Also, acetyl-CoA activates pyruvate carboxylase (and inhibits pyruvate dehydrogenase), which activates the reaction of pyruvate with malate to form oxaloacetate (instead of more acetyl-CoA). Citrate also inhibits PFK-1 in glycolysis and it also activates the first step in fatty acid synthesis. α-Ketoglutarate plays an important role in amino acid metabolism and other metabolic processes. Succinyl-CoA is required in the synthesis of heme.

45

Malonate is a competitive inhibitor of the enzyme succinate dehydrogenase because its structure is similar to that of the enzyme's normal substrate, i.e. the substrate and the inhibitor compete for entry into the active site. In a laboratory experiment malonate's inhibition of the enzyme is overcome by adding a relatively large amount of succinate to the enzyme assays, thereby reducing the possibility of the inhibitor's entry into the active site.

48

Three enzymes that require thiamine are pyruvate dehydrogenase, α-ketoglutarate dehydrogenase, and transketolase. Thiamine as a component of TPP is involved in decarboxylation and acyl group

transfer reactions. Decreased levels of decarboxylation reactions depresses pyruvate decarboxylation to yield fewer molecules of acetyl-CoA. Increased levels of pyruvate leads to lactate accumulation. Thiamine deficiency causes disruptions in several metabolic processes (e.g., carbohydrate metabolism) in most of the body's organs. The overall symptoms of thiamine deficiency include lack of energy, muscle wasting and acidosis. Severe thiamine deficiency results in beriberi, a disease process in which here are inflammatory changes in nerves, heart and digestive system organs.

51

Common two-carbon fragments that could be present are acetate and acetyl-CoA.

54

Carboxylation of pyruvate produces oxaloacetate, a citric acid cycle intermediate. Increasing the concentration of one of the intermediates stimulates the cycle, thereby resulting in increased energy production.

57

When oxygen levels are reduced products of the citric acid cycle accumulate. Energy generation is primarily by glycolysis. The end product of glycolysis is pyruvate. In order to regenerate the NAD^+ required for the reaction pyruvate is reduced to lactate.

60

b. O_2 is the final electron acceptor.

63

b. lipoic acid, NAD^+, thiamine pyrophosphate, and FAD

Chapter 10: In-Chapter Questions

10.1

a. NADH
b. FADH2
c. Cyt b (reduced)
d. NADH
e. NADH

10.2

DNP is a lipophilic molecule that binds reversibly with protons. It dissipates that proton gradient in mitochondria by transferring protons across the inner membrane. The uncoupling of electron transport from oxidative phosphorylation causes the energy from food to be dissipated as heat. DNP causes liver failure because of insufficient ATP synthesis in a metabolically demanding organ.

10.3

No, for ATP synthesis to occur, the proton concentration must be higher within the inside-out mitochondrial particles. ATP synthesis requires that protons move down a concentration gradient

through the base of the ATP synthetase across the membrane.

10.4

Disregarding proton leakage and assuming that the glycerol phosphate shuttle is in operation, 38 ATP would be produced from the aerobic oxidation of a glucose molecule. If the malate shuttle is in operation, only 36 ATP would be produced.

10.5

Sucrose is a disaccharide composed of glucose and fructose. As described (p. 383), the oxidation of 1 mol of glucose yields a maximum of 31 mol of ATP. Fructose, which like glucose is also partially degraded by the glycolytic pathway, also yields a maximum of 31 mol ATP. The total maximum energy yield is 62 mol of ATP per mol of sucrose.

10.6

The larger selenium atom holds its electrons less tightly than sulfur. Selenium is more easily oxidized and therefore acts as a better scavenger for oxygen than does sulfur.

10.7

The SH groups reduce hydrogen peroxide or trap hydroxyl radicals to form water. An example of a nonsulfhydryl group–containing molecule that should be capable of this activity is vitamin C or any of a number of other antioxidants (carotenoids, flavonoids, tocopherols, etc.).

10.8

The phenolic groups of both molecules are responsible for their antioxidant activity because of the ease of formation of phenoxy radicals with subsequent neutralization of electron-deficient ROS.

Chapter 10: End-of-Chapter Questions

3

a. complex I: NADH dehydrogenase complex; a component of the mitochondrial ETC; catalyzes the transfer of electrons from NADH to UQ
b. complex II: succinate dehydrogenase complex – a component of the mitochondrial ETC mediates the transfer of electrons from succinate via $FADH_2$ to UQ
c. complex III: cytochrome bc_1 complex; in the mitochondrial ETC mediates the transfer of electrons from UQH_2 to cytochrome c
d. complex IV: cytochrome oxidase; catalyzes the four electron reduction of oxygen to water
e. ETC: electron transport chain; a series of electron carrier proteins that bind reversibly to electrons at different energy levels

6

respirasome

9

FMN

12

There are four protein complexes in the mitochondrial ETC. Complex I (NADH dehydrogenase

complex) catalyzes the transfer of electrons from NADH to UQ. Complex II (succinate dehydrogenase complex) mediates the transfer of electrons from succinate, an intermediate in the citric acid cycle, via $FADH_2$ to UQ. Complex III (cytochrome bc_1 complex) transfers electrons from UGH_2 to cytochrome c, a mobile electron carrier protein loosely associated with the outer face of the inner mitochondrial membrane. Complex IV (cytochrome oxidase) transfers the electrons donated by cytochrome c that reduce oxygen to yield water.

15

In order for the system to operate efficiently energy should be released in stages. Each relatively large decrease in potential corresponds to the energy required to generate an ATP molecule. If the energy is released all at once much of the energy would be released as heat and little, if any, ATP would be produced. The gradual change in voltage of the cytochromes makes this gradual release of energy possible.

18

If oxygen is used as the electron acceptor for the ETC, the difference in voltage between NADH and water is +1.14V. The complete reduction of nitrate by NADH would be +1.18V. This is almost exactly the same as for oxygen. It would be reasonable to assume that the same amount of ATP would be produced.

21

a. ATP synthase: the enzyme complex that synthesizes ATP; in mitochondria; ATP synthase consists of F_1, the ATPase activity and F_o, a transmembrane channel for protons
b. submitochondrial particle: small membranous vesicles formed when mitochondria are subjected to sonication; used to investigate ATP synthesis
c. α, β-hexamer: a component of the F_1 unit of mitochondrial ATP synthesis; catalyzes ATP synthesis
d. torque: twisting force that causes rotation
e. rotor: a rotary motor; a device that revolves around a central axis

24

nonshivering thermogenesis

27

uncouplers

30

Nystatin is an ionophore because of its capacity to dissipate an ion gradient.

33

The translocation of three protons is required to drive ATP synthesis. The fourth proton drives the transport of ADP and P_i.

36

The passage of K^+ across a membrane through an ionophore such as valinomycin lowers the electrical potential gradient. As a result, some of the energy that would have been used to synthesize ATP is dissipated as heat. This explains the rise in body temperature and sweating upon

treatment with valinomycin.

39

The oxidation of one mole of ethanol to acetyl-CoA produces two moles of NADH. The conversion of acetate to carbon dioxide and water through the citric acid cycle produces 3 NADH, 1 $FADH_2$, and 1 GTP. Assuming that the aspartate-malate shuttle is in operation, each cytoplasmic NADH yields 2.25 ATP for a total of 4.5 ATP. Each mitochondrial NADH yields 2.5 ATP for a total of 7.5 ATP. Each molecule of $FADH_2$ yields 1.5 ATP for a total of 1.5 ATP. Each GTP yields 0.75 ATP for a total of 0.75 ATP. The total ATP produced by the oxidation of ethanol is therefore 14.25 ATP.

42

Once a labeled acetate molecule is converted to acetyl-CoA it is processed through the citric acid cycle (Figure 9.8 on p. 340). Because of the symmetrical structure of the intermediate succinate, both $^{14}CO_2$ molecules are not released until two or more turns of the cycle.

45

a. reactive oxygen species: a reactive derivative of molecular oxygen, including superoxide radical, hydrogen peroxide, the hydrogen radical, and singlet oxygen
b. reactive nitrogen species: nitrogen-containing radicals often classified as RNS; examples include nitric oxide, nitrogen dioxide, and peroxynitrite
c. superoxide dismutase: one of a class of enzymes that catalyze the formation of hydrogen peroxide and oxygen from the superoxide radical
d. antioxidant: a substance that prevents the oxidation of other molecules
e. oxidative stress: excessive production of reactive oxygen species

48

a. GSH: glutathione; an intracellular reducing agent
b. SOD: superoxide dismutase; an enzyme that catalyzes the conversion of superoxide radical to hydrogen peroxide and oxygen
c. redox code: a set of principles used to describe cell function in relationship to NAD(P)+/NAD(P)H, GSH/GSSG, the cysteine proteome and other redox molecules
d. RNS: reactive nitrogen species; nitrogen-containing radicals; nitric oxide, nitrogen dioxide, and peroxynitrite
e. radical: an atom or molecule with an unpaired electron

51

respiratory burst

54

Reaction equations that illustrate the production of ROS from electrons leaking from the electron transport system include:

$$O_2 + e^- \rightarrow O_2^- \text{ (superoxide)}$$

$$2H^+ + 2\,O_2^- \rightarrow O_2 + H_2O_2 \text{ (peroxide)}$$

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + \cdot OH \text{ (hydroxyl radical)}$$

$$2\,O_2^- + 2H^+ \rightarrow H_2O_2 + {}^1O_2 \text{ (singlet oxygen)}$$

$$2\,ROOH \rightarrow 2ROH + {}^1O_2$$

57

The many causes of reperfusion-induced cardiac cell damage are primarily the result of ROS production and the opening of MPTP. ROS are generated by a reenergized mitochondrial ETC, the neutrophil enzymes NADPH oxidase and xanthine oxidase, and iron-induced hydroxyl ion production. Acidosis, caused by lactate accumulation, triggers increased oxygen unloaded by hemoglobin which causes further ROS synthesis. Reperfusion also promotes increased nitric oxide synthesis. Nitric oxide reacts with superoxide to form an even more dangerous radical, peroxynitrite. MPTP opening triggered by ROS low ATP levels and high calcium levels causes the mitochondrial membrane potential to collapse.

60

The reaction pathway is as follows:

$$\text{Cys-SH} + \text{ROOH} \rightarrow \text{Cys-SOH} + \text{ROH}$$

$$\text{Cys-SOH} + \text{Cys-SH} \rightarrow \text{Cys-S-S-Cys} + H_2O$$

$$\text{Cys-S-S-Cys} + \text{TRX-(SH)}_2 \rightarrow \text{Cys-(SH)}_2 + \text{TRX-(SH)}_2$$

where Cys-SOH = sulfenic acid
and TRX-$(SH)_2$ = oxidized thioredoxin

63

c. b

Chapter 11: In-Chapter Questions

11.1

Because steroids inhibit the release of arachidonic acid, their use shuts down the synthesis of most if not all eicosanoid molecules, hence their reputation as potent anti-inflammatory agents. Aspirin inactivates cyclooxygenase and prevents the conversion of arachidonic acid to PGG_2 the precursor of prostaglandins and thromboxanes. Aspirin is not as effective an anti-inflammatory agent as the steroids because it shuts down only a portion of eicosanoid synthetic pathways. [Refer to the Chapter 11 online essay entitled the Eicosanoids for further information.]

11.2

The product of complete hydrogenation would be hard and therefore not useful as a margarine.

11.3

When soap and grease are mixed, the hydrophobic hydrocarbon tails of the soap insert (or dissolve) into the oil droplet. The oil droplet becomes coated with soap molecules. The hydrophilic portion

of the soap molecules allows the soap-oil complex to be dispersed in water.

11.4

The phospholipid of the surfactant, which possesses a polar head group and two hydrophobic acyl groups, disrupts some of the intermolecular hydrogen bonds of the water, thereby decreasing the surface tension.

11.5

Carvone and camphor are monoterpenes (two isoprene units); abscisic acid is a sesquiterpene (three isoprene units).

11.6

Bile salts are structurally similar to soap in that they contain a polar head group (e.g., the charged amino acid residue glycine) and a hydrophobic tail (the steroid ring system).

11.7

In alligators, the vascular shunt diverts additional CO_2-rich blood to the parietal cells of the stomach. The CO_2 flowing down its concentration into the parietal cell is immediately converted to HCO_3^-, which is subsequently released into the blood in exchange for chloride ion. These additional chloride ions and protons are then secreted as hydrochloric acid into the stomach lumen to assist digestion.

11.8

a. simple diffusion
b. secondary active transport or facilitated diffusion
c. primary active transport or exchanger protein
d. primary active transport or gated channel
e. fat molecules (triacylglycerols) are not directly transported across cell membranes; they must be hydrolyzed first
f. simple diffusion

11.9

The main stabilizing feature of biological membranes is hydrophobic interactions among the molecules in the lipid bilayer. The phospholiqids in the lipid bilayer orient themselves so that their polar head groups interact with water. Proteins in the lipid bilayer interact favorably in their hydrophobic milieu because they typically have hydrophobic amino acid residues on their outer surfaces.

11.10

The transport mechanisms discussed in the chapter fit into the following categories:

sodium channel: uniporter
glucose permease: passive uniporter
Na^+-K^+-ATPase: antiporter

Chapter 11: End-of-Chapter Questions

3

a. eicosanoid: a hormone-like molecule that contains 20 carbons; most are derived from arachidonic acid; examples include prostaglandins, thromboxanes, and leukotrienes
b. omega-6-fatty acid: linoleic acid and its derivatives
c. omega-3-fatty acid: α-linolenic acid and its derivatives such as eicosapentaenoic acid and docosahexaenoic acid
d. trans fatty acid: a side product of margarine production in which plant oils are hydrogenated; three-dimensional structure similar to that of saturated fatty acids
e. acyl group: a molecular group derived from a carboxylic acid by the removal of a hydroxyl group

6

a. phospholipid: an amphipathic molecule that has a hydrophobic domain (hydrocarbon chains of fatty acid residues) and a hydrophilic (a polar head group) domain
b. soap: sodium or potassium salts of fatty acids; the product of a saponification reaction involving TGs
c. surface active agent: a substance that lowers the surface tension of a liquid, usually water, so that it spreads out over a surface
d. phosphoglyceride: a membrane lipid composed of glycerol linked to two fatty acids, phosphate, and a polar group
e. sphingomyelin: a membrane lipid that contains a long-chain amino alcohol and ceramide (a fatty acid derivative of sphingosine)

9

a. ganglioside: a sphingolipid that possesses an oligosaccharide group with one or more sialic acid residues
b. sphingolipidoses: a class of inherited disorders in which an enzyme required to degrade a specific sphingolipid is missing; a type of lysosomal storage disease
c. isoprenoid: one of a class of molecules that contain repeating five-carbon structural units known as isoprene units; examples include terpenes and steroids
d. terpene: one of a class of isoprenoids classified according to the number of isoprene residues it contains; examples include carotenoids, rubber, and many of the molecules found in essential oils of plants
e. mixed terpenoid: a biomolecule composed of a nonterpene component attached to isoprene groups; examples include α-tocopherol and plastoquinone

12

steroid

15

apolipoproteins

18

Diets in which there are insufficient fats and oils can have serious health consequences. Among these are deficiencies in fat soluble vitamins (A,D,E, and K) and the essential fatty acids linoleic and linolenic acids, which together can result in dry skin, brittle hair, fatigue, high blood pressure, atherosclerosis, depressed immunity, poor wound healing and depression. In children essential

fatty acid deficiencies have been linked to impaired brain development.

21

Triacylglycerol functions include energy storage, insulation, and shock absorption. Their highly reduced, long hydrocarbon chains are a very efficient storage form of energy; in addition, their hydrophobicity results in compact storage in adipocytes. The relatively low heat conductivity helps to prevent heat loss and serves organisms as insulation.

24

The outer layer of lipoproteins consists of a single layer of phospholipids and proteins. The hydrophobic hydrocarbon chains of the phospholipids face inward, towards the neutral lipids contained within. The hydrophilic group of the phospholipids face outward and are solvated by water molecules, allowing the lipoprotein to dissolve in the bloodstream.

27

Both carbohydrates and proteins contain large numbers of electronegative atoms (oxygen and nitrogen) that are capable of forming hydrogen bonds with water molecules. In the presence of external water most of these molecules would dissolve. In contrast, waxes are composed of hydrophobic molecules that are resistant to the penetration of water into a leaf's interior. A relatively thick wax layer prevents insect penetration.

30

Dipalmitoylphosphatidylcholine is a component of surfactant, a surface active agent that is secreted into lung alveoli where it reduces the surface tension of water on their extracellular surfaces. Without surfactant alveoli collapse, thereby preventing oxygen and carbon dioxide diffusion, which results in suffocation.

33

Terpenes are a class of isoprenoids that most commonly occur in in the essential oils and resins of plants. The carotenoids, which include the carotenes and the xanthophylls, are plant pigments. Because of their desirable properties, certain terpenes are used in commercial products. For example, farnesene, found in the oil of citronella, is used in soap and perfumes. Turpentine is a distilled product of tree resin.

36

a. heterokaryon: an experimentally produced cell formed by the fusion of different types of cells; used to investigate membrane function or medical conditions; certain viruses or chemicals facilitate the fusion process
b. aquaporin: one of a class of water channels in cell membranes
c. integral protein: a protein that is embedded within a membrane
d. peripheral protein: a protein that is attached to a membrane by either a covalent bond to a lipid molecule or by noncovalent interactions with a membrane integral protein or membrane lipids
e. AE1: the band 3 anion exchanger protein – an integral membrane protein found in red blood cells; a channel protein that allows the transport of chloride in exchange for bicarbonate ion

39

a. membrane carrier: a membrane protein used in facilitated diffusion; a specific substance binds to the carrier on one side of the membrane and is delivered to the opposite side and then

released; also referred to as passive transporters

b. primary active transport: a form of transmembrane transport of substances against their concentration gradients that requires energy provided by ATP hydrolysis
c. secondary active transport: concentration gradients generated by primary active transport are harnessed to move substances across membranes against their concentration gradients
d. nephrogenic diabetes insipidis: a disease in which affected individuals cannot produce a concentrated urine; caused by a mutation in the water channel AQP2 gene
e. cystic fibrosis: a disease caused by a missing or dysfunctional chloride channel referred to as CFTR

42

fluidity

45

passive

48

atherosclerosis

51

For a phospholipid to move from one side of a bilayer to the other, the polar head must move through the hydrophobic portion of the membrane. This process requires a significant amount of energy and it therefore relatively slow.

54

In facilitated diffusion, polar, charged, or large molecules that normally cannot penetrate the cell membrane diffuse across the membrane through protein channels or carriers that “facilitate” this diffusion. Because facilitated diffusion occurs with (or down) a concentration gradient, this is a spontaneous process, so a coupled reaction that provides energy (such as the hydrolysis of ATP) is not needed. The transport of glucose across red blood cell membranes is an example of facilitated diffusion by a carrier protein, and anion channels in red blood cell membranes are examples of protein channels. (Note that the transport of glucose across the plasma membrane of kidney tubule cells is an example of secondary active transport, a distinctly different transport mechanism.)

57

To increase a cell’s resistance to mechanical stress, increase its content of cholesterol and cardiolipin in the cell membrane.

60

The carbohydrate portion of the glycolipid can form hydrogen bonds with the water. This carbohydrate is the polar group, and it is analogous to the charged portion of the phospholipid.

63

Many transmembrane and peripheral proteins are attached to the cytoskeleton and therefore are not free to move in the phospholipid bilayer.

66

The fluidity of the membrane allows for flexible movement. Any breaks that do occur expose the

hydrophobic core of the membrane to an aqueous environment. Hydrophobic interactions spontaneously move the broken ends together and, in combination with certain other components of cell membrane resealing mechanisms (e.g., cytoskeleton and calcium ions), the membrane reseals.

69

b. xanthophylls are derivatives of carotenes

Chapter 12: In-Chapter Questions

12.1

The triacylglycerols are emulsified in the small intestine by bile salts. They are then digested by lipases, the most important of which is pancreatic lipase. The products, fatty acids and mono acylglycerol, are transported into enterocytes and reconverted to triacylglycerol. Triacylglycerol is subsequently incorporated into chylomicrons, which are then transported into lymph via exocytosis and finally into the bloodstream for transport to the fat cells.

12.2

Although there is no cure for MCAD, symptoms can be managed primarily by ensuring that frequent feeding occurs, i.e., there are no prolonged fasts. This can be a challenge in patients who often manifest a poor appetite so intravenous feeding with glucose may be necessary. Since toxic products of fatty acid metabolites can accumulate, the diet must be low in fat. Daily consumption of carnitine supplements may also be beneficial because acyl-carnitine derivatives are excreted in urine.

12.3

a. phospholipid
b. acyl-CoA
c. carnitine

12.4

Unlike the oxidation of glucose to form pyruvate, fatty acid oxidation, which involves the citric acid cycle and the electron transport system, cannot operate in the absence of O_2.

12.5

The yield from the oxidation of stearyl-CoA is calculated as follows:

8 $FADH_2 \times 1.5$ ATP/$FADH_2$ =	12 ATP
8 NADH × 2.5 ATP/NADH =	20 ATP
9 Acetyl-CoA × 10 ATP/Acetyl-CoA =	90 ATP
	122 ATP

Two ATP are required to form stearyl-CoA from stearate to give a total of 120 ATP.

12.6

Propionyl-CoA can be reversibly converted to succinyl-CoA, an intermediate in the citric acid cycle. Oxaloacetate, a downstream intermediate of this cycle can be converted to PEP. PEP is then

converted to glucose via gluconeogenesis.

12.7

Adipic acid undergoes one round of β-oxidation to yield acetyl-CoA and succinyl-CoA. Succinyl-CoA is sequentially converted to oxaloacetate, PEP, and then to glucose.

12.8

Following the hydrolysis of sucrose, both monosaccharide products enter the bloodstream and travel to the liver, where fructose is converted to fructose-1-phosphate. Recall that the conversion of fructose-1-phosphate to glyceraldehyde-3-phosphate bypasses two regulatory steps. Consequently, more glycerol-phosphate and acetyl-CoA (the substrates for triacylglycerol synthesis) are produced. High blood glucose concentrations that result from this consumption of excessive amounts of sucrose trigger the release of larger than normal amounts of insulin. One of the functions of insulin is to promote fat synthesis.

12.9

a. β-Hydroxybutyrate is a product of ketone body metabolism.
b. Malonyl-CoA is the product of the reaction of acetyl-CoA and carboxybiotin that occurs during fatty acid systhesis.
c. Biotin is a carrier of CO_2 in fatty acid synthesis and several other reactions.
d. Acetyl ACP delivers acetate to the synthetic machinery of fatty acid synthesis.

12.10

The higher activity of HMG-CoA reductase in obese patients in combination with a high-calorie diet increases the synthesis of cholesterol.

Chapter 12: End-of-Chapter Questions

3

a. HSL: hormone-sensitive lipase; catalyzes the hydrolysis of diacylglycerol in adipocytes to yield monoacylglycerol and a fatty acid
b. CGI-58: a protein that binds to and activates adipose triglyceride lipase (ATGL)
c. fatty acid-binding protein: an intracellular water-soluble protein whose function is to bind and transport hydrophobic fatty acids
d. β-oxidation: the catabolic pathway in which most fatty acids are degraded; acetyl-CoA is formed as the bond between the α and β carbon atoms is broken
e. carnitine: a carrier molecule that transports fatty acids into the mitochondrial matrix across the mitochondrial inner membrane

6

a. sdLDL: small, dense LDLs; a high-risk factor for atherosclerosis
b. oxLDL: oxidized LDL
c. cytochrome P_{450}: one of a group of hemoproteins that when complexed with carbon monoxide absorb light at a wavelength of 450 nm; oxidizes a wide variety of hydrophobic molecules
d. atheroma: an atherosclerotic lesion in an artery wall, contains macrophages, lipids and cell debris

e. abetalipoproteinemia: an autosomal metabolic disorder in which dietary fat absorption is compromised; caused by failure of apoB-lipoproteins to form

9

glycerol-3-phosphate

12

carnitine

15

Peroxisomes have β-oxidation enzymes that are specific for long chain fatty acids, whereas mitochondria possess enzymes that are specific for short and moderate chain length fatty acids. The first reaction in the peroxisomal pathway is catalyzed by a different enzyme than in the mitochondrial pathway. The $FADH_2$ produced in the first peroxisomal reaction donates its electrons to oxygen directly (forming hydrogen peroxide instead of UQ). The processes are similar in that acetyl-CoA is derived from oxidation of fatty acids.

18

The yield of ATP from the oxidation of stearic acid is 122 (8 $FADH_2$ × 1.5 ATP, 8 NADH × 2.5 ATP and 9 acetyl-CoA x 10 ATP). Since the formation of stearoyl-CoA requires 2 ATP equivalents the net synthesis is 120 ATP. Stearic acid yields 120/18 or 6.6 ATP per carbon atom. Recall that the ratio for glucose is 5.2 ATP per carbon atom.

21

In the small intestine, triacylglycerols mix with bile acid salts and are emulsified. The size of the triacylglycerols and their emusification with the bile salts prevents them from crossing the enterocytes cell membranes.

24

Instead of cis-double bonds the fatty acids of trans fat, created by the partial hydrogenation of vegetable oils, contain trans-double bonds. The body can only process fatty acids with the cis-isomer. The consumption of trans fat-containing food poses a substantial risk for the development of cardiovascular disease because lipoprotein lipase can only bind molecules with the cis isomer. As a result, trans fat remains in the blood for long periods of time and is, therefore, likely to contribute to arterial plaque formation.

27

Since the $^{14}CO_2$ that is added to acetyl-CoA is removed in the reaction of malonyl-ACP with acetyl synthase to form acetoacetyl-ACP, no ^{14}C label appears in the eventual fatty acid products.

30

Enoyl-CoA isomerase converts the naturally cis double bond at Δ^3 to a trans double bond at Δ^2; the correct position for the next round of β-oxidation.

33

In the β-oxidation of fatty acids the final reaction (an Cα–Cβ cleavage), catalyzed by thiolase, is described as a thiolytic cleavage because of the cleavage mechanism involves the thiol group of

CoASH. The products of the reaction are acetyl-CoA and an acyl-CoA with two fewer carbon atoms.

36

NADH donates electrons to an electron transport system composed of cytochrome b_5 reductase (a flavoprotein) and cytochrome b_5 (which contains a heme). An oxygen-dependent desaturase uses these electrons (from NADH) to activate O_2. The activated O_2 oxidizes the fatty acyl-CoA to create an alkene in its hydrocarbon chain and is reduced to form two molecules of water.

39

The endogenous lipoprotein pathway transports lipids produced in the liver to the body's cells. VLDL assembled in the liver contain TGs, phospholipids, cholesterol, cholesteryl esters and apolipoprotein B-100. Once they are released by the liver VLDL unload TGs as they encounter lipoprotein lipase near the surface of the target cells (e.g., adipocytes and muscle cells). Once VLDL are depleted to yield IDLs, the latter lipoproteins are removed from blood by endocytosis into liver cells. IDLs are then converted to LDL as TGs are degraded by hepatic lipase. LDLs, now high in cholesterol, are released into blood. They are subsequently internalized by target cells with LDL receptors.

42

Although regular eating is not a panacea, it does provide sufficient carbohydrate to act as a fuel to sustain vital metabolic processes.

45

Fatty acids are assembled from two carbon acid synthase which ends the process at C-16. Although there is a robust capacity to elongate, the most abundant fatty acids are C_{16} and C_{18}, the product of one round of elongation process.

48

a. cholecystitis: an infection of the gallbladder
b. HMGR: HMG-CoA reductase; the rate limiting enzyme in cholesterol synthesis
c. SCAP: SREBP cleavage-activating protein; regulates SREBP2 activation
d. statins: competitive inhibitors of HMG-CoA reductase, the rate-limiting enzyme in cholesterol biosynthesis
e. SRE: sterol regulatory elements; DNA sequences that bind transcription factors involved in lipid metabolism

51

cholesterol

54

cholic acid and deoxycholic acid

57

Cells adjust the fluidity of their membrane with membrane remodeling, a process in which phospholipases and acyl transferases alter the fatty acid composition of membrane lipid molecules. The replacement of saturated fatty acids with unsaturated fatty acids increases fluidity.

60

All the lipid molecules mentioned are originally synthesized from the isoprene units in isopentenyl pyrophosphate molecules. Steroid and terpene molecules are assembled by head-to-tail condensation of these groups. Farnesyl pyrophosphate is a precursor of steroids such as estrogen and the carotenoid β-carotene. See Figure 12.25.

63

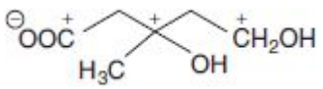
In individuals in whom cholesterol synthesis in the liver is higher than normal, cholesterol secreted in bile has a tendency to crystalize because of its insolubility in water.

66

For phospholipases to be drawn into the micelle, they must have a hydrophobic surface.

69

The ^{14}C label will appear in the mevalonate molecule as indicated by the asterisk.



71

c. inhibit HMG-CoA reductase

74

c. stimulating glucose-6-phosphate lyase

Chapter 13: In-Chapter Questions

13.1

Refer to **Figure 13.4** for an illustration of the relative locations of CF_oCF_1, P700 (in PSI), and P680 (in PSII) within thylakoid membrane. The Calvin cycle reactions occur in the stroma, the gel-like substance that surrounds the external surface of thylakoid membrane. CF_oCF_1 is the ATP synthase that utilizes the transmembrane proton gradient to drive ATP synthesis. P700 is the term used to indicate the special pair of chlorophyll a molecules in the reaction center of PSI that absorbs light energy and then donates the energized electrons that eventually reduce $NADP^+$. P680 is a special pair of chlorophyll a molecules in PSII that absorbs light energy and then donates energized electrons eventually to plastoquinone. The Calvin cycle enzymes are responsible for utilizing ATP and NADPH generated by the light reactions to incorporate $C0_2$ into carbohydrate molecules.

13.2

The energy of a photon is proportional to its frequency. Blue light has a higher frequency than green light and therefore has higher energy.

13.3

The presence of antenna pigments allows the light-harvesting systems of chloroplasts to collect energy from a wider range of frequencies than those absorbed by the special chlorophyl pairs.

Because their absorption spectra overlap, the energy absorbed by the antenna pigments is quickly transferred to the critical chlorophylls of PSI and PSII.

13.4

Excessive light promotes the formation of ROS, which damage proteins such as D_1. β-Carotene is an antioxidant that prevents some of this damage.

13.5

a. Plastocyanin is a component of the cytochrome b_6f complex; a copper-containing protein that accepts electrons from plastoquinone.
b. β-Carotene is a carotenoid pigment that protects chlorophyll molecules from ROS.
c. Ferredoxin is a mobile, water-soluble protein that donates electrons to a flavoprotein called ferredoxin-NADP oxidoreductase.
d. Plastoquinone is a component of photosystem II that accepts electrons from pheophytin a to become plastoquinol.
e. Pheophytin a is a molecule similar in structure to chlorophyll that is a component of the electron transport pathway between PSII and PSI.
f. Lutein is a carotenoid that is a component of light-harvesting complexes.

13.6

Of the herbicides discussed, paraquat and DCMU are most hazardous to humans. Paraquat generates free radicals that can attack cell components. DCMU poisons the electron transport complex.

13.7

The hydrolysis of glycerate-l,3-bisphosphate generates 1 mol of ATP. Recall that aerobic respiration is stimulated by relatively high ADP concentrations and inhibited by relatively high ATP concentrations. Any measurable increase in ATP concentration has the effect of depressing aerobic respiration. Also recall that ATP is an inhibitor of PFK-1 and pyruvate kinase, enzymes required to channel carbon skeletons into the citric acid cycle.

Chapter 13: End-of-Chapter Questions

3

a. granum: granum (singular)—the folded portion of the thylakoid membrane; grana (plural)—stacks of folded thylakoid membrane
b. stromal lamellae: a thylakoid membrane segment that interconnects two grana
c. thylakoid lumen: the internal compartment created by the formation of grana
d. PSI: a large membrane-spanning multisubunit protein-pigment complex that energizes and transfers the electrons that are eventually donated to $NADP^+$
e. PSII: a large membrane-spanning multisubunit protein-pigment complex that water molecules and donates energized electrons to electron carriers that eventually reduce PSI

6

a. cytochrome b_6f complex: a multisubunit protein complex in thylakoid membrane that is similar in structure and function to cytochrome bc_1 complex in mitochondrial inner membrane;

delivers electrons donated by plastoquinol to the water-soluble protein plastocyanin

b. CF_o: a membrane-spanning protein complex in the chloroplast ATP synthase that contains a proton channel
c. CF_1: the ATP synthesizing component of the chloroplast ATP synthase
d. LHCII: light harvesting complex II; a trimer of light harvesting proteins bound to chlorophyll a and b molecules; detachable from PSII
e. Mn_4CaO_5: a component of the oxygen-evolving complex; found on the lumenal side of PSII

9

cyanobacteria

12

cytochrome b_6f complex

15

The criteria for sustainable biofuel production are: (1) economically viable production of large quantities of energy, (2) noncompetition with food production (e.g., no use of arable land), and (3) neutral effects on the environment.

18

The three primary photosynthetic pigments are the chlorophylls, the carotenoids, and pheophytin. Chlorophyll a absorbs light energy and is involved in light harvesting. Chlorophyll b is a light harvesting pigment that passes absorbed energy to chlorophyll a. The carotenoids function as light-harvesting pigments or protect against overexcitation and ROS. Pheophytin is an electron transfer molecule in PSII.

21

The chloroplast contains the thylakoid membranes in both appressed (stacked) and unappressed format. The ATPase is oriented in the membrane so that ATP synthesis is always exposed to the stromal compartment. LCHII and PSII are richly concentrated in the appressed regions to maximize light collection and electron transfer. PSI, which should not receive the excitation energy directly from PSII, is physically separated from it in the unappressed regions. Electron replacement of PSI and PSII is mediated by mobile carriers, so physical separation is not a problem.

24

Chloroplasts possess DNA similar to that of modern cyanobacteria as well as prokaryotic-like protein synthesizing machinery. In addition, they multiply by binary fission as do bacteria.

27

a. radiationless decay: an excited molecule decays to its ground state by converting the excitation energy into heat
b. P700: the special pair of chlorophyll a molecules in PSI that absorb light energy at 700 nm
c. P680: the special pair of chlorophyll a molecules in PSII that absorb light energy at 680 nm
d. $\lambda\nu$: In the equation $\lambda\nu = c$, c (the velocity for each type of radiation) is equal to the wavelength λ multiplied by the wave's frequency (ν)
e. UV: ultraviolet light; wavelengths of 400 nm - 100 nm

30

radiationless decay

33

carotenoids

36

Both of the terms P680 and P700 refer to a special pair of chlorophyll a molecules in the reaction centers of PSII and PSI, respectively. P680 in PSII absorbs light at 680 nm. P700 in PSI absorbs light at 700 nm.

39

Conjugation is a system of alternate double and single bonds. When light with sufficient energy strikes the p electrons of a conjugated system (or any double bond), an electron is promoted from the ground state to a higher energy state, referred to as the excited state. Conjugation lowers the energy difference between the ground and the excited state; hence photons of lower energy are capable of achieving this transition.

42

a. F_A: one of several iron-sulfur clusters in photosystem I; transfers electrons that ultimately reduce $NADP^+$
b. Q_BH_2: the reduced form of QB, a plastoquinone that is loosely bound to PSII; once formed it is released into the thylakoid membrane pool of plastoquinones;t two electrons are then donated one at a time to b_6f cytochrome complex
c. plastoquinol: the reduced form of plastoquinone, an electron acceptor in PSII
d. OEC - oxygen-evolving complex; a mechanism that splits water molecules into H^+, electrons and O_2 and then transfers the electrons that will reduce $P680^+$
e. photophosphorylation: captured light energy is transduced into ATP phosphoryl bond energy

45

thylakoid lumen

48

The oxygen evolving system is referred to as a clock because it involves five oxidation-reduction reactions that must be completed in order.

51

Since dinitrophenol is an uncoupler that collapses proton gradients, it would be expected to depress synthesis of ATP, which is required to drive the light-independent reactions.

54

The ratio of energies is:

$$\frac{-hc/_{700}}{-hc/_{1000}} = 1.48$$

57

a. photorespiration: a light-dependent process occurring in plant cells actively engaged in photosynthesis that consumes oxygen and liberates carbon dioxide
b. CA1P: 2-carboxyarabinitol-1-phosphate; a competitive inhibitor of rubisco
c. reductive pentose phosphate cycle - the Calvin cycle; also referred to as light-independent reactions
d. phytochrome: a protein that possesses a red light-sensitive chromophore; activated phytochrome triggers signal transduction pathways that mediate plant responses to light
e. stomata: pores on the surface of leaves; when open CO_2, O_2, and H_2O vapor can readily diffuse down the concentration gradients between the leaf's interior and the external environment

60

glyceraldehyde-3-phosphate

63

malate

66

Carbon dioxide fixation occurs in the stroma of chloroplasts.

69

If sufficient carbon dioxide is already present to saturate all of the ribulose-1,5-bisphosphate carboxylase molecules, the presence of additional carbon dioxide molecules will not increase the rate of photosynthesis. Low light levels depress photosynthesis even with increased carbon dioxide levels.

72

Oxygen is released only after several bursts of light because the oxygen-evolving complex has five oxidation states, i.e., the removal of four electrons each from two water molecules requires the absorption of 8 photons.

75

a. glycerate-3-phosphate

Chapter 14: In-Chapter Questions

14.1

a. CH3NH2
b. NH3
c. CH3CH3

14.2

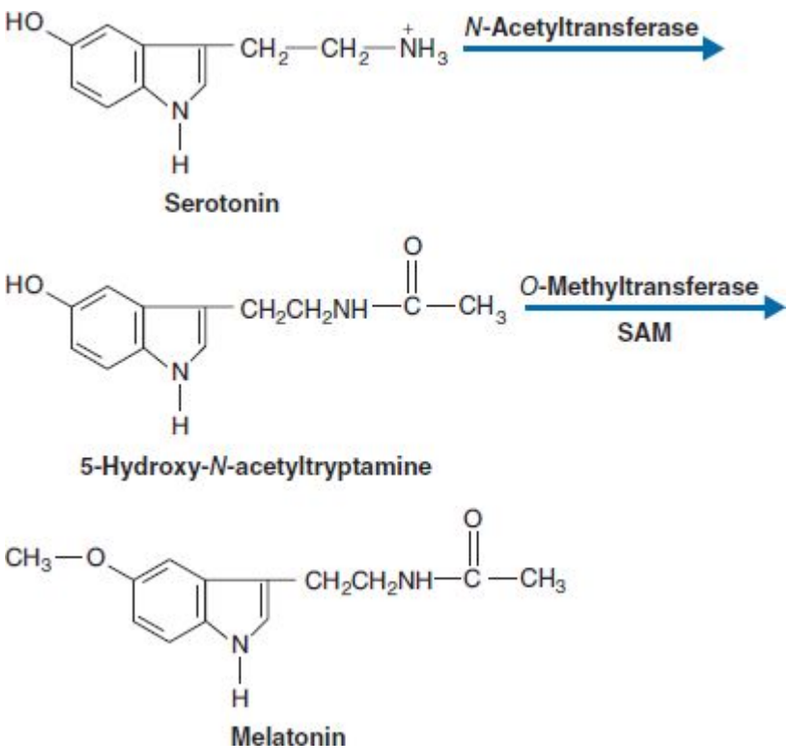
As their names suggest, hemoglobin and leghemoglobin are proteins in the globin superfamily. Recall that hemoglobin is an oxygen transport protein that contains a heme group, which binds reversibly with O_2. The heme in leghemoglobin also binds to O_2. The function of leghemoglobin, the sequestration of oxygen molecules, can be deduced from the irreversible inactivation of the

nitrogenase complex in root nodules by O_2.

14.3

Because of its close structural similarity to folic acid, methotrexate is a competitive inhibitor of the enzyme dihydrofolate reductase. (Recall that this enzyme converts folic acid to its biologically active form, THF.) Rapidly dividing cells require large amounts of folic acid. Methotrexate prevents the synthesis of THF, the one-carbon carrier required in nucleotide and amino acid synthesis. It is therefore toxic to rapidly dividing cells, especially those of certain tumors and normal cells that divide frequently such as hair and GI tract cells.

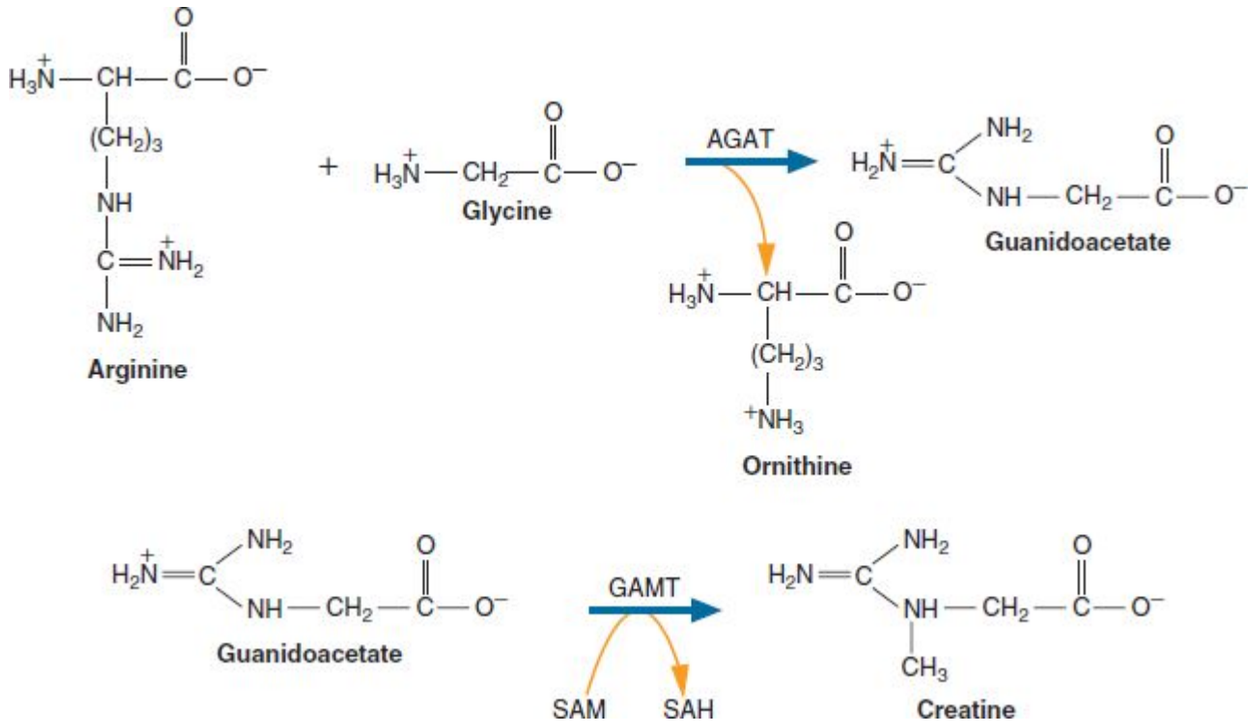
14.4



14.5

The reaction sequence is as follows:

Note that the secondary amino nitrogen alkylates more easily than the primary amino nitrogen in the guanidoacetate molecule.



Chapter 14: End-of-Chapter Questions

3

Nitrogen assimilation

6

While nitrogen incorporation into organic molecules is thermodynamically favored, the conditions in the biosphere (temperature, pressure, and pH) are such that the reactions have low kinetic probability. Only a few organisms possess the biomolecular machinery to surpass the energy barrier to nitrogen compound synthesis.

9

When plants die, organic nitrogen is mineralized or converted through the actions of numerous types of microbes into NH_3, NO_3^-, NO_2^-, (nitrite), and eventually N_2.

12

The source of the hydrogen gas produced by the nitrogen reductase system is the same as the source of electrons: NAD(P)H.

15

Nitrogen fixation is highly endergonic. One possibility of a naturally occurring method of fixing nitrogen is via lightning. The tremendous amount of energy in a lightning bolt can split N_2 molecules, allowing them to react with oxygen to form nitric oxide. Nitric oxide dissolves in, and/or reacts with water in the atmosphere to form nitrates that eventually “drop” to the earth as rain.

18

a. glutamate family: a group of amino acids that include glutamate, glutamine, proline and

arginine that are derived from α-ketoglutarate

b. THF: tetrahydrofolate, the biologically active form of folic acid; carrier of methyl, methylene, methenyl, and formyl groups in one-carbon metabolism
c. AST: aspartate transaminase; catalyzes the reversible reaction in which the amino group of glutamate is transferred to oxaloacetate
d. GOT: glutamic oxaloacetic transaminase; an alternate name for AST
e. shikimate pathway: pathway that forms benzene rings of the aromatic amino acids

21

tyrosine

24

leucine, isoleucine, and valine

27

a. Glutamine is produced in the following reactions:

$$\alpha\text{-Ketoglutarate} + NADH + NH_4^+ + H^+ \rightarrow \text{Glutamate}$$

$$\text{Glutamate} + NH_4^+ + ATP \rightarrow \text{Glutamine} + ADP + P_i$$

b. Serine is produced in the following series of reactions:

$$\text{Glycerate-3-phosphate} + NAD^+ \rightarrow \text{3-Phosphohydroxypyruvate} + NADH + H^+$$

$$\text{3-Phosphohydroxypyruvate} + \text{Glutamate} \rightarrow \text{3-Phosphoserine} + \alpha\text{-Ketoglutarate}$$

$$\text{3-Phosphoserine} + H_2O \rightarrow \text{Serine} + P_i$$

c. Arginine is produced in the following series of reactions:

$$\text{Glutamate} + \text{Acetyl-CoA} \rightarrow \text{N-acetylglutamate} + \text{CoASH}$$

$$\text{N-acetylglutamate} \rightarrow \rightarrow \rightarrow \text{L-ornithine}$$

$$\text{L-ornithine} \rightarrow \rightarrow \text{L-Arginine}$$

d. Glycine is produced in the following series of reactions:

$$\text{Glycerate-3-phosphate} + NAD^+ \rightarrow \text{3-Phosphohydroxypyruvate} + NADH + H^+$$

$$\text{3-Phosphohydroxypyruvate} + \text{Glutamate} \rightarrow \text{3-Phosphoserine} + \alpha\text{-Ketoglutarate}$$

$$\text{3-Phosphoserine} + H_2O \rightarrow \text{Serine} + P_i$$

$$\text{Serine} + THF \rightarrow \text{Glycine} + N^5, N^{10}\text{-methylene THF} + H_2O$$

e. $\text{Serine} + \text{L-Homocysteine} \rightarrow \text{Cystathione} + H_2O$

$$\text{Cystathione} + H_2O \rightarrow \text{Cysteine} + \alpha\text{-Ketoglutarate} + NH_4^+$$

30

Arginine is normally synthesized by the urea cycle. In small children, the urea cycle is not fully functional. Consequently, arginine must be obtained from external sources.

33

In ping-pong reactions, the first substrate must leave the active site before the second can enter. In the reaction of alanine with α-ketoglutarate to produce pyruvate and glutamate the following steps take place: (1) the alanine enters the active site and its amino group is transferred to pyridoxal phosphate, (2) water enters the reaction site and hydrolyses the Schiff base to produce pyridoxamine phosphate and pyruvate, (3) pyruvate diffuses from the active site, (4) α-ketoglutarate, the second substrate, enters the reaction site and forms a Schiff base with the pyridoxamine phosphate, (5) water hydrolyzes the Schiff base to give pyridoxal phosphate and glutamate, and (6) glutamate diffuses out of the active site.

36

a. SAM (S-adenosylmethionine): the major methyl donor in one-carbon metabolism; contains an activated methyl thioether group
b. SAH (S-adenosylhomocysteine): the product of reactions in which SAM is the methyl donor
c. polyamine: polycationic molecules that bind to DNA: allows DNA to be compressed into chromosomes
d. methotrexate: amethopterin; structural analogue of folate
e. glutathione: γ-glutamylcysteinylglycine; the most common intracellular reducing agent

39

a. anti-nucleoside: the conformation of a nucleoside in which the base is rotated outward away from the 6′-CH_2OH group of the ribose moiety
b. HGPRT: Hypoxanthine-guaninephosphoribosyltransferase; catalyzes nucleotide synthesis using PRPP and either hypoxanthine or guanine
c. Lesch–Nyhan syndrome: a fatal X-linked disease caused by HGPRT (hypoxanthine guaninephosphoribosyl transferase) deficiency; excessive production of uric acid causes severe neurological symptoms
d. purine salvage: purine bases reconverted into nucleotides
e. gout: a condition in which high blood uric acid concentrations result in the accumulation of sodium urate crystals in joints, especially those in feet

42

purine

45

glutathione

48

Orotic aciduria is so named because of the large amounts of orotate, an intermediate in the synthesis of UMP, that appear in the urine of affected individuals. Treatment consists of the

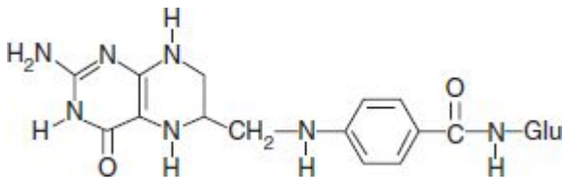
administration of pyrimidine nucleotides, which inhibit orotate synthesis.

51

Pyrimidine nucleosides occur predominantly in the *anti*-conformation. Steric hindrance between the pentose sugar and the carbonyl oxygen at C-2 of the pyrimidine ring prevents free rotation around the N-glycosidic bond. Perhaps surprisingly, the larger purine does not have similar steric interactions with the pentose. Even though the pyrimidine has only one ring, its carbonyl group interferes sterically with the pentose ring to hinder the formation of the *syn*-conformation. In contrast, the shape and functionality of the purine rings are such that steric interactions with the pentose are minimal when in the syn conformation, in spite of the overall larger size of the purines.

54

The biologically active form of folic acid, referred to as tetrahydrofolate or THF, is shown below. It is formed by the reduction of folic acid with NADPH, in two reactions catalyzed by dihydrofolate reductase.



57

The hydroxyl radical extracts a hydrogen atom from glutathione to form water in a reaction catalyzed by glutathione peroxidase.

$$GSH + \cdot OH \rightarrow GS\cdot + H_2O$$

Two molecules of oxidized glutathione then react to form GSSG. GSH is regenerated from GSSG by glutathione reductase using NADPH as the source of hydrogen atoms.

$$GSSG + NADPH + H+ \rightarrow 2GSH + NADP^+$$

60

Running is a physical activity that requires the rapid metabolism of both fatty acids and glucose. The most rapidly utilized source of glucose is blood glucose. Although certain amino acids can be absorbed easily in the intestine and used as substrates in gluconeogenesis in the liver, this process is slower than the immediate absorption of glucose into the blood.

63

Nitric oxide synthesis from nitrate begins with the reduction of nitrate to yield NH_4^+. Glutamate dehydrogenase catalyzes the NADH-requiring reaction of α-ketoglutarate and NH_4^+. Glutamate is a precursor in the synthesis of L-ornithine, an intermediate in the urea cycle. Ornithine is converted to arginine in three reactions. Nitric oxide is the product of a two-step oxidation of arginine to form citrulline catalyzed by the O_2 and NADPH-requiring enzyme nitric oxide synthase. Alternatively, nitrate can be converted to nitrite by microorganisms in the large intestine. Nitrite can then be reduced to nitric oxide by several redox enzymes (e.g., xanthine oxidase).

66

d. GSH

Chapter 15: In-Chapter Questions

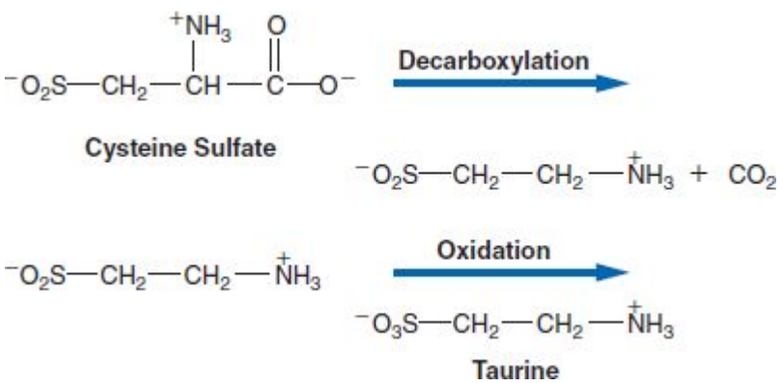
15.1

In newborn animals arginine will be an essential amino acid if the urea cycle is not yet fully functional.

15.2

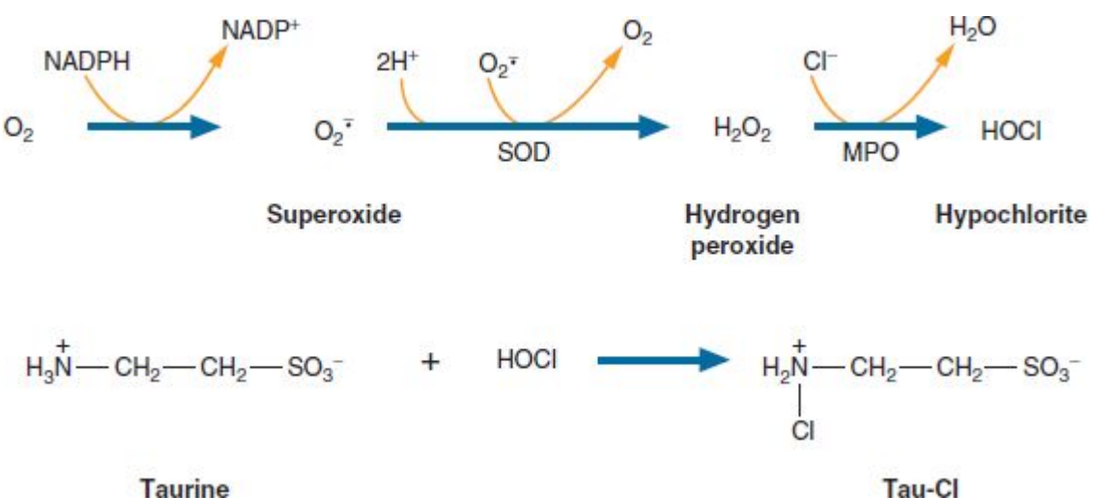
Certain intestinal bacteria can release ammonia from urea molecules that diffuse across the membrane into the intestinal lumen. Treatment with antibiotics kills these organisms, thereby reducing blood ammonia concentration.

15.3



15.4

The reactions are as follows:



15.5

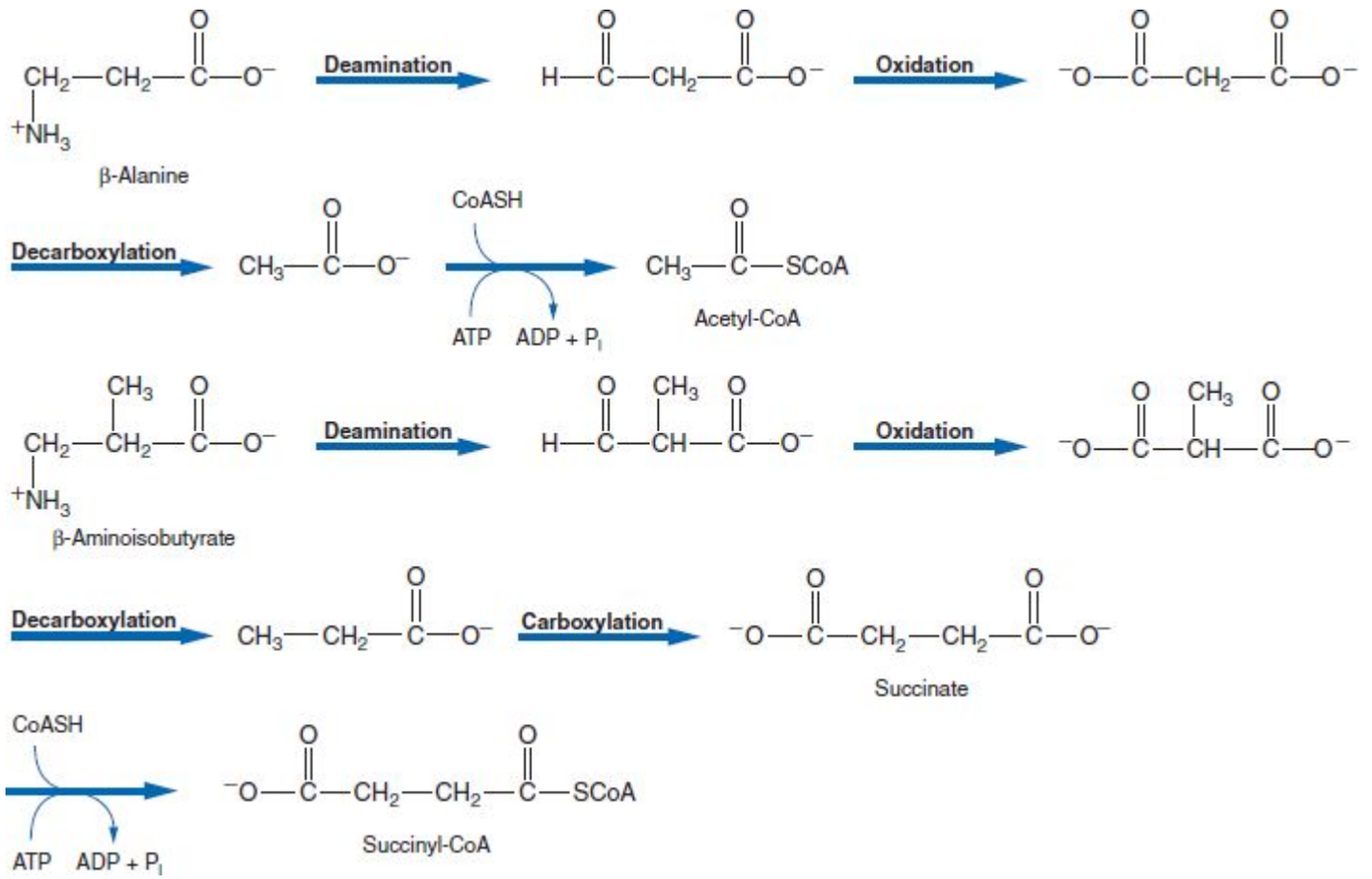
Gout is caused by high levels of uric acid. Animals that do not suffer from gout possess the enzyme urate oxidase, which converts uric acid to allantoin. Unlike uric acid, which is relatively insoluble in blood, allantoin readily dissolves and is easily excreted.

15.6

a. Urea is formed from ammonia, CO_2, and aspartate in the urea cycle.
b. Uric acid is the oxidation product of purines.
c. β-Alanine is produced in the degradative pathway of pyrimidines.

15.7

Suggested catabolic reactions of β-alanine and β-aminoisobu tyrate:



Chapter 15: End-of-Chapter Questions

3

a. chaperone-mediated autophagy: a receptor-mediated process in which specific proteins that are bound to a chaperone complex are unfolded and then translocated to a lysosome where they are degraded
b. microautophagy: a degradative process in which small amounts of cytoplasm are directly engulfed by lysosomes
c. macroautophagy - a lysosomal pathway for bulk degradation of cytoplasmic components; also referred to as autophagy
d. autophagosome: a structure formed during autophagy; an isolation membrane, which sequesters cytoplasmic components and then seals itself
e. isolation membrane: the intracellular membrane that forms at the beginning of autophagy; sequesters cytoplasmic components that will subsequently be degraded in an autophagosome

6

degron

9

amphisome

12

Structural features that mark proteins for destruction include certain N-terminal residues (Leu, Phe, Tyr, Trp and Ile) and motifs such as the PEST sequence (Pro, Glu, Ser, and Thr).

15

Lid subunits of the proteasome bind directly to the a-ring of the core particle; they participate in substrate selection.

18

Ammonotelic organisms release their nitrogenous waste as ammonia, which they release in surrounding water. Uricotelic organisms for which water conservation is a significant issue, excrete uric acid, which is less toxic than ammonia than ammonia and can be released in a nearly solid form.

21

One of the problems of aging cells is that over time the body's systems lose efficiency. Gradually autophagic processes become less effective, with the result that autophagolysosomes containing indigestible debris begin accumulating. Gradually, debris accumulation compromises cell function.

24

a. CBS: cystationine-β-synthase; an enzyme in the transsulfuration pathway that converts homocysteine and serine into cystathionine
b. CSE: γ-cystathionase; the enzyme in the transsulfuation pathwaythat converts cystathionine to cysteine, α-ketobutyrate and NH_4^+
c. transsulfuration pathway: a biochemical pathway involving the interconversion of cysteine and homocysteine; also a source of sulfate for the sulfate groups of proteoglycans and sulfate-containing glycolipids
d. bacterial urease: an intestinal enzyme that hydrolyzes urea in the bloodstream
e. serine dehydratase: a hepatic pyridoxal-requiring enzyme that converts serine to pyruvate and NH_4^+

27

glucogenic

30

hyperammonemia

33

a. Ketogenic
b. Ketogenic
c. Glycogenic

d. Glycogenic
e. Glycogenic
f. Both.

36

In the muscle, pyruvate undergoes a transamination reaction and is converted to alanine. Alanine is then transferred to the liver, where it is reconverted to pyruvate bytransferring its amino group to α-ketoglutarate to form glutamate. Glutamate is then oxidatively deaminated to form α-ketoglutarate and ammonia. The pyruvate and α-ketoglutarate are degraded by the citric acid cycle.

39

The branched chain amino acids (leu, ile, and val) are metabolized primarily in muscle tissue, there they are principally used to synthesize nonessential amino acids.

42

Nitrogen enters the urea cycle as NH_4^+ and aspartate. The two nitrogen atoms in urea most likely originated in the α-amino groups of amino acids. The pathways from α-amino acids to ammonium and aspartate are outlined below, with (1) and (2) providing most of the nitrogen in the urea cycle. (1) From muscle cell amino acids to NH_4^+ in the liver: In muscle cells, a transamination reaction with α-ketoglutarate produces glutamate (and an α-keto acid). Glutamate in muscle cells may transfer its amino group to pyruvate via another transamination to produce alanine, or it may be oxidatively deaminated to generate NH_4^+, which reacts with another glutamate to form glutamine. Alanine and glutamine transport the nitrogen atom from α-amino acids to the liver, where glutamate is regenerated (a)from alanine via transamination to α-ketoglutarate, or (b) from glutamine via hydrolysis of the amide, which also forms a free NH_4^+. The glutamate produced from both alanine and glutamine is oxidatively deaminated to form NH_4^+. (2) From liver cell amino acids to aspartate: An α-amino acid and oxaloacetate react via transamination to form an α-keto acid and aspartate. (3) From serine and threonine to NH_4^+ in the liver: Since Ser and Thr cannot undergo transamination, they are deaminated by serine dehydratase and threonine dehydratase. (4) From urea present in the blood to NH_4^+ in the liver: Urea in the bloodstream diffuses into the intestinal lumen, where it is hydrolyzed by bacteria with urease to form ammonia, which diffuses back into the blood for transport to the liver. (5) From amino acids in the liver and kidney to NH_4^+ via the action of L-amino acid oxidases. (6) From the C-6 amino group of adenine to NH_4^+ via the action of adenosine deaminase.

45

As the concentration of glutamate (as well as its deamination product ammonia) rises, the enzyme N-acetylglutamate synthase catalyzes the synthesis of N-acetylglutamate, an activator of carbamoyl phosphate synthetases I. The latter enzyme catalyzes the first committed step in urea synthesis.

48

a. Methylene (–CH2–) groups
b. Methylene groups
c. Methyl groups

d. Methyl groups

51

The α-keto acids react with ammonia to produce amino acids thus sparing α-ketoglutarate, a citric acid cycle intermediate needed to generate energy.

54

a. uric acid: the end waste product of purine nucleotide catabolism in some animals including humans
b. allantoin: the excretory product of purine nucleotide metabolism in many mammals; formed from uric acid by uric oxidase
c. allantoate: the excretory product of purine metabolism in bony fish; formed from allantoin by allantoicase
d. urate oxidase: the enzyme that converts uric acid to allantoin
e. allantoicase: the enzyme that splits allantoic acid to glyoxylate and urea

57

humoral

60

pyrimidine

63

In the absence of functional PNP, high levels of purine nucleotides (e.g., dGTP) apparently impair T cells.

66

Because of the structural similarities to purine, caffeine is converted to a variety of derivatives by xanthine oxidase. Examples include 1-methyluric acid and 7-methylxanthine.

69

Primates lack urate oxidase, which converts uric acid to allantoin, a more soluble molecule. As a result uric acid crystals can build up and gout results.

72

d. glucose–alanine cycle

75

c. urease

Chapter 16: In-Chapter Questions

16.1

Several series of signal transduction components are illustrated in **Figure 8.21**. A prominent example series consist of glucagon (primary signal), glucagon receptor (receptor), adenylate cyclase (transducer), and activated protein kinase (response).

16.2

cAMP molecules would be produced by the binding of a single hormone molecule before it diffused away from the receptor site. The amplification factor for the hormone molecule is 350; that is, 350 cAMP molecules are produced for every hormone-receptor binding event.

16.3

Approximately 100,000 (or 10^5) molecules of target molecule (E_R) can be activated by a single molecule of hormone. cAMP is generated from ATP by adenylate cyclase when a hormone molecule binds to its receptor. The interaction between the receptor and adenylate cyclase is mediated by a G protein, G_s. As a consequence of hormone binding and the resulting conformational change, the receptor interacts with a nearby G_s protein. As G_s binds to the receptor, GDP dissociates. Then the binding of GTP to G_s allows one of its subunits to interact and stimulate adenylate cyclase, thus initiating cAMP synthesis. cAMP must break down quickly so the signaling mechanism can be precisely controlled.

16.4

The inhibition of GTP hydrolysis causes the subunit of G_s protein to continue activating adenylate cyclase. In intestinal cells, this enzyme activity opens chloride channels, causing loss of large amounts of chloride ions and water. The massive diarrhea caused by this process quickly leads to serious dehydration and electrolyte loss.

16.5

Both DAG and phorbol esters promote the activity of protein kinase C, which promotes cell growth and division. Phorbol esters provide initiated cells with a sustained growth advantage over normal cells. This condition is an early stage in carcinogenesis.

16.6

The high blood glucose levels in untreated diabetics result in the loss of increasingly large amounts of glucose along with water in the urine, a condition that causes dehydration. In one form of the absence of usable glucose, the body rapidly degrades fats and proteins to generate energy, Hence Aretaeus's observation that in this disease excessive weight loss and excessive urination arc related.

16.7

Long-term fasting or low-calorie diets arc interpreted by the brain as starvation. The brain responds by lowering the body's BMR. The majority of the energy is derived from fatty acid oxidation. The glucose needed for glucose-dependent tissues is generated via gluconeogenesis at the expense of muscle protein.

16.8

As blood glucose and insulin levels in blood drop back to normal, glucagon is released from the pancreas. Glucagon acts on the liver to prevent hypoglycemia by promoting glycogenolysis and gluconeogenesis. Glucagon stimulates glycogenolysis by triggering the synthesis of cAMP, which in turn initiates a cascade of reactions that lead to the activation of glycogen phosphorylase. Increased lipolysis, hydrolysis of fat molecules, provides glycerol molecules that are substrates for gluconeogenesis.

16.9

The difference in insulin release levels that result from glucose consumption and glucose injection involves the actions of two hormones referred to as the incretins. Glucose consumption triggers the release of the incretins (glucagon-like peptide-1 and glucose-dependent insulinotropic peptide) by intestinal cells. Both incretins cause an increased release of insulin by pancreatic b-cells. A glucose injection bypasses the intestine with the result that less insulin is released.

Chapter 16: End-of-Chapter Questions

3

hypertension, kidney disease, obesity, diabetes, and cardiovascular diseases

6

endocrine

9

The nervous system, a network of neurons and support cells, provides the body with a rapid and efficient mechanism for acquiring internal and external information. The endocrine system consists of specialized glands that produce hormones that regulate metabolism. Interactions between the nervous system and the endocrine system released hormones are responsible for integrating the body's metabolic processes to maintain homeostasis (a stable internal environment). Sensor signals are received by the brain's hypothalamus which integrates nervous and endocrine signaling mechanisms.

12

The clinical trials were designed to determine the relationship between hypertension and high uric acid levels, without other factors (previous hypertension damage) complicating the study data. Younger patients are more likely to be better candidates for these studies due to the decreased likelihood of kidney damage or other diseases related to hypertension.

15

a. adenylate cyclase: the enzyme that converts ATP to cyclic AMP (cAMP), the second messenger molecule
b. $G\alpha_s$: a G protein α subunit that stimulates the activity of adenylate cyclase
c. $G\beta\gamma$: a dimer that binds to and inhibits a $G\alpha$ subunit; may also play a role in downstream effector signaling
d. PKA: protein kinase A; activated by cAMP; modified the activity of numerous enzymes via phosphorylation reactions
e. CREB: cAMP response element binding protein; a transcription factor involved in the regulation of several enzymes; activated by PKA

18

a. phosphatidylinositol cycle: PIP_2 is cleaved by phospholipase C to form DAG and IP_3; mediates, primarily via G_q proteins, the actions of hormones and growth factors
b. PIP_2: phosphatidylinositol-4,5-bisphosphate; a minor membrane phospholipid that is cleaved by phospholipase C

c. IP_3: inositol-1,4,5-triphosphate; a second messenger molecule in the phosphatidylinositol pathway that binds to IP_3 receptor, a calcium channel
d. DAG: diacylglycerol; a second messenger molecule in the phosphatidylinositol pathway that activates protein kinase C
e. PKC: protein kinase C; phosphorylates key cell signal regulatory enzymes

21

a. glucosuria: the presence of glucose in the urine
b. osmotic diuresis: a process in which high levels of solutes in the urinary filtrate cause excessive loss of water and electrolytes
c. ketoacidosis: a form of acidosis caused by the excessive accumulation of ketone bodies; occurs in type II diabetes as the result of unrestrained fatty acid oxidation
d. ketosis: accumulation of ketone bodies in blood and tissues
e. hyperinsulinemia: higher than normal blood levels of insulin; caused by insulin resistance

24

downregulation

27

insulin

30

ketoacidosis

33

second messengers

36

Anabolic steroid hormones change the expression of a specific set of genes that code for proteins (e.g., enzymes) that increase protein synthesis in skeletal muscle (among other metabolic changes).

39

Phorbol esters, found in croton oil, activate protein kinase C, an action that stimulates cell growth and division. However, unlike DAG, the molecule that they mimic, phorbol esters continue to activate protein kinase C for a prolonged time. This circumstance provides the affected cell with an advantage over unstimulated cells. Phorbol esters may transform a cell previously exposed to a carcinogenic initiating event into a cancerous cell whose unrestrained proliferation creates a tumor.

42

The most common sites on proteins that are phosphorylated during signal transduction cascades are the hydroxyl groups of serine and tyrosine residues.

45

Flies are well known for their preference for sugar. If urine has detectable sugar in it, flies will be attracted.

48

The steroid molecule is covalently bound to the matrix in a chromatographic column. The extract suspected of containing a steroid binding protein is then passed through the column. Any proteins remaining on the column are eluted by changing the salt concentration of the eluting buffer. After isolation and purification such proteins can be examined specifically for binding activity to the steroid.

51

Excessively high doses of insulin will cause the insulin target cells in muscle and adipose tissue to remove so much glucose from the bloodstream that the brain which exclusively relies on glucose for energy, may be unable to sustain its processes, resulting in the death of brain cells.

54

a. postprandial: the phase in the feeding-fasting cycle immediately after the meal; blood nutrient levels are relatively high
b. postabsorption: the phase in the feeding-fasting cycle in which nutrient levels are low
c. chylomicron remnants: chylomicrons after about 90% of the triacylglycerols have been removed by lipoprotein lipase
d. feeding behavior: a complex mechanism by which animals, including humans, seek out and consume food
e. satiety: the sensation of feeling full; the absence of hunger

57

nucleus tractus solitarius (NTS)

60

The cardiovascular system is damaged by fructose consumption by two processes, both of which stimulate atherosclerosis: (1) dyslipidemia, a consequence of hepatic fructose metabolism and (2) glycation reactions that occur about seven times as often with fructose as those of glucose and lead to AGE (advanced glycation end product) formation.

63

Diets high in fructose may result in obesity, blood dyslipidemia, hyperuricemia, hypertension, and insulin resistance. Metabolic syndrome is a term used to describe this cluster of disorders.

66

Appetite regulation in humans is a set of complex and robust mechanisms that involve several areas of the brain, such as the hypothalamus. In response to calorie restriction which is interpreted as starvation, the appetite centers of the brain respond by stimulating appetite, lowering the body’s BMR (to conserve energy), and lowering energy expenditures (resulting in lethargy). As a result of these and other hormone- and peptide-triggered responses, continued achievement of weight loss and maintenance of a reduced weight over time becomes very difficult, if not impossible.

69

Exercise promotes insulin-independent glucose uptake by muscle cells, which facilitates blood glucose control.

72

a. cholecystokinin

75

c. ANF

Chapter 17: In-Chapter Questions

17.1

The cytosine-guanine base pair with its three hydrogen bonds is more stable than the adenine-thymine base pair. The more CG bp there are, the more stable the DNA molecule. Structure b, with the fewest CG bp, will therefore denature first.

17.2

a. Ethanol will disrupt the hydrogen bonding in the base pairs and denature the DNA.
b. Heat, which easily disrupts hydrogen bonds, will cause DNA chains to separate and denature.
c. Dimethylsulfate is an alkylating agent that can cause transversion and transition mutations.
d. Nitrous acid deaminates bases.
e. Quinaerine is an intercalating agent that can cause frame shift mutations.

17.3

The brain is especially sensitive to oxidative stress because it uses a greater proportion of oxygen than other tissues. Consequently, the chance of oxidative damage is also high. In addition, when most types of brain cells are irreversibly damaged by ROS, they cannot be replaced. In addition to hydroxyl radicals, other ROS that can contribute to oxidative stress in the brain include superoxide, hydrogen peroxide, and singlet oxygen.

17.4

In A-DNA, the dehydrated form of DNA, the base pairs are no longer at right angles to the helical axis. Instead, they tilt 20 degrees away from the horizontal as compared to B-DNA. The distance between adjacent base pairs is slightly reduced, with 11 bp per helical turn instead of the 10.4 bp that occurs in the B-form. Each turn of the double helix of A-DNA occurs in 2.5 nm instead of the 3.4 nm of B-DNA. The diameters of A-DNA and B-DNA are 2.6 and 2.4 nm, respectively. The significance of A-DNA is unclear. It has been observed that its overall appearance resembles that of RNA duplexes and the RNA-DNA hybrids that form during transcription.

With a diameter of 1.8 nm, Z-DNA is considerably slimmer than B-DNA. It is twisted into a left-handed spiral with 12 bp per turn, each of which occurs in 4.5 nm instead of the 3.4 observed in B-DNA. Segments with alternating purine and pyrimidine bases are most likely to adopt the Z-DNA configuration. In Z-DNA the bases stack in a left-handed, staggered pattern that gives this form its flattened, non-grooved surface and zigzag appearance. The significance of Z-DNA is unresolved.

H-DNA (triple helix) segments can form when a poly-purine sequence is hydrogen-bonded to a polypyrimidine sequence. H-DNA, which has been observed to form under low pH conditions, is made possible by nonconventional, Hoogsteen base pairing. H-DNA may play a role in recombination.

17.5

The genome is the total set of DNA-encoded genetic information in an organism. A chromosome is a DNA molecule, usually complexed with certain proteins. Chromatin is the partially decondensed form of eukaryotic chromosomes. Nucleosomes are the repeating structural units of eukaryotic chromosomes formed by the interaction of DNA with the histones. A gene is a DNA sequence that codes for a polypeptide or an RNA molecule.

17.6

The deletion of gene C in (a) results in a deficit of a critical gene in limb development. In addition, the TAD deletion results in the aberrant contact between the enhancer with gene D, causing the inappropriate transcription of this gene. In (b) the removal of the TAD between the enhancer sequence and gene B promotes the inappropriate transcription of this gene, a circumstance that results in a different form of limb malformation than that observed in (a).

17.7

The genomes of prokaryotes are substantially smaller than those of eukaryotes. For example, the genome sizes of *E. coli* and humans are 4.6 and3200 Mb (haploid), respectively. Prokaryotic genomes are compact and continuous; that is, there are few, if any, noncoding DNA sequences. In contrast, eukaryotic DNA contains enormous amounts of noncoding sequences. Other distinguishing features of prokaryotic and eukaryotic DNA are the linkages of genes into operons in prokaryotes and intervening sequences in eukaryotic genes.

17.8

The antisense DNA sequence is 3′-CGTAAGCT TAACGTCTGAGGACGTTAAGCCGTTA-5′; the mRNA sequence is 3′-CGUAAGCUUAACGUCUGAGGACGU UAAGCCGUUA-5′, The antisense RNA sequence is 3′-GCAUUCGAAUUGCAGACUCCUGCAAUUCGGCAA U-5′.

17.9

In the original central dogma, the flow of genetic information is in one direction only, that is, from DNA to the RNA molecules, which then direct protein synthesis. The altered diagram indicates that the RNA genome of some viruses can replicate their RNA genomes (using a viral enzyme activity referred to as RNA-directed RNA polymerase) or undergo reverse transcription (i.e., synthesize DNA from an RNA sequence).

Chapter 17: End-of-Chapter Questions

3

a. point mutation: a change in a single nucleotide base in a DNA sequence
b. transition mutation: a DNA mutation that involves the substitution of a purine base by a different purine, or the substitution of a pyrimidine by a different pyrimidine
c. transversion mutation: a type of point mutation in which a pyrimidine is substituted for a purine and vice versa
d. silent mutation: a mutation that has no discernable effect on an organism
e. missense mutation: a point mutation that results in an amino acid substitution resulting in a change in a polypeptide’s function

6

a. Chargaff’s rules: a set of rules describing the base composition of DNA; posits the equality of

the concentration of adenine and the thymine and of cytosine and guanine

b. constitutive heterchromatin: sections of DNA in eukaryotes that is permanently highly condensed and transcriptionally silent; ocurrs at centromeres, telomeres, transposons and repetitive sequences
c. bacteriophage: a type of virus that infects bacteria
d. satellite DNA: DNA sequences that are highly repetitive; when genomic DNA is digested and centrifuged, a satellite band forms
e. LINE: long interspersed nuclear elements; non-LTR retrotransposons with lengths greater than 5 kb

9

a. histones: a group of basic proteins found in all eukaryotes that bind to DNA to form nucleosomes
b. heterochromatin: highly condensed chromatin that is transcriptionally inactive; constitutive chromatin is permanently condensed; facultative heterochromatin may be decompressed in specific cell types
c. euchromatin: a less condensed form of chromatin with varying levels of transcriptional activity
d. intergenic sequences: noncoding DNA sequences; includes ncRNA genes, introns, and untranslated regions of protein coding genes
e. tandem repeats: DNA sequences in which multiple copies are arranged next to each other

12

metabolome

15

nonsense mutation

18

intron

21

Nucleosomes are the structural units of chromatin. Each nucleosome consists of a left handed supercoiled DNA segment wound around eight histone molecules.

24

There are approximately 6 million base pairs in a single human cell. Assuming that there are 10^{14} body cells, the total length of the DNA in the human body is approximately 2×10^{11} km. This estimated length is about 1000 time greater than the distance for the earth to the sun. (Note that 1 nm is 10^{-9} m.)

27

Before the publication of the Watson-Crick paper in 1953, research efforts were focused on proving that DNA is the genetic material, and more recently to discover its structure. Beginning with the publication of the Watson-Crick paper, research efforts rapidly shifted to the functional properties of DNA and related processes. This work was eventually referred to as molecular biology.

30

Chargaff's rules do not apply to RNA because RNA is single-stranded.

33

Since the mitochondria's electron transport system is a major source of ROS, it would be expected to possess a substantial antioxidant mechanism. Examples of antioxidant enzymes in mitochondria include superoxide dismutase, catalase, peroxiredoxin/thioredoxin reductase, glutathione peroxidase and glutathione reductase.

36

Ethyl chloride is an alkylating agent that can react with DNA bases to form ethyl derivatives.

39

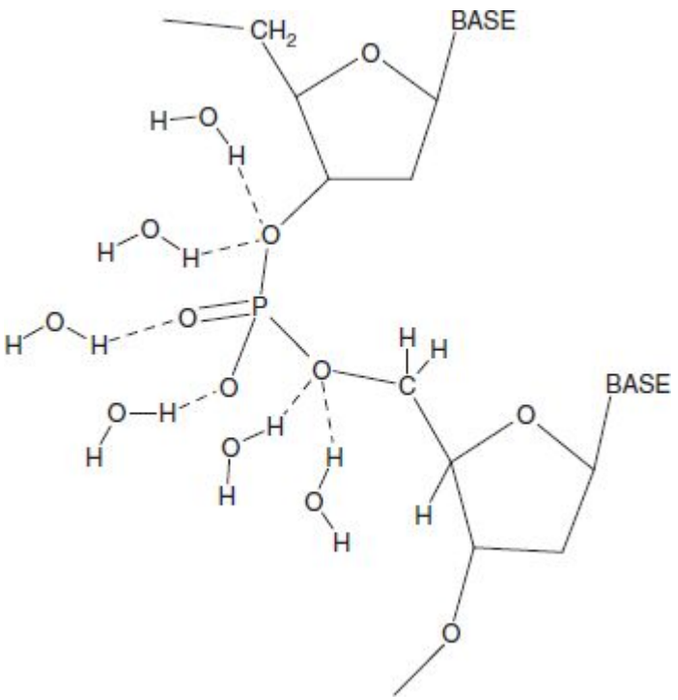
Syncytin-1, a product of the HERV-W family of endogenous retroviral sequences, plays a critical role in the formation of the placenta. Inadequate synthesis of syncytin-1 is one of several factors that contribute to pre-eclampsia.

42

The electron withdrawing effect of the bromine increases the likelihood of enol formation of uracil. This enol mimics the hydrogen bonding pattern of cytosine. Therefore, this base can be paired with guanine.

45

The hydrogen bonds to water molecules formed by the atoms in a phosphodiester linkage are as follows:



48

a. noncoding RNA: types of RNA other than the RNAs involved in protein synthesis (tRNAs, rRNAs and mRNAs) that act as an extensive genome regulatory network

b. miRNA: microRNA, short (22 to 26 nt in length) RNAs involved in gene expression; bind to complementary sequences in the 3′-UTR of target mRNAs
c. siRNA: small interfering RNA; short 21 to 23 nt in length) RNAs involved in RNA interference
d. snoRNA: small nucleolar RNA; RNAs containing 70 to 300 nt that facilitate chemical modifications of rRNA, tRNA and snRNA within the nucleolus
e. lncRNA: long noncoding RNA; functions include transcription and posttranslational modification regulation, cell cycle control, and promoting apoptosis 17

51

a. tiRNA: a 5′- or 3′-tRNA fragment; formed in mammalian cell stress
b. monocistronic: refers to a eukaryotic DNA sequence that codes for one polypeptide
c. snRNPs: small nuclear ribonucleoproteins; with other proteins form spliceosomes
d. antisense DNA strand: the DNA strand that acts as a template for RNA synthesis; the noncoding strand
e. sense DNA strand: the nontranscribed DNA sequense; the coding strand

54

open reading frame

57

The results of this experiment suggest that the methyl donors overrode the hypomethylation effects of BPA. In other words, a nutritious diet can provide some degree of protection against the effects of environmental toxins.

60

According to the histone code hypothesis, the pattern of histone modifications within each DNA sequence regulates gene expression by serving as a platform for the binding of proteins that inhibit or facilitate transcription.

63

The nucleophilic attack of the 3′-hydroxyl oxygen on the phosphorous of the α-phosphate group of a deoxyribonucleotide yields a proton.

66

Prokaryotic and eukaryotic mRNA differences include the following. Many prokaryotic mRNAs are polycistronic, whereas eukaryotic mRNAs are monocistronic. Prokaryotic and eukaryotic mRNAs are processed differently. Prokaryotic mRNAs are translated during or immediately after their synthesis. Eukaryotic mRNAs are modified extensively and are then subject to several levels of regulation. Examples include regulation by miRNAs and complex mechanisms that link mRNA translation to cell energy levels or ER stress.

69

a. lytic cycle: a viral infective process in which the host cell is destroyed
b. prophage: a viral genome that is integrated into the host cell’s DNA
c. transduction: a lysing bacterial cell releases virions that contain some bacterial DNA along with phage DNA; bacterial DNA is then introduced into the genome of a new host genome
d. lysogeny: a condition in which a phage genome is integrated into a bacterial chromosome
e. nucleocapsid: a protein complex that surround a viral genome, especially that of an enveloped

virus

72

nucleocapsid

75

If two types of virus infect the same host cell, there is the possibility of a viral particle containing genetic material of both viruses.

78

Measles has been afflicting humans in Europe, North Africa and the Middle East forcenturies. Individuals with no resistance to the virus died; those with immune systems able to fend off measles infection survived. After centuries of measles outbreaks, many Europeans had inherited some resistance to the measles virus (i.e. some individuals might be infected, but their immune systems could mount an eventually successful defense). Prior to the European discovery of the New World, its residents had never had any exposure to measles. As a result, millions of native people died from infections of measles and other viral diseases such as smallpox.

81

a. nucleosome

84

b. histone covalent modifications

Chapter 18: In-Chapter Questions

18.1

Briefly, prokaryotic DNA replication consists of DNA unwinding, RNA primer formation, DNA synthesis catalyzed by DNA polymerase and the joining of Okazaki fragments by DNA ligase. Prokaryotic DNA replication differs from the eukaryotic process in that prokaryotic replication is faster, the Okazaki fragments are longer and there is usually only one origin of replication per chromosome (eukaryotes have many per chromosome).

18.2

In excision repair short damaged sequences (e.g., thymine dimers) are excised and replaced with correct sequences. After an endonuclease deletes the damaged single-stranded sequence, a DNA polymerase activity synthesizes a replacement sequence using the undamaged strand as a template. In photoreactivation repair a photoreactivating enzyme uses light energy to repair pyrimidine dimers. In recombinational repair damaged sequences are deleted. Repair involves an exchange of an appropriate segment of the homologous DNA molecule.

18.3

When antibiotics are used in large quantities, the bacterial cells that possess resistance genes (acquired through spontaneous mutations or through intermicrobial DNA transfer mechanisms such as conjugation, transduction, and transformation) survive and even flourish. Because of antibiotic use, which acts as a selection pressure, resistant organisms (once only a minor constituent of a microbial population) become the dominant cells in their ecological niche.

18.4

Most gene duplications are apparently a consequence of accidents during genetic recombination. Examples of possible causes of gene duplication are unequal crossing over during synapses and transposition. After a gene has been duplicated, random mutations and genetic recombination may introduce variations.

18.5

Because phytochrome has been demonstrated to mediate numerous light-induced plant processes, it appears reasonable to assume that it does so in part by interacting with light-response elements (LRE) in plant cell genomes. Presumably, phytochrome influences gene expression by binding, either alone or as part of a complex, to various LREs when its chromophore is activated by light.

Chapter 18: End-of-Chapter Questions

3

a. Okazaki fragment: any of a series of deoxyribonucleotide segments that are formed during discontinuous replication of one DNA strand as the other strand is continuously replicated
b. ter region: a segment of the *E. coli* chromosome that contains DNA replication termination sequences
c. tus protein: a protein that when bound to a ter sequence facilitates DNA replication termination
d. transcription preinitiation complex: PIC; a eukaryotic complex of proteins that is necessary for the transcription of protein-coding genes
e. ORC: origin of replication complex; a protein complex that binds to the DNA replication origin during the initiation phase of eukaryotic DNA synthesis; contains analogs of the protein DnaA

6

a. nonreplicative transposition: transposition of a DNA segment occurs in cut and paste mechanism; a DNA sequence is removed from a donor site and then spliced into a target site without duplication of the transposed sequence
b. transposable element: a DNA sequence that excises itself and then inserts at another site
c. bacterial transformation: a DNA fragment of one bacterial cell and then enters another bacterial cell where it is incorporated into its genome
d. transduction: the transfer of DNA segments between bacteria by bacteriophages
e. conjugation: unconventional sexual mating between bacterial cells; a donor cell transfers a DNA segment into a recipient cell through a specialized pilus

9

a. inverted repeat: a DNA sequence that is the reverse of another downstream sequence; defines the boundary of a transposon
b. DNA microarray: a DNA chip used to analyze the expression of thousands of genes simultaneously
c. chromosomal jumping: a technique used to isolate clones that contain discontinuous sequences for the same chromosome
d. genome project: the process of determining the entire set of DNA base sequences of a particular organism
e. bioinformatics: the computer-based field that facilitates the analysis of biological sequence data

12

primosome

15

base excision repair

18

transposons

21

Briefly, prokaryotic DNA replication consists of DNA unwinding, RNA primer formation, DNA synthesis catalyzed by DNA polymerase and the joining of Okazaki fragments by DNA ligase. Prokaryotic DNA replication differs from the eukaryotic process in that prokaryotic replication is faster, and in prokaryotes the Okazaki fragments are longer.

24

“Jumping genes” is the popular name for transposons. First discovered by Barbara McClintock, transposons (transposable elements) are DNA sequences that can move around the genome.

27

In bacteria general recombination is involved in transformation (DNA from one cell enters another cell and is subsequently integrated into the recipient’s genome), transduction (a bacteriophage synthesized in a bacterial cell inadvertently carries a bacterial DNA fragment to a bacterial cell that the virus infects), and conjugation (unconventional sexual mating in which DNA from one cell enters a second cell via a sex pilus).

30

Because DNA is constantly exposed to disruptive processes, its structural integrity is highly dependent on efficient repair mechanisms. The life span of an organism is dependent on the health of its constituent cells, which is in turn dependent on the timely and accurate expression of genetic information. Consequently, the capacity of the organisms in a species to maintain the integrity of DNA molecules is an important factor in determining life span.

33

The processing steps that prepare a typical eukaryotic mRNA for its functional role include capping (the linkage of a 7-methylguanosine to the 5′-end), addition of a poly (A) tail to the 3′-end and splicing (the removal of introns).

36

In species that possess DNA photolyase, light energy captured by this enzyme’s flavin and pterin chromophores is used to break the cyclobutene ring in a thymine dimer, thus converting the dimer back to two thymine nucleotide residues. The phosophodiester bond is not affected. (Human do not possess this enzyme.)

39

The function of telomere repeat-binding factors is to bind to and secure the 3′- overhang sequence of a telomere into a knot-like T-loop.

42

Telomerase is a ribonucleoprotein that is responsible for the synthesis of short DNA segments at the ends of chromosomes after DNA replication. The added segment replaces approximately 100 to 200 nucleotides that are lost from the telomere regions during the replication process. Normal somatic cells (body cells excluding germ cells) do not usually use telomerase and, as a result, their chromosomes become gradually shorted until a critical length is reached that prevents further cell divisions. Cancer cells, in contrast, use telomerase to prevent the normal shortening process, thereby giving them a growth advantage over normal cells.

45

Antibiotic resistance arises because the overuse of antibiotics acts as a selection pressure, i.e., they provide a growth advantage for disease-causing organisms that possess resistant genes. So-called superbugs are organisms that are resistant to several types of antibiotics because they possess plasmids containing several resistance genes. If the circumstances that cause antibiotic resistance continue, antibiotics may eventually become ineffective against most infectious diseases.

48

a. promoter: the sequence of nucleotides immediately before a gene that is recognized by RNA polymerase and signals the start point and direction of transcription
b. consensus sequence: the average of several similar DNA sequences: for example, the consensus sequence of the –10 box of E. coli promoter is TATAAT
c. operon: a set of linked genes that are regulated as a unit
d. chromatin-remodeling complex: a multisubunit complex that facilitates the release of the histones from nucleosomal DNA before transcription
e. general transcription factors: the minimum number of additional proteins that are necessary for accurate transcription

51

transcription

54

poly(A) tail

57

Because RBP2 is a demethylase, if pRB cannot bind and suppress its activity, this enzyme will demethylate histone lysines, thereby promoting transcription and possibly facilitating deregulated cell division. This is one way in which a defective RB1 contributes to tumor formation.

60

DNA replication time is calculated as follows:

$$\frac{150{,}000{,}000\ \textit{base pairs}}{50\ \textit{bases/s}} = 3 \times 10^6\ \text{s} = 34.5\ \text{days}$$

Consequently, approximately one month is required for this DNA replication. Eukaryotic DNA synthesis is significantly faster than expected because each chromosome contains multiple replication units (replicons).

63

a. gene silencing: a form of posttranslational gene regulation in higher eukaryotes; involves short-22 nt RNAs called microRNAs
b. RNA interference: a cellular mechanism in which RNA molecules are degraded; functions in gene expression regulation and in defense against viral RNA genomes
c. tumor suppressor gene: one of set of genes that code for proteins that actively protect cells from progressing toward cancer
d. protooncogene: a normal gene that codes for a protein involved in cell cycle regulation; promotes carcinogenesis if mutated
e. GEF: guanine nucleotide exchange factor; a protein that causes GTPases to release GDP and then bind GTP

66

structural genes

69

In relatively simple genomes, such as those in bacteria, operons provide a convenient mechanism for regulating genes. Proteins required in the same metabolic pathway or functional process are synthesized together because their genes are controlled by the same promoter. The possibility of disabling mutations in the promoter sequences which results in the transcription of none of the operon genes, is a disadvantage.

72

The purpose of gene amplification is to rapidly produce multiple copies of specific gene products that are required in greater quantities during certain stages in a cell's development.

75

Riboswitches are regulatory devices usually located in the 5′-UTRs of bacterial mRNAs. They are composed of two structural elements, an aptamer that binds a specific metabolite and a gene expression regulator called the expression platform. Several genes involved in lysine synthesis would be expected to have a riboswitch. As each of these genes is transcribed, the aptamer sequence begins to emerge from the transcription complex and into the cytoplasm, where it may or may not encounter and bind a lysine molecule. If a molecule of lysine does bind to the aptamer (indicating that the cell's lysine level is sufficient), an allosteric rearrangement of the riboswitch's gene expression regulator prevents the translation the mRNA.

78

Because the Rb gene codes for a tumor suppressor, retinoblastoma occurs only when both copies have been damaged or deleted. Usually a long period of time is required for random mutations to cause this event. In hereditary retinoblastoma, in which an affected individual possesses only one functional Rb gene, the time necessary for a random mutation to inactivate the second Rb gene is significantly less than the required for the inactivation of both genes that cause the nonhereditary version of the disease.

81

The riboswitch must have a binding site for the metabolite. This binding usually triggers a conformational change that blocks translation. So the riboswitch must also have a sequence that changes shape.

84

b. DNA topoisomerse

Chapter 19: In-Chapter Questions

19.1

The amino acid sequence of the beginning of the polypeptide is Met–Ser–Pro–Thr–Ala–Asp–Glu–Gly–Arg–Arg–Trp–Leu–Ile–Met–Phe. The mutation types in the altered mRNA sequences are (a) insertion of one base, (b) deletion of one base, (c) insertion of two bases, (d) deletion of three bases. The consequences of these mutations are altered amino acid sequences of the polypeptides produced from mRNA. In (a), (b), and (c) a frame shift occurs. In (a) the frame shift mutation results in the amino acid sequence change to MetSerProThrCys-STOP, which truncates translation. In (b) a base deletion results in the amino acid sequence change to MetSerProThrAlaAspGluGlyGlyGlyGlyGlyLeuSerCys. Not only has the amino acid sequence changed past the deletion, a five glycine residue sequence has been generated that in all probability will cause instability in an already compromised polypeptide. In (c) the insertion of two bases results in a change in the amino acid sequence beginning at residue 12: MetSerProThrAlaAspGluGlyArgArgTrpProLeuSerCys.

In addition to the sequence change, which might compromise function by itself, the introduction of a proline residue may cause a major structural change since proline causes abrupt turns in a polypeptide that can disrupt a-helices and b-strand structures. In (d) no frame shift occurs because three bases are debated. In this case, the only difference between the normal polypeptide and the mutated version is the deletion of an aspartate residue. The importance of the deletion of this residue depends on its role in the polypeptide.

19.2

Assuming that the DNA sequence given is the coding strand, the mRNA sequence is 5′-GGUUUA-3′ and the anticodons are 5′-UAA-3′. If the DNA sequence is the template strand, the mRNA sequence is 5′-UAAACC-3′ and the anticodons are 5′-GGU-3′ and 5′-UUA-3′ and 5′-ACC-3′.

19.3

The possible choices for mRNA codon base sequences for the peptide are:

Tyr—	Leu—	Thr—	Ala—
5′-UAU-3′	CUU	ACU	GCU
UAC	CUC	ACC	GCC
	CUA	ACA	GCA
	CUG	ACG	GCG
	UUA		
	UUG		

The possible choices for the DNA sequences that code for the peptide are:

Tyr—	Leu—	Thr—	Ala—
3′-ATA-5′	GAA	TGA	CGA
ATG	GAG	TGG	CGG
	GAT	TGT	CGT
	GAC	TGC	CGC
	AAT		
	AAC		

The possible choices for the tRNA anticodons that code for the peptide are:

Tyr—	Leu—	Thr—	Ala—
3′-AUA-5′	GAA	UGA	CGA
AUG	GAG	UGG	CGG
	GAU	UGU	CGU
	GAC	UGC	CGC
	AAU		
	AAG		

19.4

The formation of an ADP-ribosylated derivative of eEF-2 affects the three-dimensional structure of this protein factor. Presumably protein synthesis is arrested because the ability of eEF-2 to interact with or bind to one or more ribosomal components is altered.

19.5

After the synthesis of the plastocyanin precursor in cytoplasm, the first import signal mediates the transport of the protein into the chloroplast stroma. After this signal has been removed by a protease, a second import signal mediates the transfer of the protein into the thylakoid lumen. Plastocyanin then binds a copper atom, folds into its final three-dimensional structure, and associates with the thylakoid membrane.

Chapter 19: End-of-Chapter Questions

3

a. wobble hypothesis: the explanation of the observation that cells often have fewer tRNAs than expected; freedom in the pairing of the third base of the codon to the first base of the anticodon allows some tRNAs to pair with several codons
b. aminoacyl-tRNA synthetase: amino acyl-tRNA synthetase, an enzyme that catalyzes the attachment of an amino acid to its cognate tRNA
c. tRNA: transfer RNA; a small RNA molecule that binds to a specific amino acid and then delivers it to a ribosome for incorporation into a polypeptide chain during translation
d. AUG sequence: (the initiation codon) that binds to a complementary sequence on the 30S ribosome subunits thereby promoting the formation of the correct preinitiation complex
e. synonymous codon: one of several codons that specify the same amino acid during the translation of an mRNA

6

amide bond

9

The sequential reactions that occur within the active site of aminoacyl-tRNA synthesis are (1) the formation of aminoacyl-AMP, which contains a high-energy mixed anhydride bond, and (2) linkage of the aminoacyl group to its specific tRNA.

12

One example of a DNA deletion mutation is the deletion of the first adenine in the DNA sequence from Question 10. As a result the deletion of the first A in the sequence 5′-AGA-3′ that codes for the second Ser in the peptide results in the following DNA sequence: 5′-AAT AAC ATC AGC TTG TTT TTT GAA TAA AAA GAA GC-3′. The deletion mutation causes a change in the codons past the deletion. The mRNA codons for this sequence become:

5′-GCU UCU UUU UAU UCA AAA
AAC UUG CUG AUG UUU AA-3′

This mRNA sequence causes the synthesis of the following peptide:

Ala-Ser-Phe-Tyr-Ser-Lys-Asn-Leu-Leu-Met -Phe-

where the next amino acid residue could be Asn or Lys depending on whether the next base is U or C or A or G, respectively. The deletion of the adenine in the DNA sequence has resulted in a peptide with an altered sequence. When such deletions occur in an organism, the molecule's function can range from no perceivable change to completely nonfunctional depending on the impact of the altered sequence on the peptide's structure. Note that the deletion mutation does not cause a change in the coding for the second serine, because the old codon (UCU) and the new one (UCA) code for the same amino acid.

15

A mixed anhydride such as the linkage between the amino acyl group and AMP is a bond formed from two different acids, in this case a carboxylic acid and a phosphate. An anhydride bond, in contrast, is formed from two carboxyl groups or two phosphate groups. The reaction between an amino acid and ATP to yield aminoacyl-tRNA is irreversible because pyrophosphate, the other product of the reaction, is immediately hydrolyzed (ΔG = -33.5 kJ/mol).

18

When errors in amino acid-tRNA bond formation do occur, they are usually the result of similarities in amino acid structure (e.g., Leu vs. Val). Aminoacyl-tRNA synthetases that catalyze the linkage of tRNAs of amino acids with similar structures possess a separate proofreading site that binds the incorrect aminoacyl-tRNA products and hydrolyzes them.

21

The coding reassignment mechanism for selenocysteine is different from that of standard amino acids. This mechanism involves a unique tRNA ($tRNA^{[ser]sec}$), a seryl-tRNA synthetase with sec-generating capacity, an mRNA binding protein (SPB2), a unique elongation factor (EFsec), and a selenocystine insertion sequence in the 3′-UTR of mRNAs for selenoproteins. Sec-$tRNA^{sec}$ is formed by the linkage of serine to the tRNA's acceptor arm, after which the seryl group is

converted to a selenocysteinyl group by selenocysteine synthase.

24

a. posttranslational modification: one of a set of reactions that alter the structure of newly synthesized polypeptides
b. transpeptidation: a reaction in which an amino acid is covalently linked to another amino acid; peptide bond formation in protein synthesis
c. translocation: the movement of the mRNA through the ribosome by one triplet length after peptide bond formation; the peptidyl chain linked to the A site tRNA shifts from the A site to the P site
d. termination: the last phase in protein synthesis; a stop codon enters the A site, resulting in the binding of a protein-releasing factor
e. targeting: the process that directs newly synthesized proteins to their correct intracellular destination

27

a. ribosome recycling factor: in bacteria, a tRNA-shaped protein that binds within the A site that facilitates the release of the mRNA the remaining tRNA and the separation of the ribosomal subunits
b. trigger factor: a protein transiently linked to the bacterial ribosome and positioned at the exit site; it guides the early steps of protein folding
c. signal peptide: short sequences in a nascent polypeptide that determines its destination
d. stop transfer signal: a peptide segment that prevents further transfer of a polypeptide across a membrane
e. ternary complex: the EF-Tu-GTP-aminoacyl-tRNA complex; positions each aminoacyl-tRNA within the A site of bacterial ribosomes

30

a. proteolytic cleavage: a form of processing of newly synthesized polypeptides in which specific peptide bonds are hydrolyzed; examples include removal of the N-terminal and signal peptides and the conversion of inactive precursor to their active forms
b. proproteins: inactive precursor proteins
c. preproproteins: inactive precursor proteins with removal signal peptides
d. hydroxylation: the introduction of a hydroxyl group; the most common hydroxylated residues are proline and lysine
e. phosphorylation: the chemical addition of a phosphate group; the most commonly phosphorylated residues are serine and threonine

33

a. cotranslational transfer: the simultaneous synthesis and transmembrane transfer of a polypeptide
b. SRP receptor protein: a GTPase heterodimer that mediates the binding of the ribosome to the ER; docking protein
c. posttranslational translocation: previously synthesized polypeptides are transported across the RER membrane
d. SECIS: selenocysteine insertion sequence; a sequence element required to code for selenocysteine for a selenocysteine-containing polypeptide; usually located in the 3′-UTR of the mRNA

e. dolichol: a polyterpene component of ER membrane; plays a role in N-glycosylation of nascent polypeptides

36

folding into its biologically active conformation

39

posttranslational modifications

42

The major differences between prokaryotic and eukaryotic translation are speed (the prokaryotic process is significantly faster), location (the eukaryotic process is physically separated from transcription, which takes place in the nucleus, while the prokaryotic process begins before mRNA synthesis is complete), complexity (eukaryotes possess more complex regulatory mechanisms than prokaryotes), and posttranslational modifications (eukaryotic reactions are far more numerous than those observed in prokaryotes).

45

Preproproteins are inactive proteins with a removable signal peptide segment. Proproteins are inactive precursor proteins, which are converted to active proteins by proteolytic cleavage reactions.

48

Eukaryotic mRNA structure affects translational control as follows: The presence of a 5′-cap and a 3′-poly(A) tail, and their associated proteins allow an mRNAs to be exported to cytoplasm through nuclear pores. Several features of mRNA structure are known to affect stability. Certain sequences confer resistance to nucleases, whereas other sequences increase the likelihood of nuclease-catalyzes hydrolysis. AU rich elements found in the 3′- UTR of some mRNAs allow a rapid response to changing environmental conditions, i.e., depending on circumstances certain mRNAs may be stabilized or destabilized.

51

The proteins involved in the initiation of prokaryotic protein synthesis are: IF-1 (binds to the A site of the 30S subunit, blocking it during initiation), IF-2 (binds to the 30S subunit and promotes the binding of the initiating tRNA to the initiation codon of mRNA), and IF-3 (prevents the 30S subunit from binding prematurely to the 5OS subunit).

54

Answer d (RNA → proteins) is the process of translation.

57

The aminoacyl-tRNA synthetases correctly attach each amino acid to its cognate tRNA and proofread the product. This process increases the accuracy of protein synthesis.

60

In eukaryotes polypeptides are targeted to their final destination via transcript localization (transport of mRNAs and their subsequent binding specific cellular receptors) and signal peptides, polypeptide segments that allow binding to specific membrane-translocating complexes.

63

The phases of protein synthesis during which each process occurs are (a) initiation, (b) elongation, (c) elongation, and (d) termination.

66

The translational efficiency of eukaryotes, as measured by the number of polypeptides that can be synthesized per unit time, is largely made possible by the circular conformation of eukaryotic polysomes. After a polypeptide's synthesis is completed the released ribosomal subunits are in close proximity to the 5′-end of the mRNA, which facilitates their immediate recruitment for another round of protein synthesis.

69

Mitochondrial proteins, synthesized as preproteins, bind to a multichaperone complex composed of the ATPases hsp70 and hsp90. The preprotein-chaperone complex then docks in an ATP-dependent process with receptors of the TOM complex, which is embedded in the outer mitochondrial membrane. As soon as the N-terminal signal sequence emerges from the TOM complex entering the intermembrane space, it binds to a receptor of the TIM complex. Transport of the preprotein through the TIM complex in the inner membrane is driven by the proton gradient. After the signal sequence is removed, the newly arrived matrix protein folds into its native conformation in a process assisted by ATP-dependent chaperones.

72

Posttranslational modification reactions prepare polypeptides to serve their specific functions or direct them to specific cellular or extracellular locations. Examples of these modifications include proteolytic processing (e.g., removal of signal peptides and the conversion of a zymogen to an active enzyme), glycosylation (e.g., targeting proteins to intracellular destinations such as lysosomes and involvement in cell-cell interactions), methylation (e.g., histone methylation in epigenetic gene expression regulation), phosphorylation (a major feature of cell transduction mechanisms), hydroxylation (e.g., hydroxylation of prolyl residues in collagen), lipophilic modifications (e.g., N-myristoylation and prenylation of peripheral proteins), and disulfide bond formation (e.g., antibodies and insulin).

75

A two subunit ribosome is essential to ensure that all of the required elements of the translation process are in place before the process begins. This is a physical ordering process much like an assembly line; the parts must be in their correct locations before the various enzymatic processes are set in motion.

78

heat shock response

81

The heat shock response (HSR) is a cellular response to heat and other stresses in which there are rapid and global changes in the expression of heat shock proteins, molecules that repair misfolded proteins and protect against subsequent stress. In prokaryotes such as *E.coli*, HSR is mediated by σ^{32}, a transcription initiation factor that directs RNA polymerase to the promoters of HSR target genes. In eukaryotes, heat shock factor 1 (HSF1), a cytoplasmic protein, is in monomer form

bound to the molecular chaperone hsp90. When heat stress occurs HSF1 releases hsp90 and forms a trimer, which then relocates to the nucleus. Once HSF1 enters the nucleus, it stimulates the activity of chromatin remodeling proteins and the subsequent transcription of HSR genes.

84

c. mTORC1

87

a. CBC20/80

Glossary

43S preinitiation complex The eukaryotic multisubunit complex composed of the 40S subunits eIF-A, eIF-2-GTP, eIF-3, and methionyl-$tRNA^{met}$ that binds to mRNA.

48S initiation complex The mature eukaryotic initiation complex that scans the mRNA in search of the start codon.

A-DNA A short, compact DNA structure in which the base pairs are not at right angles to the helical axis; occurs when DNA becomes partially dehydrated.

acetal The family of organic compounds with the general formula $R'CH(OR'')_2$; formed from the reaction of a hemiacetal with an alcohol.

acid A molecule that can donate hydrogen ions.

acidosis A condition in which the pH of the blood is below 7.35 for a prolonged time.

activation energy The threshold energy required to produce a chemical reaction.

active site The cleft in the surface of an enzyme where a substrate binds.

acyl carrier protein A component of fatty acid synthase. Intermediates of fatty acid synthesis are linked to this molecule through a thioester linkage.

acyl group Any molecular group derived from a carboxylic acid by the removal of a hydroxyl group.

addition reaction A chemical reaction in which two molecules react to form a third and there are more groups attached to carbon atoms in the product.

adduct The product of an addition reaction.

aerobic metabolism The mechanism by which the chemical bond energy of food molecules is captured and used to drive the oxygen-dependent synthesis of adenosine triphosphate (ATP).

aerobic respiration The metabolic process in which oxygen is used to generate energy from food molecules.

aerotolerant anaerobe An organism that depends on fermentation for its energy needs and possesses protection from toxic oxygen metabolites in the form of enzymes and antioxidant molecules.

affinity chromatography A technique in which proteins are isolated based on their capacity to bind to a specific ligand.

aldaric acid The product formed when the aldehyde and the terminal CH_2OH groups of a monosaccharide are oxidized to carboxylic acids.

aldimine An imine product of a reaction of a primary amine group with a carbonyl group; also referred to as a Schiff base.

aldol cleavage A reverse of the aldol condensation reaction.

aldol condensation An aldol addition reaction; the nucleophilic addition of a ketone enolate ion to an aldehyde to form a β-hydroxyketone, followed by the elimination of a water molecule.

aldonic acid The product of the oxidation of the aldehyde group of a monosaccharide.

aldose A monosaccharide with an aldehyde functional group.

aliphatic hydrocarbon A nonaromatic hydrocarbon such as methane or cyclohexane.

alkalosis A condition in which the blood pH is above 7.45 for a prolonged period of time.

alkylating agent An electrophile that reacts with a molecule that possesses an unshared pair of electrons, adding an alkyl group.

allosteric enzyme An enzyme whose activity is affected by the binding of effector molecules.

allosteric transition The ligand-induced conformational change in a protein.

allostery The control of protein function through ligand-binding events.

amino acid An organic molecule that contains an amino group and a carboxyl group.

amino acid pool The amino acid molecules that are immediately available in an organism for use in metabolic processes.

amino acid residue An amino acid that has been incorporated into a peptide molecule.

amphibolic pathway A metabolic pathway that functions in both anabolism and catabolism.

amphipathic molecule A molecule containing both polar and nonpolar domains.

amphoteric molecule A molecule that can act as both an acid and a base.

AMPK AMP-activated protein kinase; an important regulatory enzyme in energy metabolism.

amylopectin A type of plant starch; a branched polymer containing α(1,4)- and α(1,6)-glycosidic linkages.

amylose A type of plant starch; an unbranched polymer of D-glucose residues linked with α(1,4)-glycosidic linkages.

anabolic pathways A series of biochemical reactions in which larger molecules are synthesized from smaller precursor molecules.

anaerobic organisms Organisms that do not use oxygen to generate energy.

anaplerotic reaction A reaction that replenishes a substrate needed for a biochemical pathway.

anhydride The product of the condensation reaction between two carboxyl groups or two phosphate groups in which a molecule of water is eliminated.

annotation The functional identification of the genes of a genome.

anomer An isomer of a cyclic sugar that differs from another in its configuration about the hemiacetal or acetal carbon.

antenna pigment A molecule that absorbs light energy and transfers it to a reaction center during photosynthesis.

anticodon A sequence of three ribonucleotides on a tRNA molecule that is complementary to a codon on the mRNA molecule; codon-anticodon binding results in the delivery of the correct amino acid to the site of protein synthesis.

antioxidant A substance that prevents the oxidation of other molecules.

antisense strand A noncoding DNA strand that is complementary to the base sequence of an mRNA molecule transcribed from the coding DNA strand.

apoenzyme The protein portion of an enzyme that requires a cofactor to function in catalysis.

apoprotein A holoprotein without its prosthetic group.

apoptosis The genetically programmed series of events that lead to cell death.

apurinic site A nucleotide residue in a DNA molecule from which a purine base has been lost or removed.

apyrimidinic site A nucleotide residue in a DNA molecule from which a pyrimidine base has been lost or removed.

aromatic hydrocarbon A molecule that contains a benzene ring or has properties similar to those exhibited by benzene.

asymmetric (chiral) carbon A carbon bound to four different groups.

atherosclerosis A cardiovascular disease in which soft masses containing fatty material and cellular debris are formed in the lining of blood vessels.

attachment (att) site A short DNA sequence that facilitates site-specific recombination; also refers to an IS element.

autocrine A hormone-like molecule that is active within the cell in which it is produced.

autoimmune disease A condition in which an immune response is directed against an individual patient's own tissues.

autophagy A cellular degradation pathway in which cell components are degraded by enzymes in lysosomes.

autopoiesis A system that is autonomous, self-organizing, and self-maintaining.

autotroph An organism that transforms light energy or the energy of various chemicals into the chemical bond energy of biomolecules.

B cell A B lymphocyte: a white blood cell that produces and secretes antibodies, the proteins that bind to foreign substances, thereby initiating their destruction in the humoral immune response.

B-DNA The commonly found form of DNA, as the sodium salt under highly humid conditions.

base A molecule that can accept hydrogen ions.

base analogue A molecule that resembles a normal DNA nucleotide base and can substitute for it during DNA replication, leading to mutation.

base excision repair A mechanism that removes and then replaces individual nucleotides in DNA whose bases have undergone various types of damage (e.g. alkylation, deamination, or oxidation).

bile salts Amphipathic molecules with detergent properties that are important components of bile, a yellowish green liquid that aids in the digestion of fat; a conjugated derivative of the bile acids cholic acid and deoxycholic acid.

bioenergetics The study of energy transformations in living organisms.

biogenic amine An amino acid derivative that acts as a neurotransmitter (e.g., GABA and the catecholamines).

biogeochemical cycle A pathway driven by solar and geothermal energy in which a chemical element moves throughout Earth's biotic and abiotic compartments.

bioinformatics The computer-based field that facilitates the analysis of biological sequence data.

biomolecules Molecules that make up a living organism.

biotransformation A series of enzyme catalyzed processes in which toxic and/or hydrophobic molecules are converted into (usually) less toxic and more soluble metabolites.

branched-chain amino acid One of a group of essential amino acids (leucine, isoleucine, and valine) with hydrophobic side chains.

buffer A substance that resists large pH changes when small amounts of acids or bases are added; usually a solution that contains a weak acid and its conjugate base.

C3 plants Plants that produce glycerate-3-phosphate, a three-carbon molecule, as the first stable product of photosynthesis.

C4 metabolism A photosynthetic pathway in plants such as corn and sugarcane that produces a four-carbon molecule and avoids photorespiration.

C4 plants Plants that possess mechanisms that suppress photorespiration by separating rubisco (ribulose-1,5-bisphosphate carboxylase) from atmospheric O_2.

Calvin cycle The major metabolic pathway by which carbon dioxide is incorporated into organic molecules

cap-binding complex (CBC) A protein complex that binds to capped mRNA molecules and facilitates their translation; consists of eIF-4A (a helicase), eIF-4E (a translation initiation factor), and eIF-G (a scaffold protein); also referred to as eIF-4F.

carbanion A carbon with a negative charge.

carbocation A carbon with a positive charge.

carotenoid An isoprenoid molecule that functions as a light-harvesting pigment and/or protects against reactive oxygen species (ROS).

carrier protein A membrane transport protein.

catabolic pathway A series of biochemical reactions in which a larger molecule is degraded into smaller, simpler products; in some catabolic pathways, energy is captured.

catalyst A substance that enhances the rate of a chemical reaction but is not permanently altered by the reaction.

cDNA library A clone library of cDNA (complementary DNA) molecules produced from mRNA or miRNA molecules by reverse transcription.

cell division cycle protein One of a family of phosphatases that activate Cdks by removing inhibitory phosphate groups.

cell fractionation A technique involving homogenization and centrifugation that allows the study of cell organelles.

cellobiose A degradation product of cellulose; a disaccharide that contains two molecules of glucose linked by a β(1,4)-glycosidic bond.

cellulose A polymer produced by plants that is composed of D-glucopyranose residues linked by β(1,4)-glycosidic bonds.

centromere A special region of repetitive DNA that plays a critical role in cell division; it holds the two sister chromatids together during prophase and metaphase of cell division.

chain-terminating method A technique for determining DNA base sequences that uses 2′-3′-dideoxy base analogues as chain-terminating inhibitors of DNA polymerase; also referred to as the Sanger method.

channel protein A membrane protein that contains a pore through which ions are transported.

chaperone-mediated autophagy A receptor-mediated process in which specific proteins that are bound to a chaperone complex are unfolded and then translocated into a lysosome, where they are then degraded.

Chargaff's rules A set of rules describing the base composition of DNA and positing the equality of the concentration of adenine and thymine and of cytosine and guanine.

chemiosmotic coupling theory ATP synthesis is coupled to electron transport by an electrochemical proton gradient across a membrane.

chemoautotroph An organism that transforms the energy of various chemicals into chemical bond energy.

chemoheterotroph An organism that uses preformed organic food molecules as its sole source of energy.

chemosynthesis The biochemical mechanism whereby chemical energy is extracted from certain minerals.

chitin The principal structural component of the exoskeletons of arthropods and the cell walls of many fungi; a homoglycan composed of *N*-acetylglucosamine residues.

chlorophyll A magnesium-containing green pigment molecule found in plants and photosynthetic bacteria that resembles heme; absorbs light energy in photosynthesis.

chromatin remodeling complex A multisubunit complex that facilitates the release of the histones from nucleosomal DNA during transcription.

chromatin The complex of DNA and histones found in the nucleus of eukaryotic cells.

chromatin The complex of DNA, histones and other chromosomal proteins found in the nucleus of eukaryotic cells.

chromophore A molecular component that absorbs light of a specific frequency.

chromosomal jumping A cloning technique used in the physical mapping of genomes. It allows the bypassing of regions, such as repetitive sequences, that are difficult to clone.

chromosomal walking A technique in which overlapping DNA clones are used to determine the sequence of an unknown chromosomal region.

chromosome A very long DNA molecule associated with proteins that contains the genes of an organism.

chromosome territory A region of the nucleus preferentially occupied by a specific chromosome.

chylomicron A large lipoprotein of extremely low density: transports dietary triacylglycerols and cholesteryl esters from the intestine to muscle and adipose tissue.

chylomicron remnants Chylomicrons after about 90% of the triacylglycerols have been removed by lipoprotein lipase.

cistron A DNA sequence that contains the coding information for a polypeptide and the signals required for ribosome function.

clamp loader The γ complex that recognizes single DNA strands with primer and transfers β_2-clamp dimer to the core polymerase.

clathrin A protein that plays a major role in the formation and functional properties of coated vesicles.

clathrin-dependent endocytosis A receptor-mediated process whereby cells internalize molecules by the inward budding of plasma membrane vesicles.

coding strand The DNA strand that has the same base sequence as the RNA transcript (with thymine instead of uracil).

codon A sequence of three nucleotides in mRNA that directs the incorporation of an amino acid during protein synthesis or acts as a start or stop signal.

coenzyme A A carrier of acetyl and acyl groups that is composed of a 3′-phosphate derivative of ADP linked to pantothenic acid (via a phosphate ester bond), which in turn is linked to β-mercaptoethylamine via an amide bond.

coenzyme A small organic molecule required in the catalytic mechanisms of certain enzymes.

cofactor The nonprotein component of an enzyme (either an inorganic ion or a coenzyme) required for catalysis.

colony hybridization technique A method used to identify bacterial colonies that possess a specific recombinant DNA sequence.

competitive inhibition A reversible type of enzyme inhibition in which the inhibitor molecule competes with the substrate for occupation of the active site.

composite transposon A bacterial transposon composed of a gene and flanking IS elements.

conjugate base The anion (or molecule) that results when a weak acid loses a proton.

conjugate redox pair An electron donor and its electron acceptor form: for example, NADH and NAD^+.

conjugated protein A protein that functions only when it carries other chemical groups attached by covalent linkages or by weak interactions.

conjugation reaction A biochemical reaction that may improve the water solubility of a molecule by converting it to a derivative that contains a water-soluble group.

conjugation Unconventional sexual mating between bacterial cells; a donor cell transfers a DNA segment into a recipient cell through a specialized pilus.

consensus sequence The average of several similar DNA sequences: for example, the consensus sequence of the –10 box of *E. coli* promoter is TATAAT.

constitutive heterochromatin Highly condensed and permanently transcriptionally silent segments of DNA in eukaryotes, most notably repetitive sequences, telomeres, and centromeres.

contig One of a set of overlapping DNA sequences used to identify the base sequence of a region of DNA.

cooperative binding A mechanism in which binding of one ligand to a target molecule promotes the binding of other ligands.

Cori cycle A metabolic process in which lactate, produced in tissues such as muscle, is transferred to liver, where it becomes a substrate in gluconeogenesis.

cosmid Cloning vehicles that contain the γ bacteriophage cos sites incorporated into plasmid DNA sequences with one or more selectable markers.

cotranslational transfer The insertion of a polypeptide across a membrane during ongoing protein synthesis.

covalent bond The sharing of electrons between atoms.

CpG CpG dinucleotides; methylated cytosines occur predominantly in 5′-CG-3′ sequences.

CpG islands Regions of the genome where CpGs constitute more than 50% of the bases.

Crabtree effect The physiological capacity of *S. cerevisiae* cells to ferment sugar to produce ethanol, which kills their competitors, and then use the ethanol as an energy source.

CRISPR A genome editing technology adapted from a prokaryotic defense mechanism that is used to remove or add genes.

cyclin One of a family of proteins that control the progression of cells through the cell cycle.

cyclin-dependent kinase One of a family of enzymes, activated by cyclin proteins, that are involved in cell cycle regulation.

cytochrome P_{450} system An electron transport system that consists of two enzymes (NADPH-cytochrome P_{450} reductase and cytochrome P_{450}); involved in the oxidative metabolism of many endogenous and exogenous substances.

cytokine One of a group of hormone-like polypeptides and proteins secreted by certain immune system cells.

cytoskeleton A set of protein filaments (microtubules, microfilaments, and intermediate fibers) that maintains the cell's internal structure and allows organelles to move.

DAG Diacylglycerol; a second messenger molecule in the phosphatidylinositol pathway.

decarboxylation The removal of a carboxylic group from a carboxylic acid as carbon dioxide.

decoding center The position located within the bacterial 30S ribosomal subunit where codon-anticodon base pairs form.

degeneracy The capacity of structurally different system parts to perform the same or similar functions.

denaturation A disruption of protein or nucleic acid structure caused by exposure to heat or chemicals leading to loss of biological function.

denaturation In nucleic acid research a disruption of double-stranded structure caused by exposure to heat.

density-gradient centrifugation A technique in which cell fractions are further purified by centrifugation in a density gradient.

desensitization A process in which target cells adjust to changes in stimulation by decreasing the number of cell surface receptors or by inactivating those receptors.

detoxication The process by which a toxic molecule is converted to a more soluble (and usually less toxic) product.

detoxification Correction of a state of toxicity; the chemical reactions that produce sobriety in an inebriated person.

diastereomer A stereoisomer that is not an enantiomer (mirror-image isomer).

differential centrifugation A cell fractionation technique in which organelles in homogenized cells are separated by centrifugal forces.

dipole A difference in charge between atoms in a molecule resulting from the unsymmetrical orientation of polar bonds.

disaccharide A glycoside composed of two monosaccharide residues.

dissipative system A system that facilitates the reduction of an energy gradient.

disulfide bridge A covalent bond formed between the sulfhydryl groups of two cysteine residues.

disulfide exchange An enzyme-catalyzed posttranslational process in which there is an interchange of disulfide bonds in a protein until the correct biologically relevant disulfide bonds are formed.

DNA fingerprinting A laboratory technique used to compare DNA banding patterns from different individuals.

DNA glycosylase A DNA repair enzyme that cleaves the N-glycosidic linkage between the damaged base and the deoxyribose component of the nucleotide.

DNA ligase An enzyme that catalyzes the formation of a covalent phosphodiester bond between the 3′-OH end of one segment and the 5′-phosphate end of another segment during DNA replication.

DNA profile A unique DNA pattern of repeats of target sequences that is separated in an electrophoresis gel; used to identify individuals.

DNA typing A DNA analysis technique used to identify individuals; involves the analysis of several highly variable sequences called markers.

double-strand break repair model A general recombination mechanism for repairing double-strand breaks in DNA utilizing homologous chromosomes; produces both crossover and noncrossover products.

downregulation The reduction in cell surface receptors in response to stimulation by specific hormone molecules.

dyslipidemia High blood levels of total cholesterol and triacylglycerol and low HDL levels; a disorder associated with metabolic syndrome.

effector A molecule whose binding to a protein alters the protein’s activity.

eicosanoid A hormone-like molecule that contains 20 carbons; most are derived from arachidonic acid; examples include prostaglandins, thromboxanes, and leukotrienes.

electron acceptor Species that accepts electrons from an electron donor during a reaction.

electron donor Species that donates electrons to an electron acceptor during a reaction.

electron transport system A series of electron carrier proteins that bind reversibly to electrons at different energy levels.

electrophile An electron-deficient species that is preferentially attracted to a region of high electron density in another species during a chemical reaction.

electrophoresis A class of techniques in which molecules are separated from each other because of differences in their

net charge.

electrostatic interaction Noncovalent attraction between oppositely charged atoms or groups.

elimination reaction A chemical reaction in which a double bond is formed when atoms in a molecule are removed.

elongation The polypeptide chain growth phase during translation of an mRNA in a ribosome.

emergent property A new property conferred by the complexity and dynamics of the system.

enantiomer A mirror-image stereoisomer.

enantiomers Stereoisomers that are mirror-images.

endergonic process A reaction that does not spontaneously go to completion; the standard free energy change is positive and the equilibrium constant is less than 1.

endocrine system A signaling system consisting of cells that secrete hormones into the bloodstream to act on distant target cells.

endocytosis A process in which a cell takes up solutes or particles by enclosing them in vesicles pinched off from its plasma membrane.

endogenous retrovirus A decayed virus within a genome: also referred to as a LTR retrotransposon.

endomembrane system An extensive set of interconnecting internal membranes that divide the cell into functional compartments.

endoplasmic reticulum (ER) A series of membranous channels and sacs that provides a compartment separate from the cytoplasm with diverse functions including numerous chemical reactions.

endothermic reaction A reaction that requires energy.

enediol The intermediate formed during the isomerization reactions of monosaccharides. It contains a double bond with a hydroxyl group on each carbon of the double bond.

energy The capacity to do work.

enhancer A short DNA sequence that promotes the transcription of one or more genes when it is bound to an activator protein.

enthalpy The heat content of a system; in a biological system it is essentially equivalent to the total energy of the system.

entropy A measure of the randomness or disorder of a system; a measure of that part of the total energy in a system that is unavailable for useful work.

enzyme A biomolecule that catalyzes a biochemical reaction.

enzyme induction A process in which a signal molecule stimulates increased synthesis of a specific enzyme.

enzyme kinetics The study of the rates of enzyme-catalyzed reactions.

epigenetics Heritable covalent modification-induced gene activations and repressions that do not change DNA base sequences.

epigenome The current epigenetic modifications within a cell.

epimer A molecule that differs from the configuration of another by one asymmetric carbon.

epimerization The reversible interconversion of epimers.

epimutation An alteration in the normal epigenetic pattern.

epoxide An ether in which the oxygen is incorporated into a three-membered ring.

ER stress Various stress conditions cause misfolded proteins to accumulate in the ER.

ER-associated protein degradation (ERAD) A process that targets misfolded proteins for destruction.

essential amino acid An amino acid that cannot be synthesized by the body and must be supplied by the diet.

essential fatty acid A fatty acid that must be supplied in the diet because it cannot be synthesized by the body; linoleic and linolenic acids in humans.

euchromatin A less condensed form of chromatin that has varying levels of transcriptional activity.

exergonic process A reaction that spontaneously goes to completion as written; the standard free energy change is negative, and the equilibrium constant is greater than 1.

exocytosis The secretion process in eukaryotic cells; involves the fusion of membrane-bound secretory granules with the plasma membrane.

exon The region in a split or interrupted gene that encodes a segment of a final mature RNA molecule.

exonuclease An enzyme that removes nucleotides from the end of the polynucleotide strand.

exothermic reaction A reaction that releases heat.

extracellular matrix (ECM) A gelatinous material, containing proteins and carbohydrates, that binds cells and tissues together.

facultative anaerobe An organism that possesses the capacity for detoxifying oxygen metabolites; energy is generated using oxygen, when available, as an electron acceptor.

facultative heterochromatin A condensed DNA sequence that can become less condensed and transcriptionally active in response to specific signaling mechanisms.

fatty acid A monocarboxylic acid usually with an even number of carbon atoms; R-COOH where R is an alkyl group.

fatty acid–binding protein An intracellular water-soluble protein whose function is to bind and transport hydrophobic fatty acids.

feedback control The control of a self-regulating system (e.g., a metabolic process or pathway) in which product influences the output of the process.

fermentation An energy-yielding process in which organic molecules serve as both donors and acceptors of electrons; the anaerobic degradation of sugars.

fibrous protein A protein composed of polypeptides arranged in long sheets or fibers.

flavin adenine dinucleotide (FAD) A tightly bound prosthetic group consisting of riboflavin and D-ribitol linked to an adenosine group through a pyrophosphate linkage that functions in the class of enzymes called flavoproteins.

flavin mononucleotide (FMN) A tightly bound prosthetic group consisting of riboflavin and D-ribitol phosphate that functions in the class of enzymes called flavoproteins.

flavin-containing monooxygenase One of a family of NADPH- and O_2-requiring enzymes that oxidize molecules (usually xenobiotics) with nitrogen-, sulfur-, or phosphorus-containing functional groups.

flavoprotein A conjugated protein in which the prosthetic group is either FMN or FAD.

fluid mosaic model The currently accepted model of cell membranes in which the membrane is a lipid bilayer with integral proteins buried in the lipid and peripheral proteins loosely attached to the membrane surface.

fluorescence A form of luminescence in which certain molecules can absorb light of one wavelength and emit light of another wavelength.

fold A core three-dimensional structure of a protein domain.

free energy The energy in a system available to do useful work.

free radical An atom or molecule that has an unpaired electron.

functional genomics The investigation of gene and protein functions and interactions.

functional genomics The investigation of gene expression patterns.

functional group A group of atoms that undergoes characteristic reactions when attached to a carbon atom in an organic molecule or biomolecule.

G protein A heterotrimeric GTP-binding protein that acts as a molecular switch when activated by a GPCR.

G protein–coupled receptor (GPCR) A cell surface receptor that transduces the binding of a hormone or other signal molecule into an intracellular response via the activation of a G protein.

gasotransmitter An endogenous gaseous molecule that acts as a signal molecule.

gel-filtration chromatography A technique used to separate molecules according to their size and shape that employs a column packed with a gelatinous polymer.

gene A DNA sequence that codes for a polypeptide, or an RNA molecule.

gene A DNA sequence that codes for a polypeptide, rRNA, tRNA or several types of noncoding RNAs.

gene duplication The creation of a duplicate gene or part of a gene; can result from unequal crossing over during meiosis or from retrotransposition.

gene expression The mechanism by which living organisms regulate the flow of genetic information; the control of when and if genes are transcribed.

gene silencing A form of posttranscriptional gene regulation that involves 22-nt miRNAs.

general recombination Recombination involving exchange of a pair of homologous DNA sequences; it can occur at any location on a chromosome.

genetic code The set of nucleotide base triplets (codons) that code for the amino acids in proteins as well as start and stop signals.

genetics The scientific investigation of inheritance.

genome The total genetic information possessed by an organism.

genomics The investigation of entire genomes; the sequencing and characterization of genomes.

ghrelin A protein that stimulates appetite; produced by the cells of the stomach and small intestine.

globular protein A protein that adopts a globular shape.

glucagon A peptide hormone released from pancreatic α-cells; among its effects are increasing the level of glucose in blood via the breakdown of liver glycogen.

glucogenic Describing amino acids that are degraded to pyruvate or a citric acid intermediate; these amino acids are used as substrates in the synthesis of glucose in gluconeogenesis.

gluconeogenesis The synthesis of glucose from noncarbohydrate molecules.

glucose–alanine cycle A method of recycling α-keto acids between muscle and liver and for transporting ammonia to the liver.

glucosuria The presence of glucose in the urine.

glycan A polymer of monosaccharides; a polysaccharide.

glycerol-3-phosphate shuttle A metabolic process that uses glycerol-3-phosphate to transfer electrons from NADH in the cytoplasm to mitochondrial FAD.

glyceroneogenesis An abbreviated version of gluconeogenesis in which glycerol- 3-phosphate is synthesized from substrates other than glucose or glycerol.

glycocalyx A layer on the external surface of many eukaryotic cells that contains substantial amounts of carbohydrate-containing molecules.

glycoconjugate A molecule that possesses covalently bound carbohydrate components (e.g., glycoproteins and glycolipids).

glycoform One of several slightly different forms of a glycan component of a glycoprotein.

glycogen A glucose storage molecule in vertebrates; a branched polymer containing α(1,4)- and α(1,6)-glycosidic linkages.

glycogenesis A biochemical pathway that adds glucose to growing glycogen polymers when blood glucose levels are high.

glycogenolysis A biochemical pathway that removes glucose molecules from glycogen polymers when blood glucose levels are low.

glycolipid A glycosphingolipid; a molecule in which a monosaccharide, disaccharide, or oligosaccharide is attached to a ceramide through an O-glycosidic linkage.

glycolysis The enzymatic pathway that converts a glucose molecule into two molecules of pyruvate: the anaerobic process generates energy in the form of two ATP molecules and two NADH molecules.

glycome The total set of sugars and glycans that a cell or organism produces.

glycoprotein A conjugated protein in which carbohydrate molecules are covalently bound.

glycosaminoglycan A long unbranched heteropolysaccharide chain composed of disaccharide repeating units.

glycoside The acetal of a sugar.

glycosidic link An acetal linkage formed between two monosaccharides.

glyoxylate cycle A modification of the citric acid cycle that occurs in plants, bacteria, and other eukaryotes: allows growth in these organisms from two-carbon substrates such as ethanol, acetate, and acetyl-CoA.

Golgi apparatus (complex) A series of curved membranous sacs involved in packaging and distributing cell products to internal and external compartments.

GPI (glycosylphosphatidylinositol) anchor A glycolipid used to link certain proteins to membrane, preferentially in lipid rafts.

grana (pl) Stacks of thylakoid membrane.

granum (*sing*) The folded portion of the thylakoid membrane.

growth factor An extracellular polypeptide that stimulates cells to grow and/or undergo cell division.

GTPase associated region (GAR) A set of overlapping binding sites on the 50S ribosomal subunit that activate specific translation factors with GTPase activity during protein synthesis.

GTPase-activating protein (GAP) A protein molecule that stimulates the GTPase activity of a GTP-binding protein. GTP hydrolysis of a GTP- Gα subunit yields a GDP-Gα subunit.

guanine nucleotide exchange factor (GEF) A protein that mediates a conformational change in the transmembrane region of a G-protein-coupled receptor and leads to GDP/ GTP exchange during G-protein activation.

helicase An ATP-requiring enzyme that catalyzes the unwinding of duplex DNA.

hemiacetal One of the family of organic molecules with the general formula R'CH(OR'')OH that is formed by the reaction of one molecule of alcohol with an aldehyde.

hemiketal One of the family of organic molecules with the general formula R'R''C(OR''')OH that is formed by the reaction of a molecule of alcohol with a ketone.

hemoprotein A conjugated protein in which heme, an iron-containing organic group, is the prosthetic group.

heterochromatin Chromatin that is so highly condensed that it is transcriptionally inactive.

heteroglycan A high-molecular-weight carbohydrate polymer that contains more than one kind of monosaccharide.

heterotroph An organism that obtains energy by degrading preformed food molecules usually obtained by consuming other organisms.

high-density lipoprotein A type of lipoprotein with a high protein content that is believed to scavenge excess cholesterol from cell membranes and transport it to the liver.

holoenzyme A complete enzyme consisting of an apoenzyme plus a cofactor.

holoprotein An apoprotein combined with its prosthetic group.

homeostasis The capacity of living organisms to regulate metabolic processes despite variability in their internal and external environments.

homoglycan High-molecular-weight carbohydrate polymers that contain only one type of monosaccharide.

homologous polypeptide Protein molecules whose amino acid sequences are similar; implies a common evolutionary origin.

hormone A molecule produced by a specific cell that influences the function of distant target cells.

hormone response element A specific DNA sequence that binds hormone-receptor complexes; the binding of a hormone-receptor complex either enhances or diminishes the transcription of a specific gene.

humoral immune response The immunity that results from the presence of antibodies in blood and tissue fluid; also referred to as an antibody-mediated immunity.

hybridization A laboratory technique in which fragments of single-stranded DNA or RNA from different sources anneal; the rate at which a double-stranded hybrid forms is a measure of the similarity of the two strands.

hydration reaction A type of addition reaction in which water is added to a carbon-carbon double bond.

hydrocarbons Compounds that contain only carbon and hydrogen.

hydrogen bond The force of attraction between a hydrogen atom and a small highly electronegative atom (e.g., O or N) on another molecule or the same molecule.

hydrolase An enzyme that catalyzes reactions in which adding water cleaves bonds.

hydrolysis A chemical reaction in which molecules are cleaved by water.

hydrophilic Describing molecules or portions thereof that dissolve easily in water; hydrophilic molecules possess positive or negative charges or contain relatively large numbers of electronegative oxygen or nitrogen atoms.

hydrophobic Describing molecules that do not dissolve in water and possess few, if any, electronegative atoms.

hyperammonemia A potentially fatal elevation of the concentration of ammonium ions in the blood.

hyperglycemia Blood glucose levels that are higher than normal.

hyperinsulinemia Higher than normal blood levels of insulin.

hyperosmolar hyperglycemic nonketosis Severe dehydration in non-insulin-dependent diabetics; caused by persistently high blood glucose levels.

hypertonic solution A concentrated solution with a high osmotic pressure.

hypochromic effect The decrease in the absorption of UV light (260 nm) that occurs when purine and pyrimidine bases in single stranded polynucleotides are incorporated into the base pairs in double stranded polynucleotides.

hypoglycemia Blood glucose levels that are lower than normal.

hypotonic solution A dilute solution with a low osmotic pressure.

indel An insertion or deletion mutation; occurs when from one to thousands of bases are inserted or deleted, respectively, from a DNA sequence.

inhibitor A molecule that reduces an enzyme's activity.

initiation The beginning phase of translation.

inner mitochondrial membrane A selectively permeable membrane in which are embedded the respiratory complexes that are responsible for ATP synthesis.

inner nuclear membrane A membrane that encloses the nucleoplasm; contains proteins with functions such as stabilizing nuclear structure and chromatin binding.

insertional element A short DNA sequence involved in site-specific recombination; also called an IS element or att site.

insulator A DNA sequence that when bound to an insulator binding protein blocks the interaction between enhancers and the promoters of neighboring genes; also prevents the spread of heterochromatin.

insulin A peptide hormone released from pancreatic β-cells; among its many effects is the promotion of glucose uptake into the cells of certain target organs (muscle and adipose tissue).

insulin resistance The insensitivity of tissues to insulin; a common cause is the downregulation of insulin receptors.

integral protein A protein that is embedded within a membrane.

intercalating agents Planar molecules that insert themselves between base pairs; this action distorts the DNA chain.

intermediate A species produced in the course of a reaction that exists for a finite period of time.

intermediate-density lipoprotein (IDL) A lipoprotein formed when a very-low-density lipoprotein shrinks in size and becomes more dense as a result of depletion of triacylglycerol, apolipoprotein, and phospholipid molecules.

intermediate filament A cytoskeletal component that provides cells with significant mechanical support; a flexible, strong, and relatively stable polymer.

interspersed genome-wide repeats Repetitive DNA sequences that are scattered around the genome.

intrinsic termination Transcription termination that involves an RNA termination sequence that contains an inverted repeat sequence; also referred to as rho-independent termination.

intrinsically disordered proteins (IDPs) Proteins that are partially or completely lacking in a stable three-dimensional stucture.

intron A noncoding intervening sequence in a split or interrupted gene; missing in the final RNA product.

inversion mutation A deleted DNA fragment is reinserted into its original position, but in the opposite direction.

inverted repeat A DNA sequence that is a reversed complement of another downstream sequence; defines the boundary of a transposon.

ion-exchange chromatography A technique that separates molecules on the basis of their charge.

ionophore A substance that transports cations across membranes.

IP_3 Inositol-1,4,5-triphosphate; the IP3 receptor is a calcium channel.

irreversible inhibition A form of enzyme inhibition in which an inhibitor molecule permanently impairs an enzyme,

usually through binding via a covalent bond.

isoelectric point (p*I*) The pH at which a protein has no net charge.

isomerase An enzyme that catalyzes the conversion of one isomer to another.

isomerization reaction A reaction that involves the intermolecular shift of atoms or groups.

isoprenoid One of a class of biomolecules that contain repeating five-carbon structural units known as isoprene units: examples include terpenes and steroids.

isothermic reaction Reactions in which heat is not exchanged with the surroundings; $\Delta H = 0$.

isotonic solution A solution with exactly the same particle concentration as that inside cells; there is no net movement of water in or out of the cells.

ketal The family of organic compounds with the general formula R'R''C(OR''')$_2$; formed from the reaction of a hemiketal with an alcohol.

ketoacidosis Acidosis caused by the excessive accumulation of ketone bodies.

ketogenesis The condition in which excess acetyl-CoA molecules are converted to acetoacetate, β-hydroxybutyrate, and acetone (referred to collectively as the ketone bodies).

ketogenic Describing amino acids degraded to form acetyl-CoA or acetoacetyl-CoA.

ketone body One of three molecules (acetone, acetoacetate, or β-hydroxybutyrate) that are produced in the liver from acetyl-CoA.

ketose A monosaccharide with a ketone functional group.

ketosis Accumulation of ketone bodies in blood and tissues.

Krebs bicycle A biochemical pathway in which the aspartate required in the urea cycle is generated from oxaloacetate, an intermediate in the citric acid cycle.

Krebs urea cycle The cyclic pathway that converts waste ammonia molecules along with carbon dioxide and aspartate into urea; named for its discoverer, Hans Krebs.

lactone A cyclic ester.

lactose A disaccharide found in milk; composed of one molecule of galactose linked in a β(1,4)-glycosidic bond to a molecule of glucose.

Le Chatelier's principle Law that states that when a system in equilibrium is disturbed, the equilibrium shifts to oppose the disturbance.

leaving group The group displaced during a nucleophilic substitution reaction.

lectin A carbohydrate-binding protein.

leptin A 16 kDa satiety-inducing protein secreted into the bloodstream primarily by adipose tissue.

leukotriene A biologically active molecule derived from arachidonic acid; its synthesis is initiated by a peroxidation reaction.

ligand A molecule that binds to a specific site on a larger molecule.

ligase An enzyme that catalyzes the joining of two molecules.

light harvesting The capture of light energy by pigment molecules; once absorbed, the energy is channeled to a photosynthetic reaction center.

light-harvesting antenna An array of protein and chlorophyll molecules within the thylakoid membrane in chloroplasts; transfers light energy to a chlorophyll a molecule within a photosystem reaction center.

light-independent reactions A photosynthetic pathway in which CO_2 is incorporated into carbohydrate that can occur in the absence of light; also referred to as the Calvin cycle.

LINE (long interspersed nuclear elements) Retrotransposons with lengths greater than 5 kb that contain a strong promoter, an integration sequence, and the coding sequences for transposition enzymes.

lipid Any of a group of biomolecules that are soluble in nonpolar solvents and insoluble in water.

lipogenesis The biosynthesis of fatty acids and triacylglycerol (body fat).

lipoic acid A biomolecule that contains a carboxylate group and two sulfhydryl groups that are easily oxidized or

reduced; functions as an acyl group carrier in pyruvate dehydrogenase complex and α-ketoglutarate dehydrogenase complex.

lipolysis The enzyme-catalyzed hydrolysis of triacylglycerol molecules.

lipoprotein A conjugated protein in which lipid molecules are the prosthetic groups; a protein-lipid complex that transports water-insoluble lipids in the blood.

London dispersion force A temporary dipole–dipole interaction.

long ncRNA (lncRNA) Regulatory ncRNAs longer than 200 nt, that are involved in diverse cell processes serving as guides, scaffolds, or decoys.

low-density lipoprotein A type of lipoprotein that contains cholesterol, triacylglycerols, and phospholipids; transports cholesterol to peripheral tissues.

lyase An enzyme that catalyzes the cleavage of C—O, C—C, or C—N bonds, thereby producing a product that contains a double bond.

lysogeny The integration of a viral genome into a host genome.

lytic cycle A viral life cycle in which a virus destroys its host cell.

macroautophagy A cellular pathway that uses lysosomes for bulk degradation of cytoplasmic components; also referred to as autophagy.

macromolecular crowding The dense packing of an enormous variety of macromolecules and other molecules within the interior of cells that has a significant effect on biochemical processes.

macromolecule A biopolymer; examples include polypeptides and DNA.

malate shuttle A metabolic process in which the electrons from NADH in the cytoplasm are transferred to mitochondrial NAD^+; oxaloacetate is transferred by reversible conversion to malate from a mitochondrion to the cytoplasm.

malate-aspartate shuttle A metabolic process in which the electrons from NADH in the cytoplasm are transferred to mitochondrial NAD^+; oxaloacetate is transferred by reversible conversion to malate from a mitochondrion to the cytoplasm.

maltose A degradation product of starch hydrolysis; a disaccharide composed of two glucose molecules linked by an α(1,4)-glycosidic bond.

marker enzyme An enzyme known to be a reliable indicator of the presence of a specific organelle.

mass spectrometry A technique in which molecules are vaporized and then bombarded by a high-energy electron beam, causing them to fragment as cations.

mass spectrometry A technique in which molecules are vaporized and then bombarded by a high-energy electron beam, causing them to fragment as ions.

mediator A protein complex required for the transcription of almost all RNAP II promoters; a signal integration platform.

membrane potential The potential difference across the membrane of living cells; usually measured in millivolts.

Meselson–Radding model A model of general recombination that explains several phenomena not explained by the Holliday model (e.g., sometimes only one of two homologous chromosomes has a recombinant strand).

messenger RNA (mRNA) An RNA species produced by transcription that specifies the amino acid sequence of a polypeptide.

metabolic syndrome A cluster of clinical disorders that include obesity, hypertension, dyslipidemia, and insulin resistance.

metabolism The sum of all chemical reactions in an organism.

metabolism The total of all chemical reactions in an organism.

metabolome The complete set of organic metabolites that are produced within a cell under the direction of the genome.

metalloprotein Conjugated proteins containing metal ions.

methyl-CpG Methylated cytosine in CpG dinucleotides within CpG islands.

methyl-CpG–binding protein Mediates chromatin-associated gene silencing by binding preferentially to the 5-MeCpG dinucleotides and recruiting histone deacetylase to the site along with histone methylases.

micelle An aggregation of molecules having a nonpolar and a polar component, leaving the polar domains facing the surrounding water.

microautophagy A process in which small amounts of cytoplasm are directly engulfed by lysosomes.

microbiota A multicellular organism's indigenous microbial flora.

microfilament A type of cytoskeletal fiber (5–7 nm) composed of polymers of globular actin (G-actin).

microheterogeneity Variations in the glycan components of each type of glycoprotein.

microRNAs (miRNA) 22-nt ncRNAs that inhibit the translation of target mRNAs by binding to partially complementary sequences in the mRNA's 3′-UTR.

microsatellite DNA sequences of 1 to 4 bp that are tandemly repeated 10 to 20 times.

microsome A membranous vesicle derived from fragments of endoplasmic reticulum obtained by differential centrifugation.

microtubule A component of the cytoskeleton; composed of the protein tubulin.

minichromosome maintenance (MCM) complex The major DNA helicase in eukaryotes.

minisatellite Tandemly repeated sequence of about 25 bp with total lengths between 10^2 and 10^5 bp.

mismatch repair A single-strand repair mechanism that corrects helix-distorting base mispairings that are the result of replication proofreading errors or as the result of replication slippage.

missense mutation A point mutation that results in coding for a different amino acid causing a change in a polypeptide's structure and function.

mitochondrial fission The division of a mitochondrion to form two mitochondria.

mitochondrial fusion The merging of two or more mitochondria.

mitochondrion (*pl* mitochondria) An organelle possessing two membranes in which aerobic respiration occurs.

mixed anhydride An acid anhydride with two different R groups.

mixed terpenoid A biomolecule that is composed of a nonterpene component attached to an isoprenoid group.

mobile genetic element One of numerous DNA sequences that can be duplicated and move within the genome.

mobile phase The moving phase in chromatographic methods.

modular (mosaic) protein A protein that contains numerous duplicate or imperfect copies of one or more domains that are linked in series; also known as a mosaic protein.

modulator A ligand whose binding to an allosteric site of an enzyme alters the enzyme's activity.

molecular biology The science devoted to elucidating the structure and function of genomes.

molecular chaperone A molecule that assists in protein folding; many are heat shock proteins.

molecular disease A disease caused by a mutated gene.

molten globule A partially globular state of a folding polypeptide that resembles the molecule's native state.

monosaccharide A polyhydroxy aldehyde or ketone containing at least three carbon atoms.

monounsaturated Describing a fatty acid with a single double bond.

motor protein Components of biological machines that bind nucleotides; nucleotide hydrolysis drives precise changes in the protein's shape.

mutarotation A spontaneous process in which the α and β forms of monosaccharides are readily interconverted.

mutation Any change in the nucleotide sequence of a gene.

***N*-heteroglycan** An asparagine-linked oligosaccharide.

nascent Newly synthesized.

negative cooperativity The binding of one ligand to a target enzyme or protein, decreasing the likelihood of subsequent ligand binding.

neurotransmitter A molecule released at a nerve terminal that binds to and influences the function of other nerve cells

or muscle cells.

neutral fat Triacylglycerol molecules.

nicotinamide adenine dinucleotide (NAD) A coenzyme form of nicotinic acid containing an *N*-ribosyl derivative of nicotinamide and adenosine linked through a pyrophosphate group; occurs as the oxidized form, NAD^+, and the reduced form, NADH and is involved in electron transfer in a class of enzymes called dehydrogenases.

nicotinamide adenine dinucleotide phosphate (NADP) A coenzyme form of nicotinic acid containing an *N*-ribosyl derivative of nicotinamide and adenosine linked through a pyrophosphate group with an additional phosphate group attached at the 2′-OH group of the adenosine sugar; occurs as the oxidized form $NADP^+$ and the reduced form NADPH and is involved in electron transfer in a class of enzymes called dehydrogenases.

nitrogen fixation Conversion of molecular nitrogen (N_2) into a reduced biologically useful form (NH_3) by nitrogen-fixing microorganisms.

nonalkylating agents A variety of chemicals other than the alkylating agents that can modify DNA structure.

noncoding RNAs (ncRNAs) Types of RNA other than the RNAs involved in protein synthesis (i.e., tRNAs, rRNAs, and mRNAs) that act as an extensive genome regulatory network.

noncompetitive inhibition Inhibition of an enzyme in which the inhibitor binds to both the free enzyme and the enzyme-substrate complex.

nonessential amino acid An amino acid that can be synthesized by the body.

nonessential fatty acid A fatty acid that can be synthesized by the body.

nonsense mutation A point mutation that changes the code for an amino acid into a premature stop signal.

nuclear envelope The double membrane that separates the nucleus from the cytoplasm.

nuclear lamina A dense protein meshwork attached to the inner surface of the inner nuclear membrane; involved in nuclear processes such as DNA replication, transcription, and chromatin organization.

nuclear matrix The cytoskeleton-like scaffold within the nucleus in which loops of chromatin are organized.

nuclear pore complex One of thousands of pore complexes in the nuclear envelope through which pass most of the molecules that enter or leave the nucleus of a cell.

nuclease An enzyme that hydrolyzes nucleic acid molecules to form oligonucleotides.

nucleic acid A macromolecule formed from the polymerization of nucleotides.

nucleoid In prokaryotes, an irregularly shaped region that contains a long circular DNA molecule.

nucleolus A structure revealed in the nucleus when the nucleus is stained with certain dyes: it plays a major role in the synthesis of ribosomal RNA and ribosomal subunits.

nucleophile An electron-rich atom or molecule.

nucleophilic substitution A reaction in which a nucleophile substitutes for an atom or molecular group.

nucleoplasm The gelatinous substance within the nucleus that contains the cytoskeleton-like nuclear matrix and a network of chromatin fibers.

nucleoside A biomolecule composed of a pentose sugar (ribose or deoxyribose) and a nitrogenous base.

nucleosome A repeating structural element in eukaryotic chromosomes composed of a core of eight histone molecules around which about 140 base pairs of DNA are wrapped; an additional 60 base pairs connect adjacent nucleosomes.

nucleotide A biomolecule composed of a five-carbon sugar (ribose or deoxyribose), a nitrogenous base, and one or more phosphate groups.

nucleotide excision repair Bulky lesions of 2 to 30 nucleotides are removed and the resulting gap is filled; the excision enzymes appear to recognize the physical distortion rather than a specific base sequence.

nucleus A double membrane-bound organelle in eukaryotic cells that contains the cell's genome.

***O*-heteroglycan** A mucin-type polysaccharide.

obligate aerobe An organism that is highly dependent on oxygen for energy production.

obligate anaerobe An organism that grows only in the absence of oxygen.

Okazaki fragment Any of a series of deoxyribonucleotide segments that are formed during discontinuous replication

of one DNA strand as the other strand is continuously replicated.

oligomer A multisubunit protein in which some or all subunits are identical.

oligonucleotide A short nucleic acid segment that contains fewer than 50 nucleotides.

oligosaccharide An intermediate-size carbohydrate composed of 2 to 15 monosaccharides.

omega-3-fatty acid α-Linolenic acid and its derivatives, such as eicosapentaenoic acid and docosahexaenoic acid.

omega-6-fatty acid Linoleic acid and its derivatives.

oncogene A mutated version of a protooncogene that promotes abnormal cell proliferation.

one-carbon metabolism A set of reactions in which single carbon atoms are transferred from one molecule to another.

open reading frame (ORF) A series of triplet base sequences in mRNA that does not contain a stop codon.

operator A regulatory DNA sequence to which a repressor protein binds.

operon A set of functionally linked genes that are regulated as a unit; occurs in prokaryotes.

operon A set of linked genes that are regulated as a unit.

optical isomer A stereoisomer that possesses one or more chiral centers.

organelle A membrane-enclosed structure within a eukaryotic cell.

origin-of-replication (ORC) complex A protein complex that binds to a DNA replication origin during the initiation phase of eukaryotic DNA synthesis; contains analogues of the protein DnaA.

osmosis The diffusion of solvent through a semipermeable membrane.

osmotic diuresis A process in which solutes in the urinary filtrate cause excessive loss of water and electrolytes.

osmotic pressure The force necessary to resist the movement of water across a semipermeable membrane.

outer mitochondrial membrane The porous external membrane of the mitochondrion.

outer nuclear membrane A membrane surrounding the nucleus that is continuous with the endoplasmic reticulum.

oxidation The removal of electrons from an atom or molecule.

oxidative phosphorylation The synthesis of ATP coupled to electron transport.

oxidative stress Excessive production of reactive oxygen species.

oxidizing agent A substance that oxidizes (removes electrons from) another substance; the oxidizing agent is itself reduced in the process.

oxidoreductase An enzyme that catalyzes an oxidation/reduction reaction.

oxyanion A negatively charged oxygen atom.

palindrome A sequence that provides the same information whether it is read forward or backward; DNA palindromes contain inverted repeat sequences.

Pasteur effect The observation that glucose consumption is greater under anaerobic conditions than when oxygen is available.

pentose phosphate pathway A biochemical pathway that produces NADPH, ribose, and several other sugars.

peptide An amino acid polymer with fewer than 50 amino acid residues.

peptide bond An amide linkage in an amino acid polymer.

peptidyl transfer center (PTC) The location of peptidyl transferase activity within the large subunit of bacterial ribosomes; located in a domain of 23S rRNA.

perinuclear space The space between the two membranes of the nuclear envelope.

peripheral protein A protein that is not embedded in the membrane but attached either by a covalent bond to a lipid molecule or by noncovalent interactions with a membrane protein or lipid.

pH optimum The pH value at which an enzyme's activity is maximal.

pH scale A measure of hydrogen ion concentration; pH is the negative log of the hydrogen ion concentration in moles per liter.

phase I reaction A biotransformation reaction involving oxidoreductases and hydrolases that converts hydrophobic substances into more polar molecules.

phase II reaction A biotransformation reaction in which metabolites containing appropriate functional groups are conjugated with substances such as glucuronate, glutamate, sulfate, and glutathione.

phosphoglyceride A type of lipid molecule found predominantly in membrane; composed of glycerol linked to two fatty acids, phosphate, and a polar group.

phosphoprotein A conjugated protein in which phosphate is the prosthetic group.

phosphoryl group transfer potential The tendency of a phosphorylated molecule to undergo hydrolysis.

photoautotrophs Organisms that transform light energy (usually from the sun) into chemical bond energy.

photoheterotrophs Organisms that use both light and biomolecules as energy sources.

photophosphorylation The synthesis of ATP coupled to electron transport driven by light energy.

photoreactivation repair A mechanism to repair thymine dimers using the energy of visible light.

photosynthesis The trapping of light energy and its conversion to the chemical energy required to incorporate carbon dioxide into organic molecules.

photosystem A photosynthetic mechanism composed of light-absorbing pigments.

piwi-interacting RNA (piRNA) Single-stranded RNAs (26–31 nt) that recognize and silence retrotransposons in the mammalian testes. They also have roles in somatic euchromatin remodeling and developmental processes.

plasma membrane The membrane that surrounds a cell, separating it from its external environment.

plasmid A small circular DNA molecule that can exist and replicate independently of a bacterial chromosome; plasmids are stably inherited but are not required for the host cell's growth and reproduction.

point mutation A change in a single nucleotide base in a DNA sequence.

polar An unequal distribution of electrons in a bond.

polar head group A molecular group that contains phosphate or other charged or polar groups.

poly(A) binding protein (PABP) Forms a circular mRNA molecule, during the initiation phase of eukaryotic translation, by interacting with the 3′-poly(A) tail and 5′-capped end of the mRNA and eIF-G, a translation initiation factor.

polymerase chain reaction (PCR) A laboratory technique that uses a heat-stable DNA polymerase to synthesize large quantities of specific nucleotide sequences from small amounts of DNA.

polypeptide An amino acid polymer with more than 50 amino acid residues.

polysaccharide A linear or branched polymer of monosaccharides linked by glycosidic bonds.

polysome An mRNA with several ribosomes bound to it.

polyunsaturated Describing a fatty acid with two or more double bonds, usually separated by a methylene group.

positive cooperativity The mechanism in which the binding of one ligand to a target enzyme or protein increases the likelihood of subsequent ligand binding.

postabsorptive The phase in the feeding–fasting cycle in which nutrient levels are low.

postprandial The phase in the feeding–fasting cycle immediately after a meal: blood nutrient levels are relatively high.

posttranslational modification One of a set of reactions that alter the structure of newly synthesized polypeptides.

posttranslational translocation The transfer of previously synthesized polypeptides across the membrane of an organelle.

PPAR Peroxisome proliferator-activated receptors.

preinitiation replication complex A multisubunit protein complex, formed at the origin of replication during the initiation step of eukaryotic DNA synthesis.

prenylation The covalent attachment of prenyl groups (e.g., farnesyl and geranylgeranyl groups) to protein molecules.

preproprotein An inactive precursor protein with a removable signal peptide.

primary cilium A nonmotile cilium that functions as a sensory organelle on the surface of most differentiated vertebrate cells.

primary structure The amino acid sequence of a polypeptide.

primase An RNA polymerase that synthesizes short RNA segments, called primers, that are required for DNA

synthesis.

primer A short RNA segment required to initiate DNA synthesis.

primosome A multienzyme complex involved in the synthesis of RNA primers at various intervals along the DNA template strand during *E. coli* DNA replication.

processivity The prevention of frequent dissociation of a polymerase from the DNA template.

proenzyme An inactive precursor of an enzyme.

promoter The sequence of nucleotides immediately before a gene that is recognized by RNA polymerase and signals the start point and direction of transcription.

prophage A viral genome integrated into host cell DNA.

prostaglandin An arachidonic acid derivative that contains a cyclopentane ring with hydroxyl groups at C-11 and C-15.

prosthetic group The nonprotein portion of a conjugated protein that is essential to the biological activity of the protein; often a complex organic group.

proteasome A multienzyme complex that degrades proteins linked to ubiquitin.

protein A macromolecule composed of one or more polypeptides.

protein family A group of protein molecules that are related by amino acid sequence similarity.

protein folding The process in which an unorganized polypeptide acquires a highly organized and relatively stable three-dimensional structure.

protein turnover The continuous degradation and synthesis of proteins in an organism.

proteoglycan A large molecule containing large numbers of glycosaminoglycan chains linked to a core protein molecule.

proteome The complete set of proteins produced within the cell.

proteomics The investigation of protein synthesis patterns and protein–protein interactions.

proteostasis A process by which cells control the folding of proteins.

proteostasis network A series of proteins organized into pathways that control the folding, trafficking, and degradation of proteins.

protomer A component of an oligomer; may consist of one or more subunits.

protonmotive force The force arising from a gradient of protons and a membrane potential.

protooncogene A normal gene that codes for a protein involved in cell cycle regulation; promotes carcinogenesis if mutated.

pseudogene An imperfect copy of a functional gene, which is not expressed.

purine A nitrogenous base with a two-ring structure; a component of nucleotides.

pyrimidine A nitrogenous base with a single-ring structure; a component of nucleotides.

Q cycle The movement of electrons from reduced coenzyme Q, UQH_2, to cytochrome c during electron transport.

quaternary structure Association of two or more folded polypeptides to form a functional protein.

racemization The interconversion of enantiomers.

radical An atom or molecule with an unpaired electron.

reaction center The membrane-bound protein complex in a photosynthesizing cell that mediates the conversion of light energy into chemical energy.

reaction mechanism Step-by-step description of a chemical reaction process.

reactive nitrogen species (RNS) Nitrogen-containing radicals often classified as ROS; the most important are nitric oxide, nitrogen dioxide, and peroxynitrite.

reactive oxygen species (ROS) A reactive derivative of molecular oxygen, including superoxide radical, hydrogen peroxide, the hydroxyl radical, and singlet oxygen.

receptor protein A protein with binding sites for extracellular ligands (signal molecules).

receptor tyrosine kinase (RTK) A transmembrane receptor that contains a cytoplasmic domain with tyrosine kinase activity that is activated when a ligand is bound to the external domain.

recombinant DNA technology A series of techniques whose essential feature is that DNA molecules obtained from various sources can be cut and spliced together.

recombination A process in which DNA molecules are cut and rejoined in new combinations.

reducing agent A substance that reduces the oxidation number of another reactant; the reducing agent is itself oxidized in the process.

reducing sugar A sugar that can be oxidized by weak oxidizing agents.

reduction potential The tendency for a specific substance to lose or gain electrons.

reduction The addition of electrons to an atom or molecule.

releasing factor A protein involved in the termination phase of translation.

replication factor C (RFC) A clamp loader protein that controls the attachment of DNA polymerase δ to each DNA strand.

replication factories Specific nuclear compartments (or nucleoids) in which DNA replication occurs.

replication fork The Y-shape region of a DNA molecule that undergoes replication; results from separation of two DNA strands.

replication licensing factor One of several proteins that bind to the origin-of-replication complex (ORC) and complete the structure of the preRC.

replication protein A (RPA) A protein that stabilizes the separated DNA strands during replication.

replication The process in which an exact copy of parental DNA is synthesized using the polynucleotide strands of the parental DNA as templates.

replicon A DNA molecule or a segment of DNA that contains a single origin of replication.

replisome The large complex of polypeptides, including the primosome, that replicates DNA in *E coli*.

resonance energy transfer The transfer of energy from an excited molecule to a nearby molecule, thereby exciting the second molecule.

resonance hybrid A molecule with two or more alternative structures that differ only in the position of electrons.

respirasome A functional aerobic respiration unit in the inner mitochondrial membrane; the I, III_2, $IV_{1–2}$ supercomplex has been identified in animals, plants, and fungi.

respiration A biochemical process whereby fuel molecules are oxidized and their electrons are used to generate ATP.

respiratory burst An oxygen-consuming process in scavenger cells such as macrophages in which reactive oxygen species are generated and used to kill foreign or damaged cells.

respiratory control The control of aerobic respiration by ADP concentration.

response element A DNA sequence within the promoter of genes; transcription is triggered when a specific hormone receptor complex or transcription factor binds.

retrotransposon A subclass of transposons that use an RNA intermediate.

retrovirus One of a group of viruses with RNA genomes that carry the enzyme reverse transcriptase and form a DNA copy of their genome during the reproductive cycle.

reversible inhibition A form of enzyme inhibition in which the inhibitory effect of a compound can be counteracted by increasing substrate or removing the inhibitor while the enzyme remains intact.

rho-dependent termination Transcription termination in prokaryotes that requires the rho factor protein.

rho-independent termination A form of transcription termination in prokaryotes that does not involve rho factor; also referred to as intrinsic termination.

ribosomal RNA (rRNA) The RNA present in ribosomes. Ribosomes contain several types of single-stranded ribosomal RNA that contribute to ribosome structures and are also directly involved in protein synthesis.

ribosome A protein-RNA complex that is the site of protein biosynthesis.

ribosome recycling factor In bacteria, a tRNA-shaped protein; it binds within the A site and causes the dissociation of the ribosomal subunits after polypeptide synthesis.

riboswitch An RNA-based control mechanism made up of a specific untranslated sequence within an mRNA. Usually, the action of the riboswitch, triggered by a change in its tertiary structure induced by binding to a ligand, is to repress translation.

ribozyme A catalytic RNA molecule; catalyzes self-cleavage or the cleavage of other RNAs.

RNA editing The alteration of the base sequence in a newly synthesized mRNA molecule; bases may be chemically modified or deleted.

RNA interference A cellular mechanism in which RNA molecules are degraded; functions in gene expression regulation and in defense against viral RNA genomes.

RNA splicing The process in which introns are cut out and the exons are linked together to form a functional RNA product.

RNA transposon A transposable element that is formed by a mechanism that involves an RNA transcript; also referred to as retrotransposon.

RNA-induced silencing complex (RISC) A ribonucleoprotein that binds a single-stranded miRNA in the regulation of mRNA transcripts, or double-stranded siRNAs in the targeting of viral mRNAs. The antisense strand of the siRNA bound to the RISC complex targets and anneals with the viral RNA; nucleases within the RISC then degrade the viral sequence.

rough endoplasmic reticulum (RER) A type of endoplasmic reticulum that has ribosomes bound to its external surface; newly synthesized polypeptides are translocated through the RER membrane.

***S*-adenosylmethionine (SAM)** The major methyl group donor in one-carbon metabolism.

salt bridge An electrostatic interaction in proteins between ionic groups of opposite charge.

salting out The decrease in protein solubility caused by an increase in the ionic strength of the solution.

satellite DNA DNA sequences that are highly repetitive; when genomic DNA is digested and centrifuged, a satellite band forms.

saturated Describes a molecule that contains no carbon–carbon double or triple bonds.

Schiff base The imine product of a reaction between a primary amino group and a carbonyl group; also referred to as an aldimine.

SDS-polyacrylamide gel electrophoresis A method for separating proteins or determining their molecular weights that employs the negatively charged detergent SDS.

SECIS element Selenocysteine insertion sequence element: a base sequence required at the 3′-UTR (3′ untranslated region) of the mRNAs for all selenoproteins for the insertion of selenocysteine during translation.

second messenger A molecule that mediates the action of some hormones.

secondary structure The arrangement of a polypeptide chain into locally organized structures of α-helix and β-strands: secondary structure is maintained by hydrogen bonds between the amide hydrogen and the carbonyl oxygen of peptide bonds.

semiconservative replication DNA synthesis in which each polynucleotide strand serves as a template for the synthesis of a new strand.

sense strand The nontranscribed DNA strand; the DNA version of the mRNA used to synthesize the polypeptide product of a gene.

Shine-Dalgarno sequence A purine-rich sequence that occurs on an mRNA close to AUG (the initiation codon) that binds to a complementary sequence on the 30S ribosomal subunit, thereby promoting the formation of the correct preinitiation complex.

short tandem repeats DNA sequences with between 2- and 4-bp repeats; can be used to generate DNA profiles that distinguish among individuals.

shotgun cloning A cloning technique in which genomic libraries are created by the random digestion of a genome.

signal hypothesis A mechanism that explains how secreted or membrane proteins are synthesized on ribosomes bound to the rough endoplasmic reticulum; a sequence of amino acid residues on the nascent polypeptide chain that mediates the insertion of the polypeptide into the RER membrane.

signal peptide A short sequence typically near the amino terminal of a polypeptide that determines its insertion into a membrane of an organelle.

signal peptide A short sequence typically near the amino terminal of a polypeptide that determines its insertion into a prokaryotic plasma membrane or a specific eukaryotic membrane.

signal recognition particle A large multisubunit ribonucleoprotein complex that mediates the binding of the ribosome and the emerging signal peptide to the rough endoplasmic reticulum during protein synthesis; facilitates the passage of the growing polypeptide through the RER membrane.

signal transduction The mechanisms by which extracellular signals are received, amplified, and converted to a cellular response.

signaling cascade An information processing mechanism initiated by the binding of a signal molecule to a receptor and continued by a series of events involving protein conformational changes and covalent modifications that result in a response; responses include changes in enzyme activities, cytoskeletal rearrangements, cell movement, or cell cycle progression.

silencer A DNA sequence that when bound to a repressor protein prevents an RNA polymerase from binding to a nearby promoter.

silent mutation A base change in DNA that has no discernable effect.

SINE (short interspersed nuclear elements) A repeating DNA sequence less than 500 bp long interspersed in mammalian genomes; SINEs cannot undergo transposition without the aid of a functional LINE sequence.

single nucleotide polymorphism A point mutation that occurs in a population to any extent.

site-specific recombination Recombination of nonhomologous DNA sequences; a recombination mechanism between sequences with limited homology that depends on protein-DNA interactions.

small interfering RNA (siRNA) A 21- to 23-nt dsRNA that plays a crucial role in RNA interference.

small nuclear ribonucleoprotein particle (snRNP) A complex of proteins and small nuclear RNA molecules that promotes mRNA processing.

small nuclear RNA (snRNA) A small RNA molecule involved in the removal of introns from mRNA.

small nucleolar RNA (snoRNA) An RNA component of nucleolar ribonucleoprotein that facilitates chemical modifications of rRNA, tRNA and snRNA molecules.

smooth endoplasmic reticulum (SER) A type of endoplasmic reticulum involved in lipid synthesis and biotransformation processes.

solvation sphere A shell of water molecules that clusters around positive and negative ions.

Southern blotting A laboratory technique in which radioactively labeled DNA or RNA probes are used to locate a complementary sequence in a DNA digest.

specificity constant In enzyme kinetics, the term k_{cat}/K_m; the second order rate constant for a reaction in which [S] << K_m.

sphingolipid A membrane lipid molecule that contains a long-chain amino alcohol called sphingosine; ceramide is a fatty acid derivative of sphingosine; an important component of animal membranes.

sphingomyelin A type of phospholipid that contains sphingosine; the 1-hydroxyl group of ceramide is esterified to the phosphate group of phosphorylcholine or phosphorylethanolamine, and the amino group of sphingosine is in an amide linkage with a fatty acid.

spliceosome A multicomponent complex containing protein and RNA; used in the splicing phase of mRNA processing.

spontaneous chemical changes Physical or chemical processes that are accompanied by a release of energy.

SREBP (sterol regulatory element binding protein) One of several transcription factors that are membrane proteins in the endoplasmic reticulum or the Golgi apparatus.

SRP receptor protein A heterodimer with two GTPases on the cytoplasmic surface of the RER that binds SRP (signal recognition particle), the protein complex that facilitates the binding of a ribosome to the membrane surface during the synthesis of a signal peptide-containing polypeptide; also referred to as docking protein.

standard reduction potential The measurement of the capacity of a substance to gain or lose electrons in a galvanic cell fitted with a standard hydrogen electrode set at 0.00 V.

stationary phase The solid matrix in chromatographic techniques.

steady state A phase in an organism's life when the rate of anabolic processes is approximately equal to that of

catabolic processes.

steroid A derivative of triterpenes; contains four fused rings.

stereoisomer A molecule that has the same structural formula and bonding pattern as another but has a different arrangement of atoms in space.

sterol carrier protein A cytoplasmic protein carrier for certain intermediates during cholesterol biosynthesis.

stroma A dense, enzyme-filled substance that surrounds the thylakoid membrane within the chloroplast.

stromal lamella A thylakoid membrane segment that interconnects two grana.

substrate The reactant in an enzyme-catalyzed reaction that is converted to a product.

substrate-level phosphorylation The synthesis of ATP from ADP by phosphorylation coupled with the exergonic breakdown of a high-energy organic substrate molecule.

subunit A polypeptide component of an oligomeric protein.

sucrose A disaccharide composed of α-glucose and β-fructose residues linked through a glycosidic bond between both anomeric carbons.

superfamily The largest grouping of proteins for which common ancestry can be inferred from sequence homology.

supersecondary structure (or **motif**) One of a set of specific combinations of α-helix and β-pleated-strand structures of protein molecules.

supraspliceosome Formed by four active spliceosomes and a pre-mRNA complex, it increases the speed and efficiency of transcript splicing and provides opportunities for intron excision proofreading.

surface active agent A substance that lowers the surface tension of a liquid, usually water, so that it spreads out over a surface.

synthesis-dependent strand-annealing model A form of double-strand repair using homologous chromosomes in which there are only noncrossover products.

systems biology A field of study based on engineering principles in which the interactions between the components of living organisms are investigated; complex data sets used by systems biologists come from genomics, proteomics, and experimental sources such as protein–protein interactions and biochemical reaction fluxes.

T cell A T lymphocyte: a white blood cell that bears antibody-like molecules on its surface and binds to and destroys foreign cells in cellular immunity.

tandem repeats DNA sequences in which multiple copies are arranged next to one another; lengths of repeated sequences vary from 1 bp to more than 2000 bp.

target cell A cell that responds to the binding of a hormone or growth factor to a receptor protein.

targeting The process that directs newly synthesized proteins to their correct intra- or extracellular destinations.

tautomer An isomer that differs from another in the location of a hydrogen atom and a double bond (e.g., keto-enol tautomers).

tautomerization Chemical reaction by which two tautomers are interconverted by the movement of a hydrogen atom and a double bond.

telomerase A ribonucleoprotein with an RNA component complementary to the TG-rich telomere sequence.

telomere A structure found at both ends of a chromosome that buffers the loss of critical coding sequences after a round of DNA replication.

telomere end-binding protein (TEBP) A protein that binds to and stabilizes GT-rich telomere sequences.

telomere repeat-binding factor (TRF) A protein that binds to and secures the 3′ overhang sequence of a telomere.

termination The phase of translation in which a newly synthesized polypeptide is released from the ribosome.

terpene A member of a class of isoprenoids classified according to the number of isoprene residues it contains.

tertiary structure The globular three-dimensional structure of a polypeptide that results from interactions between the side chains (R groups) of the amino acid residues.

tetrahydrofolate The biologically active form of the B vitamin folic acid; a carrier of methyl, methylene, methenyl, and formyl groups in one-carbon metabolism.

thermodynamics The study of energy and its interconversion.

thiamine pyrophosphate The coenzyme form of thiamine, also called vitamin B1.

thiolytic cleavage Cleavage of a carbon–sulfur bond.

thromboxane A derivative of arachidonic acid that contains a cyclic ester.

thylakoid lumen The internal compartment created by the formation of grana.

thylakoid membrane An intricately folded internal membrane within the chloroplast.

topologically associated domain A self-interacting region of DNA sequences that preferentially interact with each other.

transamination A reaction in which an amino group is transferred from one molecule to the α-carbon of an α-keto acid; the amino acid that donates the amino group is converted to the corresponding α-keto acid.

transcript An RNA molecule that is produced by the transcription of a DNA sequence.

transcript localization The binding of mRNAs to certain cellular structures within cytoplasm that permits the creation of protein gradients within the cell.

transcription coupled repair A form of DNA repair in which repair only occurs on the DNA strand that is being transcribed.

transcription factor A protein that regulates or initiates the synthesis of specific mRNAs by binding to DNA sequences called response elements.

transcription preinitiation complex A multisubunit protein complex, formed during the initiation phase of eukaryotic transcription.

transcription The process in which single-stranded RNA with a base sequence complementary to the template strand of DNA is synthesized.

transcriptome The complete set of RNA molecules that are produced within a cell.

transduction The transfer of DNA segments between bacteria by bacteriophage.

transduction The transfer of DNA segments between bacteria by bacteriophages.

transfection A mechanism by which bacteriophage inadvertently transfers bacterial chromosome or plasmid sequences to a new host cell.

transfer RNA (tRNA) A small RNA molecule that binds to an amino acid and delivers it to the ribosome for incorporation into a polypeptide chain during translation.

transferase An enzyme that catalyzes the transfer of a functional group from one molecule to another.

transformation A process in which DNA fragments enter a bacterial cell and are introduced into the bacterial genome.

transition mutation A DNA mutation that involves the substitution of a purine base by a different purine, or the substitution of a pyrimidine by a different pyrimidine.

transition state In catalysis, the unstable intermediate formed by the enzyme that has altered the substrate so that it now shares properties of both the substrate and the product.

translation Protein synthesis; the process by which the genetic message carried by mRNAs directs the synthesis of polypeptides with the aid of ribosomes and other cell constituents.

translocation Movement of the ribosome along the mRNA during translation; also refers to a chromosomal abnormality in which a DNA fragment inserts into another site in the same chromosome or in a nonhomologous chromosome.

translocation Refers to a chromosomal abnormality in which a DNA fragment inserts into another site in the same chromosome or in a nonhomologous chromosome.

translocon An integral membrane protein that mediates translocation of a polypeptide.

transposable DNA elements A DNA sequence that excises itself and then inserts at another site.

transposition The movement of a DNA sequence from one site in a genome to another.

transposon (transposable element) A DNA segment that carries the genes required for transposition and moves about the chromosome; sometimes the name is reserved for transposable elements that also contain genes unrelated to transposition.

transulfuration pathway A biochemical pathway that converts methionine to cysteine.

transversion mutation A type of point mutation in which a pyrimidine is substituted for a purine or vice versa.

triacylglycerol cycle A metabolic mechanism that regulates the level of fatty acids that are available to the body for energy generation and the synthesis of molecules such as phospholipids; the constant synthesis and degradation of TG.

tumor promoter A molecule that provides cells with a growth advantage over nearby cells.

tumor suppressor gene One of a set of genes that code for a protein that protects cells from progressing toward cancer; may inhibit a transcription factor required for cell cycle progression or may facilitate DNA repair.

turnover The rate at which all molecules in a cell are degraded and replaced with newly synthesized molecules.

type 1 diabetes An autoimmune disease that destroys the insulin-producing β-cells in the pancreas; a metabolic disease in which the most obvious symptoms are hyperglycemia and dyslipidemia.

type 2 diabetes A form of diabetes mellitus in which patients are resistant to insulin.

ubiquitin A protein that is covalently attached by enzymes to proteins destined to be degraded.

ubiquitin proteosomal system An elaborate mechanism for the rapid destruction of proteins.

ubiquitination The covalent attachment of ubiquitin to proteins that are to be degraded.

uncompetitive inhibition An inhibitor binds only to the enzyme-substrate complex; a rare type of noncompetitive inhibition.

uncoupler A molecule that uncouples ATP synthesis from electron transport; it collapses a proton gradient by transporting protons across the membrane.

uncoupling protein A molecule that dissipates the proton gradient in mitochondria by translocating protons; UCP1 is also called thermogenin.

unfolded protein response The inhibition of new protein synthesis in the RER except for molecular chaperones; triggered by severe ER stress.

unsaturated Describes a molecule that contains one or more carbon–carbon double or triple bonds.

uronic acid The product formed when the terminal CH_2OH of a monosaccharide is oxidized.

van der Waals force A class of relatively weak, transient electrostatic interactions between permanent and/or induced dipoles.

vector A cloning vehicle into which a segment of foreign DNA can be spliced, ready for introduction into host cells and expression in them.

velocity The rate of a biochemical reaction; the change in the concentration of a reactant or product per unit time.

very-low-density lipoprotein (VLDL) A type of lipoprotein with a very high relative concentration of lipids; transports lipids to tissues.

vesicles Membranous sacs that bud off from a donor membrane and subsequently fuse with the membrane of another organelle or with the plasma membrane.

vitamin An organic molecule required by organisms in minute quantities; some vitamins are coenzymes required for the function of certain enzymes.

vitamin B_{12} A complex cobalt-containing molecule that is required for the N^5-methyl THF-dependent conversion of homocysteine to methionine.

wax A complex mixture of nonpolar lipids including wax esters.

wax ester One of numerous types of ester consisting of long-chain fatty acids and long-chain alcohols; a prominent constituent of most waxes.

weak acid An organic acid that does not completely dissociate in water.

weak base An organic base that has a small but measurable capacity to combine with hydrogen ions.

wobble hypothesis The hypothesis that explains why cells often have fewer tRNAs than expected; freedom in the pairing of the third base of the codon to the first base of the anticodon allows some tRNAs to pair with several codons.

work A physical change caused by a change in energy.

yeast artificial chromosome A cloning vector that can accommodate up to 3000 kb of DNA; contains eukaryotic sequences that function as centromeres, telomeres, and a replication origin.

Z scheme A mechanism whereby electrons flow between photosystems II and I during photosynthesis.

Z-DNA A form of DNA that is twisted into a left-handed spiral; named for the zigzag conformation, which is slimmer than that of B-DNA.

zwitterion A neutral molecule that bears an equal number of positive and negative charges simultaneously.

zymogen The inactive form of an enzyme that is converted into an active enzyme by a proteolytic cleavage.

***α*-tocopherol** A radical scavenger belonging to a class of compounds called phenolic antioxidants.

***β*-carotene** A plant pigment molecule that acts as an absorber of light energy and as an antioxidant.

***β*-oxidation** The catabolic pathway in which most fatty acids are degraded; acetyl-CoA is formed as the bond between the α and β carbon atoms is broken.

Credits

Chapter 1

Opener: Elaine Thompson/Associated Press

Chapter 2

Opener: STEVE GSCHMEISSNER/SCIENCE PHOTO LIBRARY; **2.5:** Adapted from Bruce Alberts "Essential Cell Biology" 1998 Routledge/Taylor & Francis; **2.6:** Reprinted from Trends in Biochemical Sciences 26(10), R. John Ellis, "Macromolecular crowding: obvious but underappreciated", pp. 597-604, copyright 2001, with permission of Elsevier.; **2.8:** Adapted from Prescott, et al., Microbiology 4e. Copyright 1999 by the McGraw-Hill Companies; **2.9:** Adapted from Prescott, et al., Microbiology 4/e. Copyright 1999 by the McGraw-Hill Companies. Reprinted with Permission of Springer-Verlag Germany; **2.12:** Adapted from 4/e Pearson/Benjamin Cummings, Campbell & Reece, Biology, 7th ed. 2005. Fig 7.7, pg. 127; **2.13:** Courtesy of Audrey M. Glauert and GMW Cook. University of Cambridge.; **2.24b:** Don W. Fawcett/Science Source; **2.26a:** JENNIFER WATERS / SCIENCE PHOTO LIBRARY; **2.26b:** Photo Researchers/Science Source; **2.26c:** ALVIN TELSER/SCIENCE PHOTO LIBRARY; **2.27b and c:** Adapted from Hardin J. et al., Becker's World of the Cell; **2.28:** Donald E. Ingber, Cellular Tensegrity: Defining New Rules of Biological Design that Govern the Cytoskelteton J. Cell Science 104: 613-627 1993 fig. 2; **2A:** Adapted from Cooper GM and Hausman, RE, "The Cell: A Molecular Approach" 1997 Sinauer; **2B:** Adapted from Lehninger Principles of Biochemistry by David Nelson and Michael Cox. Copyright 2000, 1993, 1982 by Worth Publishers permissiondept@worthpubishers.com

Chapter 3

Opener: World Perspectives/Getty Images; **3.1:** ruivale sousa/Shutterstock.com; **3.2:** Adapted from Silverberg, Chemistry 2e, Copyright The McGraw Hill Companies; **3.9:** Adapted from R. Chang, Chemistry 7e McGraw Hill; **3.12:** Adapted from Pearson/Benjamin Cummings, Campbell & Reece, Biology, 7th ed. 2005. Fig 6.27b, p. 117; **3A:** Adapted from Linda Huff, American Scientist, September-October 1997 Volume 85

Chapter 4

Opener: Elana Erasmus; **4.1:** Louise Gubb/Getty

Chapter 5

Opener: MaslovMax/iStock; **5.14:** Adapted from "Molecules of Life" Purdue University; **5.16:** Adapted from Carl Brandon & John Tooze, Intro to Protein Structure, 2nd ed 1999. Garland 3.14.44.; **5.18b:** Adapted from Garrett and Grisham, Biochemistry, 1996, Brooks Cole; **5.19c:** Adapted from C. Brandon & J Tooze, Intro to Protein Structure, 2nd ed 1999, Garland fig 5.3 p. 68; **5.19d:** Adapted from C. Brandon & J Tooze, Intro to Protein Structure, 2nd ed 1999, Garland fig 4.14c p. 58; **5.19e:** Adapted from C. Brandon & J Tooze, Intro to Protein Structure, 2nd ed 1999, Garland fig 10.1 p. 176; **5.20a:** Adapted from Annual Review of Biochemistry, vol 45, 1976; **5.20c:** Adapted from copyright 1999 from Introduction to Protein Structure by Carl Brandon and John Tooze. Reproduced by permission of Routledge Inc., part of Taylor and Francis Group; **5.20d:** Adapted from copyright 1999 from Introduction to Protein Structure by Carl Brandon and John Tooze. Reproduced by permission of Routledge Inc., part of Taylor and Francis Group; **5.20e:** Adapted from copyright 1999 from Introduction to Protein Structure by Carl Brandon and John Tooze. Reproduced by permission of Routledge Inc., part of Taylor and Francis Group; **5.21:** Adapted from Lain D. Campbell and AK Downing, NMR of Modular Proteins Nature Structure Biology NMR Supplement p 496 fig 1; **5.23:** Adapted from a figure published in Biophysical Journal, 72, C. Reid and RP Rand, "Probing Protein Hydration and Conformational States in Solution" p. 1022-1030. Copyright Elsevier (1997); **5.24:** Republished with permission of Harcourt College Publishers from Principles of Biochemistry with a Human Focus 1e by Reginald Garrett, 2002. Permission conveyed through Copyright Clearance Center.; **5.26:** Adapted from HJ Dysom & Pe Wright, Instrinsically Unstructured Proteins and Their Functions, Nature reviews: molecular cell biology, p. 200 Box 2 fig. a; **5.29a:** Adapted from FEBS Letters 498 (2-3), Sergio T. Ferreira, Fernanda G. GeFelices, "Protein dynamics, folding and

misfolding: from basic physical chemistry to conformational disease", copyright 2001, with permission from Elsevier; **5.29b:** from Trends in Biochemical Sciences 25(12), Sheena A. Radford, "Protein Folding: Progress Made and promises ahead" p. 611-618, copyright 2000, with permission from Elsevier; **5.30:** Reprinted from Trends in Biochemical Sciences 25(12), Sheena E. Radford, "Protein folding: progress made and promises ahead", pp. 611-618, copyright 2000, with permission from Elsevier.; **5.31:** Adapted from Xu, Horwich and Stigler, "The Crystal Structure of the Asymmetric Gro-El-Gros-Es (ADP) 7 Chaperonen Complex" Nature Vol 388 August 21, 1997, p. 741-50; **5.33:** Adapted from Garrett & Grisham Biochemistry 1996 Brooks Cole; **5.36:** Adapted from KA Piez in BD Betlander. Ed., The Protein Folding Problem AAAS Selected Symposium 89 American Association for the Advancement of Science, Washington DC 1984 pp. 47-61. Courtesy of the Collagen Corporation. Reproduced by permission of American Association for the Advancement of Science; **5A:** Journal of Cell Bio, Rockefeller U Press; **5B:** Adapted from Volrath, F. and Knight DP "Structure and Function of the Silk Production Pathway in the Spider Nephla ebulis" Int. J. Bio. Macro 24: 243-249, 1999; **5C:** Dennis Kunkel Microscopy, Inc.; **5E:** Adapted from Christopher K. Matthews and KE Van Holde, "Biochemistry 2e" 1996 Benjamin Cummings; **5G:** Adapted from Nelson and Cox Lehringer, Principles of Biochemistry, 4th edition. Freeman/Worth permissionsdept@worth publishers.com fig 1 p 102

Chapter 6

6.6B: Wikimedia Commons/ Protein Data Bank
https://commons.wikimedia.org/wiki/File:Carbonic_anhydrase_1CA2_active_site.png; **6.10:** Leonid Andronov/123RF

Chapter 7

7.40: Adapted from Lehninger's Principles of Biochemistry by David Nelson and Michael Cox 2000, 1993, 1982. WH Freeman

Chapter 8

Opener: MBP-one/Alamy Stock Photo; **8A:** Adapted from Trends in Biochemical Sciences 23(5), Bas Teusink, Michael C. Walsh, Karel van Dam, Hans V. Westerhoff, "The Danger of Metabolic Pathways with Turbo Design" p. 162-169 copyright 1998, with permission from Elsevier; **8.7:** "Illustration of free energy changes in chemical reactions" by Eugene Hamori in Journal of Chemical Education 52(6), June 1, 1975. © American Chemical Society. Reprinted with permission; **8B:** Adapted from JM Thomson et al. Nature Genetics 37:630-35 1 May 2005

Chapter 9

Opener: David Goodsell; **9.10:** With permission of Dr. Z Hong Zhou, Director, Electron Imagery Center for Nanomachiens (EICN)

Chapter 10

Opener: Adapted from Hutcheon, Duncan, Ngai and Cross, Proceedings of the National Academy of the Sciences of the US, Vol 98. 2001. Reprinted with the permission of Richard L. Cross; **10.8:** WH Freeman: Adapted from Nielson and Cox, Lehninger Principles of Biochemistry, 4th edition 2005 WH Freeman Fig 9.14, p. 702; **10.11:** WH Freeman: Adapted from Nielson & Cox, Lehninger's Principles of Biochemistry, 4th edition 2005. WH Freeman Fig 19.14, p. 702; **10.16:** Adapted with permission of the American Association for the Advancement of Science from "Mitochondrial Structure: Two Types of Subunits on Negatively Stained Mitochondrial Membranes" by Donald F Parsons in Science, Volume 140, issue 3570, May 31, 1963. Permission conveyed through Copyright Clearance Center; **10.17:** Adapted from Trends in Biochemical Sciences, Vol 22, Jung, Hill, Engelbrecht, p. 420-423. Reprinted with permission from Elsevier Science; **10.18:** Adapted with permission of Harcourt College Publishers from Principles of Biochemistry with a Human Focus 1e by Reginald Garrett, 2002. Permission conveyed through Copyright Clearance Center.; **10.20:** Schlattner U., Tokarska-Schlattner M., Willimann T., "Mitochondrial creatine kinase in human health and disease." US National Library of Medicine. DOI: 10.1016/j.bbadis.2005.09.004

Chapter 11

Opener: Adapted from Geoffrey Cooper, "The Cell: A Molecular Approach" 1997 Sinauer; **11.3a:** Adapted from Leonard Lessin/Peter Arnold. Inc.; **11.3b:** Adapted from Leonard Lessin/FBPA; **11.6a and b:** Adapted from Leonard Lessin/Peter Arnold Inc; **11.13:** Adapted from Leonard Lessin/Peter Arnold Inc.; **11.8:** Adapted from Leonard Lessin/Peter Arnold Inc.; **11.21:** Adapted from Geoffrey Cooper: The Cell: A Molecular Approach" 1997 Sinauer; **11.27:** Adapted from Geoffrey Cooper: The Cell: a Molecular Approach, 1997 Sinauer; **11.31:** Adapted from Thomas Zeuthan, Trends in Biochemical Sciences, Vol. 26, No 2 p. 77-79 copyright 2001 Elsevier; **11.32:** Reprinted by permission from Springer: Cellular and Molecular Life Sciences, From the endoplasmic reticulum to the plasma membrane: mechanisms of CFTR folding and trafficking, Carlos M. Farinha and Sara Cnato, 2016.; **11A:** Mahley RW., "Apolipoprotein E: from cardiovascular disease to neurodgeneratigve disorders." US National Library of Medicine. doi: 10.1007/s00109-016-1427-y

Chapter 12

Opener: Henn Photography/Getty Images; **12.15:** Adapted from Annual Review of Biochemistry, Vol 52, 1983; **12.19a:** With permission of Harcourt College Publishers from Principles of Biochemistry with a Human Focus 1e by Reginald Garrett 2002. Permission conveyed through Copyright Clearance Center.; **12.19b:** Adapted from International J. Biochem Mol Biol 1 (1): 69-89, 2010, H. Liu et al. fig 2B p. 70

Chapter 13

Opener: MAREK MIS/SCIENCE PHOTO LIBRARY; **13.3:** Adapted from Becker et al, The World of the Cell, 2000 Addison Wesley Longman; **13.4:** Adapted from Anderson and Anderson, Trends in Biochemical Sciences Vol 7, 1982, Elsevier Science; **13.5:** Adapted from Nelson, N and Ben-Shem, A. the Complex Architecture of Oxygenic Photosynthesis Nature Reviews: Mol Cel Biol 5:971-82, p. 978 fig 5A; **13.6:** Adapted from A. Melkozernov; **13.8:** Adapted from Trends in Biochemical Sciences, Vol 22, Jung, Hill, Engelbrecht, p. 420-423; **13.16a:** Adapted from Nelson, N. and Ben-Shem, A. The Complex Architecture of Oxygenic Photosynthesis Nature Reviews: Mol Cell Biol 5: 971-82 p. 978 fig 5b; **13.17:** Adapted from Becker et al. "The World of the Cell 4e" 2000 Addison Wesley Longman; **13.18:** Adapted from Becker et al. "The World of the Cell 4e" 2000 Addison Wesley Longman; **13.24:** Adapted from Becker et al, "World of the Cell, 4e 2000, Addison Wesley Longman; **13A:** MacMillan Publishers Ltd: Nature reviews: microbiology, Molloy, S. Down in the Depths Nature Reviews: Microbiology 3:582, 2005. p. 582

Chapter 14

Opener: Dr. Robert Calentine/Getty Images; **14.1a:** NASA; **14.1b:** NASA; **14.2:** Adapted with permission from PC Dossantos et al., "Formation and Insertion of the Nitrogenase Iron-Molybdenum Cofactor" Chemical Review 104: 1159-1173, 2004, fig no. 1, pg no. 1161. Copyright 2005 American Chemical Society

Chapter 15

15.2: Adapted from Richsteiner, M. and Hill, CP Mobilizing the Proteolytic Machine: Cell Biological Roles of Preoteasome Activators & Inhibitors, Trends Cell Biol 15 (1): 27-33 2005 p. 27, 28 Fig 1.3

Chapter 16

Opener: Georgia Bennet/Bennet Images; **16.1:** Adapted from J. Coolman KH Rahn, "Color Atlas of Biochemistry" 1999, Thieme Medical Publishers; **16.9:** Adapted from Geoffrey Cooper, "The Cell: a Molecular Approach", 1997, the Sinauer Associates; **16.11:** Adapted from TM Delving, "Textbook of Biochemistry with Clinicial Correlations" 1999, John Wiley and Sons; 16A: Science/Visuals Unlimited

Chapter 17

Opener: monkeybusinessimages/iStock

17.1: PHOTOGRAPH BY A. BARRINGTON BROWN, COPYRIGHT GONVILLE AND CAIUS COLLEGE, CAMBRIDGE / COLOURED BY SCIENCE PHOTO LIBRARY; **17.9:** KING'S COLLEGE LONDON ARCHIVES / SCIENCE PHOTO LIBRARY; **17.10:** Adapted from Lehninger, Nelson & Cox, "Principles" 2000 Worth Publishers; **17.13:** Adapted from TA Brown, Genomes, 1999, BIOS Scientific Publishers Oxford UK; **17.15:** Adapted from Taylor and Francis; **17.17:** Adapted from Watson et al. Mol Biol. Of the Gene, Pearson/Benj Cummings, p. 154-155, 7.19a&b, 17.20; **17.19:** Adapted from CK Matthews and KE Van Holde Biochemistry, 2nd ed. 1996 Benjamin Cummings; **17.20:** Szalaj, P. & Plewczynski, D. Cell Biol Toxicol (2018) 34: 381. https://doi.org/10.1007/s10565-018-9428-y; **TA 17.5:** Matharu NK., Ahanger SH., "Chromatin Insulators and Topological Domains: Adding New Dimensions to 3D Genome Architecture." US National Library of Medicine. doi: 10.3390/genes6030790; **17.22:** Adapted from TA Brown, Genomes, 1999, BIOS Scientific Publishers Oxford UK; **17.23:** Haggstrom, Mikael. "Medical gallery of Mikael Haggstrom 2014" Wikiversity Journal of Medicine 1(2) DOI: 10.15347/wjm/2014.008 ISSN 20018762 - PANTHER pie chart at the Panther Classification System homepage; **17B:** Adapted from Watson et al. Mol Biol of the Gener. Pearson/Benjamin Cummings fig 7.27 p 160; **17C:** Cooper GM & Housman RE, "The Cell: A Molecular Approach" 7th edition, 2017. With Permission of Sinauer Associates, an Imprint of Oxford University Press; **17D:** Reprinted with permission from Randy L. Jirtle, Duke University - Discover Magazine; **17E:** Adapted from Benjamin Lewin, Oxford University Press and Cell Press, 200. Reprinted from Genes VII; **17F:** Adapted from Prescott et al, Microbiology 4e 1999 by the McGraw Hill Companies; **17G:** Adapted from Recombinant DNA by JD Watson, M. Gilman J. Witkowski, M. Zoller 1992, 1983 by JD Watson, M. Gilman, J. Witkowski, M. Zoller. Used with the permission of WH Freeman and Company; **17H:** Adapted from Prescott et al, Microbiology 4e 1999 by the McGraw Hill Companies; **17I:** Adapted from Watson et al Recombinant DNA 1992, WH Freeman; **17J:** Tek Image/Science Source

Chapter 18

18.3: Adapted from Becker et al. "The World of the Cell 4e" 2000. Addison Wesley Longman; **18.4:** Adapted from TA Steitz,

A Mechanism for All Polymerases Nature 391:231-32, p. 231 fig. 1; **18.6:** Adapted from Oxford University Press and Cell Press 2000. Reprinted from Genes VII by Benjamin Lewin - by permission of OUP; **18.7:** Adapted from Benjamin Lewin, Oxford University Press and Cell Press, 2000. Reprinted from Genes VII.; **18.8:** Adapted from Erzberger, JP, Mott, ML & Berger, JM Structural Basis for ATP dependent AnA Assembly and Replication Origin Remodeling Nature Struc Mol Biol 12(8):676-683. p. 677 fig 1; **18.11:** Adapted from Garrett and Grisham Biochemistry, 3rd ed 2005. Thomson/Nrooks Cole fig 28.10, p. 907; **18.12:** Adapted from TA Brown, Genomes 3, Garland Science New York. Fig 15.22, p. 488; **18.15:** Adapted from TA Brown, Genomes 3, Garland Science p. 483 fig 15.15; **18.17:** Adapted from TA Brown Genomes 3 Garland Science p. 491 fig 15.25; **18.18:** Adapted from TA Brown Genomes 3 Garland Science p. 491 fig 15.25; **18.19:** Adapted from Molecular Biology of the Gene, 5th edition by James D. Watson, et al. Copyright 2004 by Pearson Education; **18.26:** Adapted from A. Landy and RA Weisber, RW Hendrix, Lambda II, Coldspring Harbor Laboratory Press, 1983; **18A:** Adapted from Prescott et al. Microbiology 4e 1999 McGraw Hill; **18B:** Adapted from EW Nester, "Microbiology: a Human Perspective" 3e 2001 McGraw Hill; **18C:** Keith V. Wood; **18D:** Adapted from RH Tamarin, Principles of Genetics 6e 1999 The McGraw Hill Companies; **18E:** Adapted from EW Nester, "Microbiology: a Human Perspective" 3e 2001 McGraw Hill; **18F:** Adapted from Molecular Cell Biology by J Darnell, H. Lodish, D. Balitimore, P. Matsudaira, S Zipursky, and A Berk, copyright 2000, 1995, 1990; **18Ga:** Adapted from RH Tamarin, Principles of Genetics 6e 1999 The McGraw Hill Companies; **18Gb:** AFFYMETRIX Inc.; **18.31:** Cooper GM & Housman RE, "The Cell: A Molecular Approach" 7th edition, 2017. Fig 8.1, p. 260. With Permission of Sinauer Associates, an Imprint of Oxford University Press; **18.32:** Adapted from Lehninger, Nelson and Cox "Principles…." 200 Worth Press; **18.34:** Adapted from Korzheva, N. et al. A Structural Model of Transcription Elongation Science 289: 619-625, fig 3 p. 622; **18.35:** Adapted from TA Brown, Genomes 3 2007 Garland Science p. 337 fig 12.5; **18.36:** Adapted from TA Brown, Genomes 3 2007 Garland Science p. 337 fig 12.5; **18.37:** Adapted from G. Karp, Cell and Molecular Biology Concepts and Experiments, 4th edition John Wiley 2005 fig 12.47 p. 534; **18.40:** DR MARK J. WINTER / SCIENCE PHOTO LIBRARY; **18.41:** Adapted from Chadick, JZ and Asturias, FS Structure of Eukaryotic Mediator Complexes, Trends in Biochemical Sciences 30(4):264-271, 2005. p. 267 fig 3; **18.43:** Adapted from Cooper, GM, and Hausman RE The Cell: A Molecular Approach, 6e. 2013 Sinauer Associates; **18.47:** Cooper GM & Housman RE, "The Cell: A Molecular Approach" 7th edition, 2017. Page 302. With Permission of Sinauer Associates, an Imprint of Oxford University Press; **18.48:** Adapted from Azubel M. et al. Native Spliceomes Assembled with Pre-mRNA to Form Supraclisesomes J. Mol. Biol. 356:955-66, 2006 Fig 7b p. 963; **18H:** Lino, Christopher, et al. "Delivering CRISPR: a review of the challenges and approaches." US National Library of Medicine. DOI: 10.1080/10717544.2018.1474964; **18.51:** Adapted from Jeffrey Baereck & Ronald R. Breaker, The Power of Riboswitches, Scientific American 296 (1): 50-57, 2007 p. 55; **18.52:** Adapted from Terence A Brown, Genomes 1999. Reprinted with permission of BIOS Scientific Publishers, Oxford UK; **18.53:** Adapted from Alberts et al. Mol Biol of the Cell, 4th ed. Garland p. 319, fig 6.27, Routledge Taylor and Francis

Chapter 19

Opener: Adapted from Harry Noller, University of California, Santa Cruz - via Phyllis Tveit Centre for Molecular Biology of RNA; **19.1:** Adapted from Lodish et al. "Molecular Cell Biology" 4e 2000; **19.6:** Adapted from Molecular Cell Biology by J. Darnell H. Lodish, D. Baltimore, P. Matsudaira, S. Zipursky, and A. Berk. Copyright 2000, 1990, 1986 by Scientific America Books. Used with permission of WH Freeman and Company; **19.13:** Adapted from L.I. Slbin "Polypeptide Chain Elongation" in Translation in Eukaryotes, edited by H. Trachsel, 1991. Copyright CRC Press, Boca Raton, Florida; **19.19:** Adapted from "The Eorld of the Cell" Becker et al. 2000, Addison Wesley and Longman; **19.23:** Adapted from Kim, YE et al., Molecular Chaperone Functions in Protein Folding and Protestasis, Annual Review of Biochemistry, 82: 323-355 (2013)

Index

Page numbers followed by *f* and *t* refer to figures and tables, respectively.

A

ABC transporters, 431
Abetalipoproteinemia, 444
ABO blood group, 259, 269
Acanthocytes, 444
Acetals, 255, 255*f*
Acetic acid:
  dissociation constants and pK_a values for, 94*t*
  oxidation-reduction reactions and, 18
  titration of with NaOH, 96*f*
Acetoacetyl-ACP, 465, 466*f*
Acetyl-CoA, 339, 344
  amino acids forming, 582–85
  biosynthesis of ketone bodies from, 456, 457*f*
  lipid metabolism and, 445, 454, 456, 458–59
  pyruvate conversion to, 341–42
Acetyl-CoA carboxylase (ACC), 461–64, 464*f*
N-Acetylglutamate (NAG), 540, 580
Acid-base catalysis, 207–9
Acid-base equilibria, P-17–18
Acid-base reactions, P-19
Acid hydrolases, 52
Acidic amino acids, 137*f*, 139
Acidosis, 94
Acids:
  amino acids as, 136
  general, 208*f*, 209
  pH and, 93–94
  strong, 168
  structure and significance of, 6*t*
  weak (*see* Weak acids)
Acquired immune deficiency syndrome (AIDS), 226
Actin-binding proteins, 87
Actin filaments, 426
Activation energy, 198
Activation theory, P-14
Active sites, 199
Active transport, 428
Activity coefficient, 200
Activity series of metals, P-19
Acylated proteins, 406
Acyl carrier protein (ACP), 461, 463*f*

Acyl-CoA, 452–54, 453*f*
Acyl-CoA:cholesterol acyltransferase (ACAT), 418
Acyl-CoA dehydrogenase, 453
Acyl groups, 406
Addition reactions, 17, 17*f*, 148
Adducts, 256
Adenine, 11
  structure of, 11*f*
  tautomers of, 556*f*
Adenine-thymine (AT) pair, 638, 640, 640*f*, 643, 644–45
Adenosine deaminase, 576
Adenosine deaminase deficiency, 592
Adenosine triphosphate. *See* ATP
S-Adenosylmethionine (SAM), 138, 213, 214*t*, 472, 550, P-32
  formation of, 547*f*
  one-carbon metabolism and, 543, 546–48
  pathway of, 548*f*
  phospholipid metabolism and, 472
S-Adenosylhomocysteine, 547, 548f
Adenylate kinase, 154*f*
Adipic acid, 460*f*
Adipocytes, 408, 446–51
Adipose tissue, brown. *See* Brown adipose tissue
A-DNA, 646, 647*f*, 647*t*
ADP-ATP translocator, 378, 379*f*
Advanced glycation end products (AGEs), 256–57, 473
Adventitia, 473
Aerobic metabolism, 58, 329–61, 362–400. *See also* aerobic respiration; Citric acid cycle; Electron transport; Oxidation-reduction reactions; Oxidative phosphorylation
  in mitochondria, 332*f*
  overview of, 331*f*
  process of, 339
Aerobic respiration, 283, 330, 363
Aerotolerant anaerobes, 331
Affinity chromatography, 185–86
AGEs. *See* Advanced glycation end products
Aggrecan, 267
Agouti mice, 664, 664*f*
Agre, Peter, 431
AIDS. *See* Acquired immune deficiency syndrome
Akt. *See* Protein kinase B
Alanine, 138, 532*t*, 542
  abbreviations, 138*t*
  β-, 7, 8*f*
  catabolic pathway of, 582, 582*f*
  enantiomers, 141*f*
  ionization state, 137*f*
  pK_a values, 142*t*
Albinism, 589
Alcohol dehydrogenases (ADHs):
  catalysis and, 216, 217, 219*f*
  functional groups of the active site, 219*f*
  human, 218*f*, 219*f*, 237–38
  overview of, 135*f*
  shape of, 135*f*
  yeast, 237

Alcohols:
properties of, P-27
structure and significance of, 6*t*
Aldaric acid, 252
Aldehydes:
properties of, P-27–28
structure and significance of, 6*t*
Aldimines, 148
Alditols, 253
Aldol cleavage, 287
Aldol condensation, 175
Aldonic acid, 252
Aldose dehydrogenase, 615
Aldoses, 202, 247, 249–50
defined, 8
D family of, 249*f*
isomerization reactions, 17, 18*f*
structure of, 247*f*
Aldosterone, 419*f*
Aliphatic elimination reactions, P-32–34
Aliphatic hydrocarbons, 138
Aliphatic substitution reactions, P-31–32
Alkali metals, 211, P-7
Alkaline metals, 211, P-7
Alkalosis, 94–95
Alkanes, P-23, P-25
Alkaptonuria, 589
Alkenes, 6*t*, P-24, P-25, P-25*f*
Alkylating agents, 644–45
Alkyl groups, P-23
Alkynes, P-24
Allopurinol, 590, 600
Allosteric enzymes, 231, 231*f*, 233–35
Allosteric regulation:
of gluconeogenesis, 308, 308*f*
of glycolysis, 298, 299*t*, 308*f*
Allosteric sites, 234
Allosteric transitions, 164, 179*f*
Allostery, 163, 178
Allyl groups, 477
Alpha-alpha units (supersecondary structure), 158
Alpha-carbonic anhydrases (CAs), 212
Alpha-helices, 156–57, 156*f*, 157–58, 172
Alpha-ketoglutarate, 536
amino acids forming, 585–86
oxidation of, 346
Alpha-ketoglutarate dehydrogenase, 351, 353
Alpha-ketoglutarate/glutamate pair, 533
Alpha-linolenic acid, 404, 404*t*, 405*f*
Alpha-oxidation, 459, 460*f*
Alpha-tocopherol. *See* Vitamin E
ALS. *See* Amytrophic lateral sclerosis
Alternative splicing, 271, 658, 736, 737*f*
Alu element, 660
Alzheimer’s disease, 40, 172, 384, 436–37, 570, 786
Amadori product, 256

Amethopterin, 548
Amides:
properties of, P-30
structure and significance of, 6*t*
Amido groups, 6*t*
Amines:
biogenic, 551
neurotransmitter, 551*t*
properties of, P-30
structure and significance of, 6*t*
Amino acid carbon skeletons:
catabolism of, 581–88
defined, 532
Amino acid composition, 136
L-Amino acid oxidases, 576
Amino acid pool, 533
Amino acid residues, 136, 146
Amino acids, 136–48
acidic, 137*f*, 139
basic, 137*f*, 139
biologically active, 140
biosynthesis of, 532–43, 537*f*
biosynthetic reactions involving, 543–64
branched-chain, 532–33, 537
catabolism of, 575–89, 576*t*
catalysis and, 209–11
classes, 136–39
classification as α, β, or γ, 7
deamination of, 575–76
essential, 532, 532*t*
general formula for α-, 7*f*
general function of, 6*t*
general structure of α-, 138*f*
glucogenic, 575
ionizable group of, 142–43
ketogenic, 575
metabolism of, 532–33
modified, in proteins, 140–41, 141*f*
nonessential, 532, 532*t*
nonpolar, 136, 137*f*, 138
nonstandard, 136
polar, 137*f*, 138–39
properties and functions of, 7
reactions, 145–48
semi-essential, 532*t*
standard, 136, 137*f*, 138*t*
stereoisomers, 141
transport of, 533
Amino acid sequences, 146
Aminoacyl-tRNA synthetase, 665, 757–59, 758*f*, 762, 764
p-Aminobenzoic acid, 213
Amino groups, 5, 136
reactions, 533–37
significance of, 6*t*
Aminopropionitrile, 176
Amino sugars, 259, 259*f*

Aminotransferases, 533
Ammonium ions, 536–37
Ammonotelic organisms, 571
Amoeboid motion, 87, 87*f*
AMP, 557, 558*f*, 559
AMP-activated protein kinase. *See* AMPK
Amphibolic pathways, 291, 349, 350*f*
Amphipathic molecules, 88, 168
Amphisomes, 575
Amphoteric molecules, 136
AMPK:
  carbohydrate metabolism and, 470*f*
  fatty acid metabolism and, 470
  feeding behavior and, 625
  glycolysis regulation and, 300–301
  lipid metabolism and, 470*f*
AMY1 gene, 634–35
Amylopectin, 262, 263*f*
Amylose, 262, 262*f*
Amytal, 373*f*
Amytrophic lateral sclerosis (ALS; Lou Gehrig's disease), 384, 389
Anabaena azollae, 528
Anandamine, 406
Anabolic pathways, 20, 20*f*
Anaerobic organisms, 283–84
Anaerobic pathways, 237
Anaplerotic reactions, 349
Anemia, 168
  lead poisoning and, 230
  pernicious, 545
  sickle-cell, 155–56, 155*f*, 432
Anfinsen, Christian, 167, 168
Anfinsen experiment, 167, 167*f*, 168
Angular momentum quantum number l, P-3–4
Anhydrides, 14, 758
Anions, P-2, P-9
Ankyrin, 426
Annotation, 717
Anomeric carbon atoms, 249
Anomers, 249–50, 251*f*
Antenna pigments, 497
Antibiotic resistance, 709, 710, 713, 752
Antibiotics, 581, 767*t*
Antibodies, 135
Anticodon loop, 665
Anticodons, 756–57, 756*f*
Anti conformations, 555
Antimycin, 373*f*
Antioxidants, 309, 384, 393*f*, 419
  in diet, 392–94
  enzyme systems, 389–91
Antiparallel β-pleated sheets, 157, 158, 172
Antiporters, 434
Antisense strand, 675
AP-1, 741
Apoenzymes, 200

Apolipoprotein E4, 436–37
Apoproteins (apolipoproteins), 153, 419
Apoptosis, 49, 58, 59*f*, 740
Aptamer, 735
Apurinic sites, 700
Apyrimidinic sites, 700
Aquaporin 1 (AQP1), 430, 431*f*
Aquaporins, 425, 430–31
Aqueous solutions, 410, 410*f*
Arabinose, 250*f*
Arachidic acid, 404*t*
Arachidonic acid, 404*t*
Archaea, 34, 36
Arcuate nucleus (ARC), 624–25, 625*f*
Aretaeus, 614
Arginine, 140, 532*t*, 537, 540
  abbreviations, 138*t*
  biosynthesis from glutamate, 538*f*
  catabolic pathway of, 585–86, 585*f*
  ionization state, 137*f*
  pK_a values, 142*t*
Argonaute, 738
Aromatic Amino Acid family, 536*f*, 542–43
Aromatic hydrocarbons, 138, P-23, P-25–26
Artemis, 703
Arthritis, 257, 407, 548
Artificial sweeteners, 35
Ascorbate, 393, 393*f*
Ascorbic acid (vitamin C), 392, 778
  coenzyme forms and process promoted, 214*t*
  functions of, 253, 392
  structure of, 253*f*, 393*f*
Asparagine, 138–39, 532*t*, 542
  abbreviations, 138*t*
  ionization state, 137*f*
  pK_a values, 142*t*
Asparagine-linked oligosaccharides, 261, 268–69
Aspartame, 204
Aspartate (aspartic acid), 139, 532*t*, 542
  abbreviations, 138*t*
  ionization state, 137*f*
  pK_a values, 142*t*
Aspartate family, 536*f*, 541–42
Aspartate-malate shuttle, 381*f*
Aspartate transcarbamoylase, 135*f*
Aspartic acid. *See* Aspartate
Aspergillus nidulans, 705
Aspirin, 196–97, 407
ASPM gene, 634
Asymmetric carbons, 141
Ataxia telangiectasia, 700
Ataxia telangiectasia and Rad3-related protein (ATR), 700, 703
Ataxia telangiectasia mutated protein (ATM), 700, 703, 709
Atherosclerosis, 257, 384, 395, 422, 473
ATM. *See* Ataxia telangiectasia mutated protein

Atomic number, P-2–3
Atomic radius, P-6
Atomic structure, P-2–8
Atomic theory, P-3–4
ATP:
functions of, 6–7
glucose oxidation and, 380–83, 381*t*
glucose reaction with, 16*f*
hydrolysis of, 14–15, 16*f*, 122, 124–27, 124*f*, 171–72
oxidative phosphorylation and (*see* Oxidative phosphorylation)
role of in energy, 124–27, 125*f*
structure of, 11*f*, 124, 125*f*
synthesis of, 374, 376–78, 378*f*
ATP synthase, 374
chloroplast, 498–99, 499*f*
E. coli, 376, 377*f*
role of in ATP synthesis, 376–78
structure of, 362*f*, 376, 376*f*
ATR. *See* Ataxia telangiectasia and Rad3-related protein
Atrial natriuretic factor (ANF), 610–11
amino acid sequence of, 148*t*
functions of, 150
Attachment (att) sites, 709
Aufbau principle, P-4
Autoantibodies, 407
Autocrine regulators, 406
Autocrine signaling, 603
Autoimmune diseases, 407, 548
Autophagosomes, 574–75
Autophagy, 40, 49, 570, 572, 574
Autophagy-lysosomal system, 573*f*, 574–75
Autopoiesis, 13
Autoradiography, 69
Autotrophs, 19
Average atomic mass unit, P-3
Avogadro's number, P-21
Axoneme, 64, 64*f*
Aztobacter vinelandii, 528

B

Bacteria:
beneficial and harmful, 36
cyano-, 526*f*, 528
Gram-negative and -positive, 42
structure of typical cell, 42*f*
symbiotic, 528
Bacterial artificial chromosomes (BAC), 713
Bacterial conjugation, 709
Bacterioides, 34*f*
Bacteriophage T4, 676–77, 677*f*
Bacteriophage λ, 709, 709*f*, 712–13
Bacteroidetes, 35
Ball-and-stick model, 153
Band 4.1, 426
Band 3 anion exchanger macrocomplex, 426–27
Banting, Frederick, 615

Bardet–Biedl syndrome (BBS), 67
Basal ganglia, 634
Base analogues, 644
Base excision repair, 700, 701*f*
Bases:
amino acids as, 136
conjugate, 93, P-19
general, 208*f*, 209
nitrogenous, 10, 11*f*
pH and, 93–94
Schiff, 148, 256, 534, 534*f*, 535*f*
strong, 102*f*, 168
weak, 93, P-17–18
Base stacking, 640
Basic amino acids, 137*f*, 139
Basic Local Alignment Search Tool (BLAST), 188
B blood group, 272
B cells (B lymphocytes), 592
B-DNA, 638, 641, 646, 647, 647*f*, 647*t*
Benard cells, 123
Benedict's reagent, 253, 253*f*, 255
Benzene, P-26, P-26*f*
Best, Charles, 615
Beta-alpha-beta units (supersecondary structure), 158, 158*f*
Beta-aminopropionitrile, 176
Beta-Arrestins, 610
Beta-barrel motif (supersecondary structure), 158, 158*f*, 159*f*
Beta-carotene, 392, 393*f*, 416*t*, 494, 506
Beta-meander motif, 158, 158*f*
Beta-oxidation, 451–61, 467, 470
acyl-CoA, 453*f*
fatty acid, synthesis compared with, 468*t*
oleoyl-CoA, 458*f*
in peroxisomes, 456
Beta-oxidation spiral, 454
Beta-pleated sheets, 157–58, 157*f*, 172
Beta-strands, 157–58
Beta-turn, 158
BH_4. *See* Tetrahydrobiopterin
BHT, 394
Bicarbonate, 94*t*
Bicarbonate buffer, 101–2
Bifidobacterium, 34*f*
Bile, 480–81
Bile salts, 418, 445, 481, 481*f*
Bimolecular reactions, 220
Biochemical pathways, 3, 20*f*
energy transfer, 20
metabolic, 20
signal transduction, 21
Biochemical reactions, 13–19
elimination, 15, 17*f*, P-32–34
isomerization, 17, 18*f*
nucleophilic substitution, 14–15, 16*f*
oxidation-reduction (*see* Oxidation-reduction reactions)
Bioenergetics, 113

Biofilms, 43
Biofuels, 492
Biogenesis, 58
Biogenic amines, 551
Biogeochemical cycles, 111–12
Bioinformatics, 26, 717–18
Biological membranes. *See* Membranes
Biomimetics:
  photosynthesis and, 520–21
  spider silk and, 181–83
Biomolecules, 5–13
  defined, 3
  functional groups of, 5, 6*t*
  major classes of, 6–13, 6*t*
  synthesis of, 21
Biopterin, 213, 214*t*
Bioremediation, 111
Biosphere, 111
Biotin, 213, 214*t*, 302
Biotransformation reactions, 49, 485–86
2,3-Bisphosphoglycerate (BPG), 180*f*
Bisphosphonates, 484
Bisubstrate reactions, 225
BLAST, 188
Blobel, Gunter, 780
Blood clotting, 25
Blood pressure, 149–50
Bohr effect, 179, 427
Botulinum toxin, 152
Botulism, 415
Brain insulin resistance, 617
Branched-chain amino acids (BCAA), 532–33, 537
Branched-chain hydrocarbons, P-23
Branched-chain ketoaciduria, 589
BRCA1, 709, 743
BRCA2, 709, 743
Broca's area, 634
Bronsted–Lowry theory, P-19
Brown adipose tissue, 364, 384
Brown, Michael, 421, 435
Buffering capacity, 96
Buffers, 94–102
  bicarbonate, 101–2
  function of, 95
  Henderson–Hasselbalch equation and, 96–97
  phosphate, 102
  physiological, 101–2
  protein, 102
Butylated hydroxytoluene (BHT), 394

C

Cadherins, 269
Caenorhabditis elegans, 717, 738
Cajal bodies, 58
Calcium:
  citric acid cycle regulation and, 353

phosphatidylinositol cycle and, 609
Calcium ions, 41
Calcium signaling, 60
Calmodulin, 135*f*
Calnexin-calreticulin cycle, 777
Calvin cycle, 493, 509, 510–13, 517
defined, 510
reduction phase, 512
regeneration phase, 512–13
schematic of, 511*f*
cAMP, 317, 317f, 448, 450f, 451, 607–8, 612
cAMP-dependent protein kinase (PKA), 607–8
Cancer, 25, 26, 66, 738
amethopterin for, 548
carcinogenesis, 356–57, 743–45
citric acid cycle and, 353
glycolysis and, 356–57
LncRNAs and, 744–45
mutations and, 645, 664, 703
obesity and, 402
oxidative damage and, 364, 384
p53 and, 166, 356–57
proteostasis network and, 786
telomerase and, 699
tumor promoters and, 614
uncoupling proteins and, 384
Cap-binding complex (CBC), 768–69
Capping of mRNA, 727
Capsid, 676
Capture silk, 134
Carbamoylated amino groups, 179
Carbanions, 205, P-33–34
Carbocation, 205, P-25
Carbohydrate metabolism, 280–328, 284*f*, 470*f*. *See also* Gluconeogenesis; Glycogen metabolism; Glycolysis; Pentose phosphate pathway
Carbohydrates, 245–79. *See also* Disaccharides; Glycoconjugates; Monosaccharides; Polysaccharides
general function of, 6*t*
properties and types of, 8–9
Carbomylation, 518
Carbon atoms:
anomeric, 249
in citric acid cycle, 348–49
Carbon cycle, 111–12
Carbon dioxide (CO_2):
formation via alpha-ketoglutarate oxidation, 346
formation via isocitrate oxidation, 345
Carbon dioxide compensation point, 515
Carbon fixation, 510–12
Carbohydrate response element binding protein (ChREBP), 300, 315, 471
Carbonic acid, 94*t*
Carbonic anhydrase, 211, 427
Carbon monoxide (CO), 180, 552–54
Carbon skeletons. *See* Amino acid carbon skeletons
Carbonyl groups, 6*t*, P-27
2-Carboxyarabinitol-1-phosphate, 519, 519*f*
Carboxylate groups, 136

Carboxylation, 779
Carboxyl groups, 5, 6*t*
Carboxylic acids, P-28
Carcinogenesis, 356–57, 743–45
Carcinomas, 743
Cardiac glycosides, 246, 418, 420*f*
Cardiolipin, 410
Cardiovascular disease, 40, 402, 664. *See also* Heart disease
Carnitine, 213, 452, 452*f*
Carnitine acyltransferase, 452, 456, 464
Carotenoids, 392, 416–17, 494
Carrier proteins, 37
Carriers, 429
Catabolic pathways, 20, 20*f*, 281
Catabolism:
  amino acid, 575–89
  nucleotide, 591*f*
  purine, 590–92
  pyrimidine, 592
  uric acid, 590, 592, 592*f*
Catabolite gene activator protein (CAP), 733, 734
Catalase, 135*f*, 390
Catalysis, 151, 204–18
  acid-base, 207–9
  amino acid roles in, 209–11
  cofactors in, 211–13
  covalent, 209
  detailed mechanisms of, 216–18
  electrostatic effects and, 207
  mechanisms of, 207–9
  organic reactions and the transition state, 204–5
  orientation effects and, 207
  pH effects on, 214–15, 215*f*
  proximity effects and, 207
  temperature effects on, 214, 215*f*
  transition state stabilization, 205–6
Catalysts, 3, 198, 198*f*
  defined, 198, P-14
  reaction rate and, P-16
Catalytic perfection, 223
Catalytic residues, 210
Catecholamines, 542
Catechol-*O*-methyltransferase (COMT), P-33*f*
Cations, P-2, P-9
Caveolar endocytosis, 52, 54
CBS deficiency, 786
CCAAT box, 724
CCCTC-binding factor (CTCF), 651, 653, 654, 655f
cDNA libraries, 715
Celiac disease, 545
Cell adhesion molecules (CAMs), 269
Cell fractionation, 68, 68*f*
Cell membranes. *See* Membranes
Cellobiose, 260–61, 261*f*
Cell cycle, 696
Cells, 33–75

animal, structure of, 46*f*
as basic unit of life, 3
biochemical reactions of (*see* Biochemical reactions)
biological order in, 21–22
energy and (*see* Energy)
eukaryotic (*see* Eukaryotes)
hormones and communication in, 603–17
macromolecular crowding in (*see* Macromolecular crowding)
metabolism in (*see* Metabolism)
movement of, 21
oxidative damage to, 363, 384
oxygen and function of, 384–94
plant, structure of, 46*f*
prokaryotic (*see* Prokaryotes)
proteostasis in, 39–40
sugars and, 245*f*
technology used in research, 68–70
as thermodynamic systems, 117
water importance in, 36–37
Cell-surface receptors, 605–10
Cellular immunity, 590
Cellulose, 8, 261
disaccharide repeating unit of, 264*f*
properties and structure of, 264
Central dogma, 638, 677
Centromeres, 659
Ceramide, 413, 413*f*
Cerebrosides, 414
cGMP, 611
Chain-terminating method (DNA sequencing), 669, 669*f*
Channel proteins, 37
Channels, 428
Chaperone-mediated autophagy, 574
Chaperonins, 171–72, 172*f*
Chargaff, Erwin, 646
Chargaff's rules, 646, 665
Chemical bonding, P-9–14
Chemical reactions, P-14–22
equilibrium constants, P-16–17
measuring, P-21–22
reaction kinetics, P-15–16
reaction types, P-18–20
Chemiosmosis, 364, 374
Chemiosmotic coupling theory, 374–76, 375*f*
Chemoautotrophs, 19
Chemoheterotrophs, 19
Chemolithotrophy, 113
Chemoorganotrophy, 113
Chemosynthesis, 19
Chemotaxis, 70, 766
Chimeras, 687
Chimpanzee genome, 634
Chiral carbons, 141
Chitin, 8–9, 261
Chloride shift, 426–27
Chlorophyll, 494–99, 501

Chloroplasts, 494–99, 507
functions of, 62
membrane organization of light reactions in, 508*f*
structure of, 62, 62*f*, 494–96, 495*f*
Cholecystitis, 482
Cholera, 415, 433
Cholera toxin, 273, 614
Cholesterol:
degradation of, 480–81
heart disease and, 484
homeostasis, 482–84
properties and functions of, 417–18
structure of, 418*f*
synthesis of, 476–80, 480*f*
Cholesterol biosynthetic pathway, 484
Cholesterol metabolism, 475–76
cholesterol degradation, 480–81
cholesterol homeostasis, 482–84
cholesterol synthesis, 476–80, 480*f*
Cholic acid, 419*f*
Cholecystokinin (CCK), 619, 624f, 625
Chorismate, 543, 543*f*
Chromatid, 709
Chromatin, 736, 744
defined, 651
electron micrograph of, 651*f*
structure of, 652–55, 654*f*
Chromatin fibers, 55
Chromatin-remodeling complexes, 723, 723*f*
Chromatography, 184–86
Chromium, 301
Chromophores, 501–2, 502*f*
Chromoplasts, 62
Chromosomal jumping, 715
Chromosomes, 44, 635, 650–56
bacterial artificial, 713
defined, 650
eukaryote, 651
prokaryote, 650
telomerase-catalyzed extension of, 699*f*
yeast artificial, 713
Chromosome territories, 55, 652
Chromosome walking, 715, 716*f*
Chylomicron remnants, 446, 619
Chylomicrons, 421, 445–46
Chymotrypsin:
activation of, 233*f*
catalysis and, 210–11, 214, 216–18, 217*f*, 233*f*
Cilia, 63–64, 67
Ciliopathies, 67
Circadian rhythms, 549
Cis-isomers, 404
Cisternal space, 48
Cistrons, 673
Citrate:
isomerization to form a secondary alcohol, 344–45

metabolism of, 352–53, 352*f*
synthesis of, 345*f*
Citrate synthase, 351–53
Citric acid cycle, 291, 298, 339–55, 454, 534
amphibolic, 349, 350*f*
defined, 331
disease and, 353
fate of carbon atoms in, 348–49
the glyoxylate cycle and, 354–55, 354*f*, 355*f*
pyruvate-to-acetyl-CoA conversion, 341–42
reactions of, 344–48
regulation of, 351–53, 351*f*
schematic of, 340*f*
Citrulline, 140, 140*f*
Clamp loader (DNA synthesis), 690, 692–93*f*
Clathrate, 88
Clathrin, 50, 52, 434
Clathrin-dependent (receptor-mediated) endocytosis, 52, 53*f*, 54*f*
Clathrin-independent caveolar endocytosis, 52
Cleland, W. W., 225
Client proteins, 171
Climate change, 492
Cloning:
molecular, 712–14, 713*f*
shotgun, 715, 715*f*
Closed thermodynamic systems, 114
Clostridium, 34*f*
Clostridium acetobutylicum, 237, 293
Clostridium botulinum, 152
Clostridium pasteurianum, 528
Clostridium perfringins, 413
C4 metabolism, 515–17, 516*f*
Coated pits, 434
Co-chaperones, 171
Cockayne syndrome, 701
Coding strand of DNA, 718, 719*f*
Codon-anticodon interactions, 756–57, 756*f*
Codons, 673
initiating, 754
start, 718, 754
stop, 718, 732, 754, 755
Coenzyme A (CoASH), 14, 213
in citric acid cycle, 339, 341*f*, 342
in lipid metabolism, 449f, 452, 453, 454, 457f, 461, 462f
Coenzyme Q (ubiquinone), 213
electron transport and, 364, 367, 368*f*, 369, 373
process promoted by, 214*t*
statin interference with, 484
structure and oxidation states of, 365*f*
Coenzymes, 19, 200, 213
in citric acid cycle, 341*t*
redox, 335–36
Cofactors:
catalysis and, 211–13
types of, 200
Cohesin, 654, 725

Collagen:
fibrils, 177*f*
functions and structure of, 174–75
lathyrism and, 176
as a shock absorber, 88
Collision theory of chemical reactions, P-14
Colonization resistance, 35
Colony hybridization technique, 714
Combustion, 116–17, 118*f*, 363, P-20, P-23
Compartmentation (cellular), 235–36
Competitive inhibitors, 226–29, 229*f*
Complex I. *See* NADH dehydrogenase complex
Complex II. *See* Succinate dehydrogenase complex
Complex III. *See* Cytochrome bc_1 complex
Complex IV. *See* Cytochrome oxidase
Composite transposons, 709
Concerted model (allosteric regulation), 234, 234*f*, 235
Conformational models:
of cholesterol, 418*f*
of fatty acids, 405*f*
of monosaccharides, 251, 251*f*
of sphingomyelin, 414*f*
of triacylglycerols, 408*f*
Conjugate acid-base pair, P-19
Conjugate bases, 93, P-19
Conjugated proteins, 153
Conjugate redox pair, 333
Conjugation, 44, 708, 709
Conjugation reactions, 481
Consensus sequence, 720
Conservation of energy, 114, 232. *See also* First law of thermodynamics
Conservation of matter, P-15
Conservative amino acid substitutions, 154, 155
Constitutive exocytosis, 51
Constitutive genes, 732
Constitutive heterochromatin, 55, 653
Context-dependent coding reassignment, 775
Cooperative binding, 178
Cooperativity
positive, 234
negative, 234, 235
Coordinate covalent bonds, P-9
COPI (coat protein complex I), 49
COPII (coat protein complex II), 49
Copper deficiency, 212
Copy number variations (CNVs), DNA, 634
Core histones, 652, 653*f*
Core promoter elements (CPEs), 724
Core promoters, 724–25, 724*f*
Core proteins of proteoglycans, 266
Cori cycle, 306, 307*f*
Cori's disease, 323
Cortisol, 419*f*
Corynebacterium, 34*f*
Corynebacterium diphtheriae, 774
Cosmids, 713

Cotranslational transfer, 780, 782*f*
Coupled reactions, 120, 121*f*
Covalent bonds, 80, 81*t*, 160–61, P-9
Covalent catalysis, 209
Covalent modification of enzymes, 233
COX Enzymes, 197
CpG, 662, 664
CpG islands, 662
C3 plants, 510, 515
C4 plants, 515–17
Crabtree effect, 296–97
Creatine, 378–79, 549
Creatine kinase (CK), 378–80
Creatine Kinase/Phosphocreatine (CK/PCr) shuttle system, 379–80, 380*f*
CREB (cAMP response element binding protein), 165, 607
Crenation, 90
Crick, Francis, 635–36, 635*f*, 645–46, 754, 756
CRISPR, 731–32, 731*f*
Cristae, 59
Crohn's disease, 548
Crossing over (DNA recombination), 708
Croton tiglium, 614
Cruciforms, 648
C-terminal residue (polypeptides), 146
Cyanobacteria, 526*f*, 528
Cyanocobalamin, 547*f*
Cyanogenic glycosides, 246
Cyclic AMP (cAMP), 317, 317f, 448, 450f, 451, 607-08,
Cyclic electron transport (photosynthesis), 507, 508*f*
Cyclic GMP (cGMP), 610–11
Cyclic hydrocarbons, P-23, P-25
Cyclin-dependent protein kinases (Cdks), 740
Cyclin destruction box, 572
Cyclins, 571
Cyclooxygenase-2 (COX-2), 195, 195f
γ-Cystathionase (CSE), 540, 541f, 553f, 554
Cystathonine β-synthase (CBS), 540, 541f, 553f, 554
Cysteine, 139, 532*t*, 540
  abbreviations, 138*t*
  biosynthesis of, 541*f*
  catabolic pathway of, 582*f*, 583
  ionization state, 137*f*
  oxidation of, 147, 147*f*
  pK_a values, 142*t*
Cysteine proteome, 391
Cystic fibrosis, 26, 429, 431–33, 786
Cystic fibrosis transmembrane conductance regulator (CFTR), 431–33
Cystine, 147*f*
Cystinuria, 148
Cytochrome bc_1 complex (complex III), 368–69, 368*f*, 370*f*, 372–73, 372*t*, 385
Cytochrome b_6f complex, 493, 498, 507
Cytochrome c, 155
  electron transport and, 369, 372, 372*t*
  shape of, 135*f*
  structure of, 369*f*

Cytochrome oxidase (complex IV), 370–73, 371*f*, 372*t*
Cytochrome P_{450}, 451
Cytochrome P_{450} electron transport systems, 485–86, 486*f*
Cytochromes, 368–69
Cytokines, 40, 603
Cytoplasm, 44, 44*f*
Cytosine, 11–12
  methylation of, 662*f*
  structure of, 11*f*
Cytoskeleton, 21, 45, 62–66
  functions made possible by, 65–66
  major components of, 63*f*
  reorganization model, 66*f*

D

DAG (diacylglycerol), 609
Dalton, John, P-3
Daltons (atomic mass units), P-3
Dark reactions, photosynthesis, 493, 510
D-block elements, P-7–8
DCMU, 510
Deamination, 575–76
Debranching enzyme, 317, 320*f*, 323
Decarboxylation, 293
Decoding center, 762
Decomposition reactions, P-19
Deforestation, 77–78
Degeneracy, 23, 754–55
Degradation:
  cholesterol, 480–81
  fatty acid, 451–54
  glycogen, 319*f*–21*f*
  isoleucine, 588*f*
  nucleotide, 590–92
  pyrimidine, 593*f*
  valine, 588*f*
  of α-amino acids found in proteins, 581*f*
Degrons, 572
Delayed response genes, 741
Denaturation of DNA, 666, 667*f*
Denaturation of proteins:
  conditions for, 168
  protection of proteins by carbohydrates, 269
De novo amino acid biosynthesis pathways, 536
De novo cholesterol synthesis, 475
De novo fatty acid synthesis, 475
De novo methyltransferases, 662
De novo peroxisome biogenesis, 61–62
De novo purine synthesis, 555, 557
Density-gradient centrifugation, 69, 69*f*, 666
Deoxyribonuclease. *See* DNase
Deoxyribonucleic acid. *See* DNA
Deoxyribonucleotides, 560–63, 563*f*
Deoxyribose, 9*f*, 10, 259–60, 554–55
Deoxy sugars, 259–60, 259*f*
Depolarization, 428

Desensitization (target cell), 605
Desmosine, 163, 163*f*
Detergents, 168
Detoxification, 485
Dextrorotary isomers, 141
DHAP. *See* Dihydroxyacetone phosphate
Diabetes insipidis, 430–31
Diabetes mellitus, 257, 402, 456, 605, 614–17
  long-term complications of, 617
  mechanisms of damage caused by, 615
  type 1, 615
  type 2, 40, 615–17, 786
  type 3, 617
Diabetic neuropathy, 617
Diacylglycerol (DAG), 609
Dialysis, 184, 184*f*
Dialyzing membrane, 90
Diamagnetic atoms, P-4
Diarrhea, 102, 260, 433, 611
Diastereomers, 248
Dicer, 738
Dielectric constant, 86
Diet:
  amino acids in, 532
  antioxidants in, 392–94
  hypertension-uric acid connection, 600
  low-fat, 402
Dietary fiber, 34
Differential centrifugation, 68
Diffusion control limit, 223
Digestion, 413, 445–46
Digitalis, 418
Digitoxin, 246, 418
Dihydrogen phosphate:
  dissociation constants and pK_a values for, 94*t*
  titration by a strong base, 102*f*
Dihydroxyacetone, 247, 247*f*
Dihydroxyacetone phosphate (DHAP), 17, 18*f*, 287–88
3,4-Dihydroxyphenylalanine. *See* L-DOPA
Dimers, protein subunits, 162
Dinitrophenol (DNP), 374, 375*f*, 376
Dipalmitoylphosphatidylcholine, 412
Diphosphatidylglycerol, 411*t*
Diphtheria, 774
Dipole-dipole interactions, 82, 82*f*
Dipole-induced dipole interactions, 82, 82*f*
Dipoles, 79, 80*f*, P-12
Diradical molecules, 363
Direct repairs, DNA, 700–704
Disaccharides, 255, 260–61
Disease:
  amethopterin and, 548
  amino acid catabolism, 589
  citric acid cycle and, 353
  degradative pathways and, 570
  linked to purine catabolsim, 590, 592

molecular, 154–56
primary cilia and, 67
proteostasis network and, 786
Dispersed core promoters, eukaryotic transcription, 724–25
Displacement reactions, P-19. *See also* Substitution reactions
Dissipative energy systems, 123
Disulfide bond formation, 779
Disulfide bridges, 147
Disulfide exchange, 779
Diterpenes, 416, 416*t*
D loop, tRNA, 665
DNA, 271, 590, 638–60
A-, 646, 647*f*, 647*t*
B-, 638, 641, 646, 647, 647*f*, 647*t*
chimeras and, 687
damage to, 643–45
denaturation of, 666, 667*f*
diagrammatic view of, 12*f*
double helix, 11–12, 12*f*, 636, 636*f*, 638, 640, 692
double-strand breaks (*see* Double-strand breaks)
epigenetics and (*see* Epigenetics)
in forensic investigations, 671–72
functions of, 6*t*, 635, 636
human sequences of, 634–35
information flow, 637*f*, 638
"junk," 634, 658
mitochondrial, 59, 671–72
mutations in (*see* Mutations)
noncoding, 12, 658
organelle, 656
recombination (*see* Recombination, DNA)
regulatory sequences of, 658
renaturation of, 666, 667*f*
repair in, 699–704
repetitive, 659–60
replication in (*see* Replication, DNA)
RNA compared with, 664–65
satellite, 659
strain effect on, 650*f*
strand structure, 639*f*
structure models, 636*f*
structure of, 11–12, 638–48, 639*f*, 640*f*
structure of discovered, 635–36, 645–46
supercoiling, 648–50, 649*f*, 650*f*, 692
synthesis, 689–92, 695–97
Z-, 647–48, 647*f*, 647*t*
DnaA, 693*f*, 694
DnaA boxes, 694
DnaB, 690, 694
DNA-binding proteins, 688–90, 688*f*, 694
DnaC, 694
DNA fingerprinting, 671
DNA fragments, joining, 692
DNA glycosylases, 700
DNA ligase, 692
DNA loops, 726*f*

DNA methylation, 548, 662
DNA microarrays, 716
DNA polymerase III, 692*f*, 694–95
DNA polymerases, 690, 691*f*, 692*f*, 696
DNA profiles, 671, 671*f*, 687
DNase, 135*f*, 590
DNA sequencing, 667–70, 670*f*
DNA typing, 671
DNA unwinding helicases, 690
Docking protein, 780
Dolichols, 416, 416*t*
Domains, 159–60, 159*f*
L-DOPA, 589
Dopamine, 543
Double bonds, 6*t*, 458, P-13, P-24, P-26
Double-displacement reactions, 225–26, 534–35, P-19
Double helix, DNA. *See* DNA, double helix
Double-strand break repair (DSBR) model, 706–7, 707*f*
Double-strand breaks (DSBs), 700, 703–4
Downregulation (receptor reduction), 605
Dragline silk, 134, 181–83, 182*f*
Drosha, 738
Drosophila, 717, 779
Drosophila melanogaster, 651
Duchenne muscular dystrophy, 728
Dynamin, 52
Dysbiosis, 35
Dyslipidemia, 615
Dystrophin gene, 728

E

Early endosomes, 52
Early response genes, 740–41
Ebola, 678–79, 678*f*
E1cb reactions, P-33–34
Ectoplasm, 87
Edman degradation method, 184
Effective concentrations, 200
Effectors, 164
EF hand, 159*f*
EF-Tu, 764, 765
EF-Tu-EF-Ts cycle, 764*f*
Eicosanoids, 406–7
Einstein's equation, 112, 501
Elasticity, 134
Electrochemical cells, 333, 333*f*
Electrolytes, P-8
Electromagnetic spectrum, 500*f*
Electron acceptors, 113, 335
Electron affinity, P-6, P-7
Electron donors, 113
Electronegativity, P-6, P-7
Electron groups, P-11
Electron microscopy (EM), 69
Electrons:
  configuration in atoms, P-4–5

flow and energy, 339*f*
valence, P-4, P-9
Electron transport, 364–73
components in, 364–71
cyclic, 507, 508*f*
fluid and solid state models, 372–73
inhibitors of, 373, 373*f*
noncyclic, 507, 508*f*
in photosystem I, 507, 507*f*
in photosystem II, 505, 505*f*
through cytochrome b_6f complex, 498, 498*f*
uncoupled, 383–84
Electron transport chain/pathway (ETC), 19, 331
electron transfer through complex II, 367*f*
energy relationships in, 372*f*
properties and structure of, 365*f*
supramolecular components of, 372–73
to UQ via $FADH_2$, 368*f*
Electron transport system, 364
cytochrome P_{450}, 485–86, 486*f*
defined, 293
Electrophiles, 14, 204, P-24
Electrophilic addition reactions, P-24
Electrophilic aromatic substitution reactions, P-26
Electrophoresis, 186–87, 186*f*
Electroporation, 713
Electrospray ionization, 187
Electrostatic effects, 207
Electrostatic interactions, 80, 160, 641
Elimination reactions, 15, 17*f*, P-32–34
Elongation. *See* Protein synthesis, elongation phase; Transcription, elongation phase
Elongation Factors (protein synthesis,) 135*f*, 764, 764f, 772, 772f
Embden–Meyerhof–Parnas pathway, 283. *See also* Glycolysis
Emergent property, 23
Emulsifying agents, 409
Enantiomers, 141, 141*f*, 248
Encephalopathy, 353
ENCODE project, 658, 686*f*, 717
Endergonic processes, 119
Endocrine hormones, 601–2
Endocytic cycle, 52, 54
Endocytosis, 52–54, 570
Endogenous pathway, 445, 472
Endogenous retroviruses, 660
Endomembrane system, 45
Endoplasm, 87
Endoplasmic reticulum (ER), 48–49, 780, 782*f*
rough, 48, 48*f*, 49
smooth, 48, 48*f*, 49
transitional, 49
Endorphins, 402
Endosomal lysosomal system, 570
Endothermic reactions, 115, P-15
Endotoxins, 42
Enediols, 205, 206f, 207, 254

Energy, 110–32
ATP role in, 124–27, 125*f*
conservation of, 114, 232
defined, 19, 112
free (*see* Free Energy)
oxygen role in generation of, 330
processes in, 19
protein folding and, 169*f*
and reducing power, 127
renewable, 492
secrets of life and, 111–12
thermodynamics of (*see* Thermodynamics)
transformation of, 110
Energy transfer pathways, 20
Enhancers, 658, 724
Enol form, 291, 554, 556f
Enterotoxin, 611
Enthalpy, 113–15
Entropy, 113, 114, 117–18, 124
Envelope proteins, 676
Enzyme cascade, 604
Enzyme induction, 233
Enzyme kinetics, 218–32
defined, 219
double-displacement reactions (*see* Double-displacement reactions)
Lineweaver–Burk plots, 224, 225*f*, 230
metabolism, macromolecular crowding, and, 231–32
Michaelis–Menten (*see* Michaelis–Menten kinetics)
multisubstrate reactions, 225–26
studies, 220*f*
Enzyme-linked receptors, 610–14
Enzymes, 195–244
allosteric, 231, 231*f*, 233–35
antioxidant systems, 389–91
catalysis by, 151, 204–18
as catalysts, 3, 198, 198*f*
classification of, 201–4
core, 725, 725*f*
debranching, 317, 320*f*, 323
defined, 3
holo-, 200
inhibition of, 226–31
marker, 69
properties of, 199–201
recommended names, 201
regulation of, 232–36
restriction, 667, 668*f*, 669
ubiquitin-activating, 572
ubiquitin-conjugating, 572
Enzyme specificity, 200–201
Epidermal growth factor (EGF), 611
Epigenetics, 652
defined, 662
processes in, 662–64
Epigenomes, 662
Epimerization, 254

Epimers, 248
Epimutations, 664
Epinephrine, 542
  COMT-catalyzed methylation of, P-33*f*
  glycogen metabolism and, 319
Epoxides, 486, 645
Equilibrium constants, 200, P-16–17
Erasers (epigenetic), defined, 663
ER (endoplasmic reticulum)-associated protein degradation, 49
E1 reactions, P-32
E1cb reactions, P-32–34
E2 reactions, P-33
ER-Golgi intermediate compartment (ERGIC), 49
ER (endoplasmic reticulum) stress, 48–49, 395
Escherichia, 34*f*
Escherichia coli, 36, 39, 172*f*, 233, 234, 649, 676, 692*f*, 713, 714*f*, 717
  ATP synthase from, 376, 377*f*
  chromosomes, 650, 651*f*
  citrate synthase inhibition in, 351
  DNA recombination in, 704, 705, 707, 709, 709*f*
  DNA repair in, 700, 702
  DNA replication in, 689, 690, 693–95, 695*f*, 696*f*
  EF-Tu-EF-Ts cycle in, 764*f*
  food poisoning and, 709
  gene expression in, 732, 733
  genetic code in, 652
  genomes of, 657, 657*f*
  glycolipid binding and, 415
  heat shock response in, 786
  lac operon in, 733*f*
  protein synthesis in, 753, 766–67
  pyruvate dehydrogenase complex in, 342, 342*t*
  ribosomes of, 673, 762*f*
  RNA polymerase in, 719, 719*f*
  rRNA processing reactions in, 722–23, 722*f*
  transcription in, 719, 719*f*, 720*f*, 767*f*
  translation in, 767*f*
  traveler's disease and, 611
Essential amino acids (EAAs), 532, 532*t*
Essential fatty acids (EFAs), 402, 405
Essential oils, 416
Ester hydrolysis, 207–9, 208*f*
Esterification, 254
Esters:
  properties of, P-28
  structure and significance of, 6*t*
17-*β*-Estradiol, 419*f*
ETC. *See* Electron transport chain/pathway
Ethane, 5*f*
Ethanol (ethyl alcohol), 237–38, 293
  oxidation-reduction reactions and, 18
  S. cerevisiae metabolism of, 296–97, 296*f*
Ethene, P-14*f*
Ethers, P-28–29
Ethyl alcohol. *See* Ethanol
Euchromatin, 55, 653

Eukaryotes:
characteristics of, 36
chloroplasts (*see* Chloroplasts)
chromosomes, 651
cytoskeleton (*see* Cytoskeleton)
DNA recombination in, 708–9
DNA replication in, 697–99, 697*f*, 698*f*
DNA synthesis in, 695–97
endocytic pathway, 52–54
endoplasmic reticulum (*see* Endoplasmic reticulum)
gene expression in, 735–41, 742*f*
genomes of, 657, 657*f*, 658
Golgi apparatus, 49–51, 50*f*
methylated cap of mRNA, 728*f*
mitochondria, 58–61, 61*f*
nucleus, 55–58, 55*f*
peroxisomes, 61–62
plasma membrane, 45–48
protein synthesis in, 768–85
ribosomes of, 672–73
structure of, 44–66
transcription in, 723–30
Evapotranspiration, 77
Evolution, 22
nonequilibrium thermodynamics and, 123
oxygen and, 330
primary protein structure, molecular diseases, and, 154–56
Excision repair, 700, 702*f*
Excitatory neurotransmitters, 551
Excluded volume, 39, 40*f*
Exergonic processes, 119
Exocytosis, 41, 51, 51*f*
Exogenous pathway, 445
Exon junction complex (EJC), 728*f*
Exons, 658
Exonucleases, 690
Exothermic reactions, 115, P-17, P-20
Exportins, 57
Expression platforms (riboswitches), 735
Extracellular matrix (ECM), 45

F

Facilitated diffusion, 428
Facultative anaerobes, 296, 331
Facultative heterochromatin, 55, 653
FAD. *See* Flavin adenine dinucleotide
$FADH_2$, 19, 127, 341, 342, 347–48
Familial hypercholesterolemia (FH), 435
Farnesene, 416*t*
Fasting. *See* Feeding-fasting cycle
Fats:
dietary, 445–46
neutral, 408
properties of, 408
Fatty acid-binding proteins, 451
Fatty acids, P-28

activation of, 14, 15*f*
biosynthesis of, 460–68, 462*f*, 468*t*
degradation of, 451–54
elongation and desaturation of, 466
essential, 402, 405
examples of, 404*t*
general function of, 6*t*
metabolism regulation in, 468–71, 469*f*
monounsaturated, 405
nonessential, 405
odd-chain, 458–59
omega-3, 405–6
omega-6, 405
oxidation of, 455–60, 467
polyunsaturated, 405
properties and types of, 10, 403–6
reactions of, 406
saturated, 10, 404, 404*t*
structure of, 10*f*, 403, 403*f*
unsaturated (*see* Unsaturated fatty acids)
Fatty acid synthase (FAS), 464–66, 465*f*
Feedback control, 24–25, 25*f*
Feedback inhibition, 234, 235*f*, 604–5
Feeding-fasting cycle, 618–25
fasting phase, 622–23
feeding behavior, 623–25
feeding phase, 619–22
Fe protein, 529–31, 529*f*
Fermentation, 237, 280*f*, 293, 296–97
Ferredoxin-thioredoxin system, 518, 518*f*
Fibrilogenesis, 175
Fibroin, 172, 174*f*
Fibronectin, 160, 161*f*
Fibrous proteins, 152, 172–75
Firmicutes, 35
Fischer, Emil, 200, 247
Fischer projections, 247, 249, 250, 251*f*
Fission (mitochondria), 59
Flagella:
eukaryotic, 63–64
prokaryotic, 44
Flavin adenine dinucleotide (FAD), 19, 213, 214*t*, 335–36, 337*f*
Flavin-containing monoxygenases, 485
Flavin mononucleotide (FMN), 213, 214*t*, 335–36, 337*f*
Flavoproteins, 335–36
Flippase, 423
Floppase, 423
Fluid mosaic model, 422
Fluid state model, 372–73
Fluorescence, 502
Fluorescence microscopy, 70, 70*f*
Fluorophores, 70
Flux, 235
Flux control, 305
FMN. *See* Flavin mononucleotide
Focused core promoters, 724

Folds, 159
Folic acid, 214*t*, 544
Food poisoning, 709
Forensic investigations, 671–72
Formaldehyde, 543, 544*t*
Formate, 544, 544*t*
Fos gene, 740–41
Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), 188
FOXP2 gene, 634
Fragility (robust systems), 25
Frameshift mutations, 643, 755
Franklin, Rosalind, 646
Free energy, 114, 118–24
  coupled reactions, 120, 121*f*
  defined, 113
  glycolysis and, 294–95
  hydrophobic effect, 122–24
  standard changes, 119
Free energy of activation, 198
Free radicals, 205
Fructofuranose, 258*f*
Fructose, 8
  Fischer and Haworth representations of, 251*f*
  hypertension and, 600
  hypertriglyceridemia and, 467–68
  metabolism of, 313–15
  obesity and, 626–27
  properties and functions of, 258
  structure of, 9*f*
Fructose-1,6-bisphosphate, 120, 121*f*, 286–87, 298–99
  in Calvin cycle, 512–13
  cleavage of, 287
  conversion to fructose-6-phosphate, 303
Fructose-2,6-bisphosphate, 298–300, 299*f*
Fructose-6-phosphate, 120, 121*f*, 309
  fructose-1,6-bisphosphate conversion to, 303
  glucose-6-phosphate conversion to, 286
  phosphorylation of, 287
Fumarate:
  hydration of, 17*f*, 348
  succinate oxidation to form, 348
Functional genomics, 26, 712
Functional groups, 5, 6*t*, P-26
Furan, 250, 251*f*
Furanoses, 250, 251
Furanosides, 255
Fusion, membrane (exocytosis), 59

G

GABA. *See* Gamma-aminobutyric acid
GADP. *See* Glyceraldehyde-3-phosphate dehydrogenase
Galactopyranose, 258*f*
Galactose, 258–59
Galactosemia, 259
Gallstones, 482
Gamma-aminobutyric acid (GABA), 7, 8*f*, 140

Gamma-carboxyglutamate, 141*f*
Gamma-glutamyl cycle, 533, 550*f*, 551
Gamma-linolenic acid, 404*t*
Gamma-synuclein gene, 744
Gangliosides, 414
Gangrene, 413
GAP. *See* Glyceraldehyde-3-phosphate and GTPase activating protein
Garrod, Archibald, 589
Gas gangrene, 413
Gasoline combustion, 116–17, 118*f*
Gasotransmitters, 552–54
Gastritis, 272
Gaucher's disease, 415*t*
GC box, 724
GEF. *See* Guanine nucleotide exchange factor
Gel electrophoresis, 186, 186*f*
Gel-filtration chromatography, 184, 185*f*
Gels, 86–87
Gene duplication, 643, 711
Gene expression, 732–41
  defined, 637
  epigenetics and, 664
  in eukaryotes, 735–41, 742*f*
  glucose-induced, 300
  process of described, 13
  in prokaryotes, 657–58, 732–35
General acids, 208*f*, 209
General bases, 208*f*, 209
General chemistry, P-2–22
General (homologous) recombination, 700, 704–7, 705*f*
General transcription factors, 725
Genes, 12, 635
  constitutive, 732
  defined, 4, 636
  delayed response, 741
  early response, 740–41
  enzyme control and, 233
  housekeeping, 724, 732
  inducible, 732
  marker, 713
  pseudo-, 659
  tumor suppressor, 740, 743, 744
Gene silencing, 738–39
Genetic code, 754–59
  aminoacyl-tRNA synthetase reaction, 757–59, 758*f*
  codon-anticodon interactions, 756–57, 756*f*
  codon assignments, 754*t*
  defined, 754
Genetic information, 637*f*, 638, 689–711, DNA Genetics, DNA; Repair, DNA; Replication. *See also* Recombination
  defined, 635
  reverse, 675
Genomes:
  defined, 12, 636
  eukaryotic, 657, 657*f*, 658
  hierarchical nuclear organization of, 655*f*

human, 661, 661*f*
projects, 716–17
prokaryotic, 657–58, 657*f*, 736
structure of, 656–60
Genomic libraries, 715
Genomics, 26
defined, 712
functional, 26, 712
technology of, 712–18
Geometric isomers, P-24
Geothermal energy, 112
Geraniol, 416*t*
Ghrelin, 619, 626
Gibbs free energy equation, 119, 119*f*
Gibbs, Josiah, 119
Glioblastoma, 349
Global genomic repair (GG-NER), 700
Global warming, 492
Globin, 177
Globin fold, 159
Globular proteins, 152, 176–80
Glucagon, 299, 319, 323
Glucogenic amino acids, 306, 575
Glucokinase, 297
Gluconeogenesis, 283, 301–9
defined, 301
glycolysis and, 304*f*
glyoxylate cycle in, 355*f*
reactions in, 301–5
regulation of, 308–9
substrates, 306
Glucopyranose, 258*f*
Glucose:
anomers of, 250*f*
ATP reaction with, 16*f*
ATP synthesis from oxidation of, 380–83, 381*t*
complete oxidation of, 380–83
conformational representations of, 251*f*
equilibrium mixture of, 252*f*
ethanol synthesis from, 237*f*
formation from glucose-6-phosphate, 305
gene expression induced by, 300
gluconeogenesis and (*see* Gluconeogenesis)
glycolysis and (*see* Glycolysis)
Haworth structures of, 251*f*
importance as an energy source, 8
isomerization of, 254
oxidation products of, 252*f*
properties and functions of, 257
removal from glycogen nonreducing ends, 317
structure of, 9*f*
Glucose-alanine cycle, 306, 307*f*, 534
Glucose-1-phosphate, 316, 317
Glucose-6-phosphate, 15, 120, 121*f*, 283
conversion to fructose-6-phosphate, 286
glucose formation from, 305

synthesis of, 286
Glucose-6-phosphate dehydrogenase deficiency, 156
Glucose sensors, 298
Glucose tolerance, 301
Glucose transporter, 429, 430*f*
Glucosuria, 615
Glutamate (glutamic acid), 139, 140, 532*t*, 536
abbreviations, 138*t*
biosynthesis of proline and arginine from, 538*f*
catabolic pathway of, 585, 585*f*
ionization state, 137*f*
5-oxyproline conversion to, 550, 550*f*
pK_a values, 142*t*
titration of, 143, 143*f*
Glutamate family, 536*f*, 537–40
Glutamic acid. *See* Glutamate
Glutamine, 138–39, 532*t*, 536
abbreviations, 138*t*
catabolic pathway of, 585, 585*f*
ionization state, 137*f*
pK_a values, 142*t*
Glutathione (GSH):
amino acid sequence of, 148*t*
biosynthesis of, 550*f*
functions of, 149, 549–50
Glutathione-centered system, 389, 390*f*
Glutathione peroxidase, 389, 390*f*
GLUTs, 619
GLUT4 transporter, 612
Glycans, 261–66
heteroglycans, 261, 264–66
homoglycans, 261–63
N-Glycans, 265
O-Glycans, 266
Glycation reactions, 256–57
Glyceraldehyde, 141, 142*f*, 247, 247*f*
Glyceraldehyde-3-phosphate (GAP), 284, 309
in Calvin cycle, 512, 513
interconversion from DHAP, 288
isomerization reactions and, 17, 18*f*
oxidation of, 288
Glyceraldehyde-3-phosphate dehydrogenase (GADPH), 152, 289*f*
Glycerate-2-phosphate, 290–91, P-34*f*
Glycerate-3-phosphate, 290
Glycerol phosphate shuttle, 380, 381*f*
Glyceroneogenesis, 448, 450*f*
Glycine, 138, 140, 532*t*, 540
abbreviations, 138*t*
biosynthesis of, 540*f*
catabolic pathway of, 582, 582*f*
ionization state, 137*f*
pK_a values, 142*t*
structural formula for, 7*f*
Glycocalyx, 43, 45, 47*f*, 269, 270*f*
Glycoconjugates, 266–69

Glycoforms, 273
Glycogen, 262, 283
  degradation of, 319*f*, 320*f*, 321*f*
  nonreducing ends, 263*f*, 317
  properties and structure of, 263
  reducing ends, 263*f*
  synthesis of, 318*f*
Glycogenesis, 283, 316–17
Glycogenin, 317
Glycogen metabolism, 315–23
  glycogenesis and, 316–17
  glycogenolysis and, 317
  major factors affecting, 322*f*
  regulation of, 317–23
Glycogenolysis, 283
  adenylate cyclase second messenger system and, 608*f*
  reactions in, 317
Glycolate-2-phosphate, 514*f*
Glycolipids, 266, 413–14, 414*f*
Glycolysis, 15, 17, 24, 283–301, 308*f*
  carcinogenesis and, 356–57
  defined, 20
  energetics of, 294–95
  fates of pyruvate in, 291–94, 294*f*
  gluconeogenesis and, 304*f*
  pentose phosphate pathway and, 312
  reactions in, 285*f*, 286–91, 292*f*
  regulation of, 295–301
  stages of, 284
  turbo jet engines compared with, 281, 281*f*
Glycolytic pathway:
  equation, 284
  reactions in, 285*f*, 286–91, 292*f*
Glycomes, 273
Glycoproteins, 153, 268–69
  functions of, 269
  heteroglycans linked to, 265*f*
  types of, 270*t*
Glycosaminoglycans (GAGs), 266
Glycoside formation, 255
Glycosidic bonds, 260, 260*f*, 317
Glycosidic linkage, 255
Glycosphingolipids, 413
Glycosylation, 49, 271
  in eukaryotic protein synthesis, 776–77
  reactions, 256–57
Glycosylphosphatilyinositol anchors. *See* GPI anchors
Glyoxylate cycle, 283, 354–55, 354*f*, 355*f*
Glypicans, 267, 267*f*
GMP, 557, 558*f*, 559
Goldstein, Joseph, 421, 435
Golgi apparatus/complex, 49–51, 50*f*
Golgi membrane, 780, 782*f*
Gonorrhea, 415
Gout, 559, 590, 592
GPI anchors, 264, 266, 411, 412*f*, 413

G-protein-coupled-receptors (GPCRs), 605–7, 606*f*
G proteins, 606, 606*f*
Gramicidin, 374
Gram-negative bacteria, 42
Gram-positive bacteria, 42
Grana/granum, 62, 495
Greek key motif (supersecondary structure), 158, 158*f*
Griffith, Frederick, 645, 708
Gro-ES-GroEL complex, 172, 172*f*
Growth factors, 603
GSH. *See* Glutathione
GTPase-activating protein (GAP), 607, 610, 741
GTPase-associated region (GAR), 762–63
Guanine, 11, 11*f*
Guanine-cytosine (GC) pair, 638, 640, 640*f*, 643, 644–45
Guanine nucleotide exchange factor (GEF), 606, 741, 764
Guanosine monophosphate. *See* GMP
Guanylate cyclase receptors, 610–11
Guide strand (gene silencing), 738–39
Gulf of Mexico dead zone, 527, 527*f*

H

Haber–Bosch reaction, 529, P-17, P-20
Haber process, P-14
Hairpin turn, 158
Half-cells, 333
Half-lives, 220, 571, 573*f*
Half-reactions, 333, 334, 334*t*
Halogenation reactions, P-23
Hatch–Slack pathway, 515
Haworth structures, 250, 251*f*, 258*f*
Haworth, W. N., 250
Heart disease, 364, 484. *See also* Cardiovascular disease
Heart failure, 418, 588
Heat, 114
Heat capacity, 83–84
Heat of fusion, 83, 84*t*
Heat of vaporization, 83–84
Heat shock factors (HSF), 786
Heat shock proteins (hsps), 152
  hsp70s, 171
  hsp90s, 171
  hsp60s, 172
Heat shock response (HSR), 786
Heat-stable enterotoxin, 611
Heat stress, 518
Heavy metal ions, 168
Hedgehog pathway, 67
Helicases, 690
Helices. *See* Alpha-helices
Helicobacter pylori, 272, 664
Helix-loop-helix units, 158
Heme, 176–77
  oxygen-biding site of, 178*f*
  properties of, 563–64
  structure of, 177*f*

Hemiacetals, 248, 250*f*, P-27, P-27*f*
Hemiketals, 248, 250*f*
Hemoglobin:
  allosteric transition of, 179*f*
  BPG effect on oxygen affinity of, 180*f*
  overview of, 135*f*
  oxygen dissociation curve of, 178, 178*f*
  pH and, 102
  properties and functions of, 176–80
  structure of, 135*f*, 179f
  sickle-cell, 155–56, 155*f*
  subunits of, 179*f*
Hemolysis, 90
Hemophilia A, 660
Hemophilia B, 660
Hemoproteins, 153
Henderson–Hasselbalch equation, 96–97
Henri, Victor, 220
Henseleit, Kurt, 576
Heparan sulfate proteoglycans (HSPGs), 266–67
Hepatomegaly, 323
Herbicides, 510
Hereditary nonpolyposis colorectal cancers (HNPCC), 703
Heterochromatin, 55, 653
Heteroduplex, 709
Heteroglycans, 261, 264–66
  glycoproteins linked to, 265*f*
Heterokaryon, 423
Heterotrophs, 19
Heterotropic effects (allosteric regulation), 234
Heterozygotes, 156, 435
Hexane, 5*f*
Hexokinases:
  ATP-binding domain of, 159*f*
  glycolysis regulation and, 295–98
  specificity of, 200–201
Hexose monophosphate shunt, 313. *See also* Pentose phosphate pathway
Hexoses, 247
HGPRT, 557, 559
HIF-1, 356
High-density lipoproteins (HDL), 421–22
High-energy bonds, 126
High-fructose corn syrup, 467–68, 600
High performance liquid chromatography, 186, 787
High-throughput technologies, 717, 718
Histidine, 139, 532*t*
  abbreviations, 138*t*
  acid-base catalysis and, 209
  biosynthesis of, 544*f*
  catabolic pathway of, 585*f*, 586
  ionization state, 137*f*
  pK_a values, 142*t*
Histidine family, 536*f*, 542
Histone acetyltransferases (HAT), 663
Histone code hypothesis, 662–64
Histone deacetylases (HDAC), 663–64

Histone fold, 652
Histones, 55, 653*f*
modifications of, 662–64, 663*f*
properties and structure of, 651–52
HIV. *See* Human immunodeficiency virus
HMG-CoA, 476, 484
HMG-CoA reductase (HMGR), 476, 478*f*, 482–84, 482*f*
Holliday junction, 709
Holliday model, 705, 705*f*, 706
Holliday, Robin, 705
Holoenzymes, 200
Holoproteins, 153
Homeostasis, 149
cholesterol, 482–84
defined, 3, 282
Homoglycans, 261–63
Homologous modeling, 188
Homologous polypeptides, 153
Homologous recombination. *See* General recombination
Homotropic effects, 234
Homozygotes, 156, 435
Hormone response elements (HREs), 614
Hormones, 40, 140
endocrine, 601–2
fatty acid metabolism and, 470–71
gluconeogenesis regulation and, 309
glycolysis regulation and, 299
intercellular communication and, 603–17
lipid-soluble, 603
thyroid, 602, 604
water-soluble, 604
Hormone-sensitive lipase, 447
HOTAIR, 744–45
Housekeeping genes, 654, 724, 732
hsps. *See* Heat shock proteins
Huckel's rule, P-26
Human accelerated regions (HARs), 634
Human alcohol dehydrogenases (ADHs), 218*f*, 219*f*, 237–38
Human endogenous retroviruses (HERVs), 660
Human Genome Project, 26, 271, 686*f*, 716–17
Human immunodeficiency virus (HIV), 638, 676
Humans:
DNA sequences in, 634–35
feeding behavior in, 624*f*
genome of, 661, 661*f*
hierarchical organization of systems, 4*f*
microbiota in systems of, 34–35
Hund's rule, P-4
Huntington's disease, 40, 172, 786
Hurler's syndrome, 268
Hutchinson–Guilford progeria syndrome, 699
Hybridization, 667, 668*f*
Hydration, 17, P-25, P-25*f*
defined, 79
DNA and, 640–41
fumarate, 17*f*, 348

protein, 161, 162*f*
Hydrocarbons, P-23–26
aromatic, 138, P-23, P-25–26
branched-chain, P-23
cyclic, P-23, P-25
defined, 5
saturated, P-23
structural formulas of selected, 5*f*
substituted, P-22, P-26–31
unsaturated, P-23, P-24
Hydrogenation, P-24
partial, 408
Hydrogen bonds, 79–80, 80*f*, 81*t*, 640, P-27
defined, 11–12
properties of, 81–82
tertiary protein structure and, 160
between water molecules on ice, 83, 83*f*
Hydrogen selenide:
heat of fusion and, 84*t*
melting and boiling points of, 83*t*
Hydrogen sulfide (H_2S), 552
biosynthesis of, 553*f*
heat of fusion and, 84*t*
melting and boiling points of, 83*t*
properties of, 554
Hydrogen telluride, 83*t*
Hydrolases, 202, 202*t*
Hydrolysis, 14–15, 16*f*
ATP (*see* ATP, hydrolysis of)
defined, 14
ester, 207–9, 208*f*
of phosphorylated biomolecules, 126*t*
triacylglycerol, 448–51
Hydrolytic reactions, 643
Hydronium ions, 92, 93*f*
Hydrophilic molecules, 7, 37, 85–87, P-27
defined, 36–37
sol-gel transitions, 86–87
structured water, 86, 86*f*
Hydrophobic effect, 87–88, 89*f*
defined, 87
free energy and, 122–24
Hydrophobic interactions:
DNA and, 640
protein folding and, 169
tertiary protein structure and, 160
between water and a nonpolar substance, 37*f*
Hydrophobic molecules, 5, 7, 38, P-23
defined, 37
properties of, 87–88
Hydrosphere, 111
Hydroxide ion catalysis, 207–9, 208*f*
Hydroxylation, 777–78, 778*f*
Hydroxyl groups, 5, 6*t*
5-Hydroxylysine, 141*f*
4-Hydroxyproline, 141*f*

Hyperammonemia, 579
Hyperglycemia, 615
Hyperinsulinemia, 616
Hyperlipoproteinemia, 615
Hyperosmolar hyperglycemic nonketosis (HHNK), 616–17
Hypertension, 384, 600
Hyperthermia, malignant, 305
Hypertonic solutions, 90, 91*f*
Hypertriglyceridemia, 467–68
Hypochromic effect, 666
Hypoglycemia, 323
Hypotonic solutions, 90, 91*f*
Hypoxanthine-guaninephosphoribosyltransferase. *See* HGPRT
Hypoxia inducible factor 1 (HIF-1), 356

I

Ideal solutions, 200
Illumina sequencing, 669
Immune system, 35
Immunoglobulin G (IgG), 163*f*
IMP, 557, 558*f*, 559
Importins, 57
Impotence, 386
Incretins, 621
Indels, 642–43
Induced dipole-induced dipole interactions, 82, 82*f*
Induced-fit model, 200, 201*f*
Inducible genes, 732
Inferior frontal cortex, 634
Influenza, 227
Inhibition/inhibitors (enzyme), 226–31
  competitive, 226–29, 229*f*
  irreversible, 226, 229–30
  noncompetitive, 226–28
  protease, 226
  reversible, 226
  uncompetitive, 226, 228, 229, 229*f*
Inhibitory neurotransmitters, 551
Initiating codons, 754
Initiation. *See* Protein synthesis, initiation phase; Transcription, initiation phase
48S Initiation complex, 769, 769*f*
80S Initiation complex, 771*f*
Initiator tRNA, 759
Inner mitochondrial membrane, 59
Inner nuclear membrane (INM), 56
Inosine-5'-monophosphate. *See* IMP
Inositol-1,4,5-trisphosphate (IP_3), 609
Insertional (IS) elements, 709, 710*f*
Insig, 484
In silico processes, 232
Insulators, 658, 725
Insulin, 40–41, 184, 309, 447, 468, 611, 615, 616, 621, 624
  fatty acid metabolism and, 470–71
  glucose tolerance and, 301
  glycogen metabolism and, 319
  glycolysis and, 299

proteolytic processing of, 776, 776*f*
shape of, 135*f*
simplified model of signaling, 613*f*
structure of, 300*f*
Insulin-like growth factor 1 (IGF-1), 611
Insulin receptors, 269, 611, 611*f*
Insulin receptor substrate 1 (IRS-1), 611
Insulin resistance, 605, 616, 626
Integral proteins, 37, 38*f*, 425–26, 425*f*, 426*f*
Integrins, 269, 273
Intercalating agents, 645
Interchromatin channel network, 652
Intermediate-density lipoproteins (IDL), 421
Intermediate filaments, 63*f*, 65
Intermediates, reaction, 148, 204–5
Intermembrane space, 59
International units (IU), 223
Interspersed genome-wide repeats, 659
Intestinal (lower gastrointestinal) tract, 34*f*
Intima, 473
Intraflagellar transport (IFT), 64
Intrinsically unstructured proteins (IUPs), 165
Intrinsic apoptosis, 58
Intrinsic termination, 721, 722*f*
Introns, 658
Invariant residues, 153
Inversion, 643
Inverted repeats, 648, 709
In vitro processes, 231
In vivo processes, 197, 200, 232
Iodine test, 262
Iodoacetamide, 231
Ion-exchange chromatography, 184–85
Ionic bonds, P-9
Ionic interactions, 81, 81*t*
Ionization, 92–102
acids, bases, and pH in, 93–94
buffers in, 94–102
electrospray, 187
Ionization energy, P-6–7
Ionizing radiation, 643–44
Ionophores, 374
Ion product of water, 93
Ions, 92, P-9
Ion torrent sequencing, 669–70
IP_3, 609
Iron-sulfur clusters, 366*f*, 368
Irreversible inhibition, 226, 229–30
Ischemia, myocardial, 395
Isocitrate, 345
Isocitrate dehydrogenase, 351, 352–53
Isoelectric point, 143
Isolation membrane, 574–75
Isoleucine, 138, 532*t*, 542
abbreviations, 138*t*
catabolic pathway of, 586*f*, 587

degradation of, 588*f*
ionization state, 137*f*
pK_a values, 142*t*
Isomerases, 202, 202*t*
Isomeric forms of unsaturated molecules, 404, 404*f*
Isomerization:
citrate, 344–45
glucose, 254
Isomerization reactions, 17, 18*f*
Isomers:
cis-, 404
dextrorotary, 141
geometric, P-24
levorotary, 141
optical, 141
positional, P-24
trans-, 404
Isoniazid, 236
Isopentenyl pyrophosphate, 415*f*
Isoprene, 415*f*
Isoprene units, 415, 415*f*
Isoprenoid metabolism, 475–76
cholesterol biosynthetic pathway, 484
cholesterol metabolism, 475–76
Isoprenoids:
properties and types of, 415–18
synthesis of, 475
Isoprenyl groups, 416–17
Isothermic reactions, 115
Isotonic solution, 90
Isotopes, P-3

J

Jacob, François, 732
Jeffreys, Alec, 671
Jun gene, 740–41
"Junk" DNA, 634, 658

K

Katal, 223
Kayser–Fleischer rings, 212
Keratin, 172, 174*f*
Ketals, 255, 255*f*
Ketoacidosis, 615
Keto form, 291, 554, 556f
Ketogenesis, 456
Ketogenic amino acids, 575
Ketogenic diet, 457–58
Ketoglutarate. *See* Alpha-ketoglutarate
Ketone bodies, 456, 457*f*
Ketones:
properties of, P-28
structure and significance of, 6*t*
Ketoses, 202, 247
defined, 8
isomerization reactions and, 17, 18*f*

structure of, 247*f*
Ketosis, 456, 615
Khorana, Har Gobind, 754
Kinetically favorable processes, 114
Kinetics:
first-order reactions, 219
in chemical reactions, P-15–16
enzyme (*see* Enzyme kinetics)
second-order reactions, 220
zero-order reactions, 220
Klebsiella, 34*f*
Kornberg, Arthur, 690
Koshland, Daniel, 200, 235
Krabbe's disease, 415*t*
Krebs bicycle, 579, 580*f*
Krebs, Hans, 348, 576
Krebs urea cycle (Krebs–Henseleit cycle), 576. *See also* Urea cycle

L

Lac operon, 732–35, 734*f*, 767
Lactase, 635
Lactate, 293
Lactic acid, 94*t*
Lactic acid fermentation, 293
Lactic aciduria, 349
Lactobacillus, 34*f*
Lactone, 252
Lactose, 260, 260*f*, 732–35
Lactose intolerance, 261
Lactose persistence, 635
Lagging strand, DNA, 693–94
Lamina associated domains (LADs), 655
Large buoyant LDL, 421
Lariats, 728–30
Laser capture dissection microscopy (LCDM), 787
Late endosomes, 52
Late onset AD (LOAD), 436–37
Lathyrism, 176
LCT gene, 635
Leading strand, DNA, 693, 694
Lead poisoning, 230
Leaving groups, 14
Le Chatelier, Henri Louis, P-17
Le Chatelier's principle, 95, P-17
Lecithin, 410, 411*t*
Lecithin:cholesterol acyltransferase (LCAT), 422*f*
Lectins, 271–73
Leghemoglobin, 532
Leptin, 468, 624, 626
Lesch–Nyhan syndrome, 559, 660
Leucine, 138, 532*t*, 542
abbreviations, 138*t*
catabolic pathway of, 583*f*, 585
ionization state, 137*f*
pK_a values, 142*t*
Leucine zipper, 159*f*

Leukemia, 548, 592, 743
Leukocyte rolling, 271, 273
Leukotriene C_4, 407*f*
Levorotary isomers, 141
Lewis acid and base theory, P-19–20
Lewis acids, 211
Lewis dot notation, P-9–10, P-12
Lewis, G. N., P-9
Licensing factors, replication, 697
Life:
  characteristics of, 3–5
  energy and secrets of, 111–12
Ligands, 40, 160
Ligases, 202, 202*t*
Light, 500–503, 517–18
  amplitude, 500
  frequency, 500
  wavelength, 500–501, 501*f*
Light harvesting, 494
Light-harvesting antenna, 493
Light-harvesting complex I (LHCI), 497
Light-harvesting complex II (LHCII), 497
Light-independent reactions, 510–17
Light-induced repair, 699
Light reactions, 503–10
Light-responsive elements (LRE), 741
Lignoceric acid, 404*t*
Limit dextrin, 317
Linear molecules, P-12
LINEs (long interspersed nuclear elements), 653, 660
Lineweaver–Burk plots, 224, 225*f*, 230
Linezolid, 752
Linoleic acid, 404*t*
Linolenic acid, 404, 404*t*, 405*f*
Lipid bilayer, 37, 422, 424*f*
Lipid exchange, 60
Lipid metabolism, 443–90
  AMPK-regulated pathways in, 470*f*
  dietary fat digestion, absorption, and transport, 445–46
  fatty acid biosynthesis, 460–68
  fatty acid degradation, 451–54
  fatty acid metabolism regulation, 468–71
  fatty acid oxidation, 455–60
  isoprenoid metabolism (*see* Isoprenoid metabolism)
  lipoprotein metabolism, 472
  membrane, 472–75
  triacylglycerol metabolism, 446–51
Lipid rafts, 427, 428*f*
Lipids, 401–42. *See also* Eicosanoids; Fatty acids; Isoprenoids; Lipoproteins; Phospholipases; Phospholipids; Sphingolipids; Triacylglycerols; Wax esters
  classes of, 403–22
  defined, 403
  functions of, 403
  membrane, 423–25
  properties of, 10
Lipid-soluble hormones, 603

Lipid-soluble vitamins, 213
Lipofuscin granules, 575
Lipogenesis, 299, 448, 622
Lipoic acid, 213, 214*t*, 341*t*, 342, 344*f*
Lipolysis, 448, 450*f*
Lipophilic modifications, 778
Lipoprotein lipase, 445,
Lipoproteins, 153
  components of, 421*f*
  high-density, 421–22
  intermediate-density, 421
  low-density, 422, 434–35
  metabolism of, 472
  properties and types of, 419–22
  structure of, 420*f*
  very-low-density, 421, 422
Lipotoxicity, 626
Lithosphere, 111
Liver failure, DNP and, 376
Liver functions, 4*f*
lncRNA, 673, 744–45
Loci (DNA markers), 671
Lock-and-key model, 200
London dispersion forces, 82
Lone pairs (electrons), P-11, P-12
Long interspersed nuclear elements (LINEs), 653, 660
Long ncRNA (lncRNA), 673, 675, 744
Loop extrusion model (chromatin), 653
Loops (polypeptide structural elements), 158
Lou Gehrig's disease. *See* Amytrophic lateral sclerosis
Low-density lipoprotein receptors, 434–35
Low-density lipoproteins (LDL), 422
Low-fat diet, 402
Luciferase, 714*f*
Lumen, 48, 62, 495
Lyases, 202, 202*t*
Lymphocytes, 35
Lymphoma, 743
Lysine, 139, 532*t*, 542
  abbreviations, 138*t*
  catabolic pathway of, 583, 583*f*
  hydroxylation of, 777–78
  ionization state, 137*f*
  pK_a values, 142*t*
Lysinonorleucine, 163, 163*f*
Lysogeny, 677
Lysophosphatide, 412
Lysosomal storage diseases, 40, 54
Lysosomes, 52
Lysozyme, 135*f*
Lytic cycle, 677
Lytic phase, 677

M

Macroautophagy, 574
Macromolecular crowding, 39, 200, 231–32

Macromolecules, 3
Macrophages, 35
Madagascar spider silk tapestry, 181
Magma, 112
Magnesium, photosynthesis and, 518
Magnetic quantum number m, P-4
Maillard reaction, 256, 257*f*
Malate:
  fumarate conversion to, 348
  hydration and, 17*f*
  oxidation to form OAA and NADH, 348
Malate-aspartate shuttle, 382
Malate shuttle, 303
Malignant hyperthermia, 305
Malignant tumors, 743
Malonate, 227*f*
Malonyl-CoA, 461, 463*f*
Maltose, 260, 260*f*
Mammalian target of rapamycin. *See* mTOR
Mammals:
  fatty acid metabolism regulation in, 468–71
  feeding-fasting cycle in, 618–25
  nutrient metabolism in, 620*f*
MAPKK, 741
Maple syrup urine disease, 589
Marker enzymes, 69
Marker genes, 713
Markovnikov's rule, P-25
Mass action, law of, 231
Mass number, P-2–3
Mass spectrometry (MS), 187–88, 187*f*, 787
Matrix (mitochondrial), 59
Matrix-assisted laser desorption ionization (MALDI), 187
Matrix proteins (mitochondrial), 783, 783*f*
Matthaei, Heinrich, 754
Mayan civilization, 77–78, 78*f*
MCAD deficiency, 453
McClintock, Barbara, 710
MCM complex, 697
Mechanical stress, 168
Media, 473
Mediator, 725, 725*f*, 727
Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, 453
Meiosis, 704, 706
Melanin, 589
Melatonin, 140, 549
Melissyl cerotate, 409*f*
Membrane fluidity, 423–24
Membrane lipid metabolism, 472–75
Membrane potential, 92
Membrane remodeling, 413
Membranes, 422–35
  asymmetry of, 425
  band 3 anion exchanger macrocomplex, 426–27
  chemical composition of selected, 423*t*
  chloride shift and, 426–27

functions of, 37, 428–35
Golgi, 780, 782*f*
importance of, 3
ion distribution across, 92
isolation, 574–75
lateral diffusion in, 423, 424*f*
lipid, 423–25
microdomains of, 427
mitochondria-associated, 60
plasma (*see* Plasma membranes)
protein, 425–26
pseudo, 774
receptors, 37, 434
selectively permeable, 90, 90*f*, 424–25
self-sealing, 425
structure of, 37, 38*f*, 401*f*, 423–27
thylakoid, 62, 495
transport across, 21, 428–33, 429*f*
Membrane skeleton, 45
Menkes' syndrome, 212
Menten, Maud, 220
Meselson, Matthew, 689
Meselson–Radding model, 706, 706*f*
Messenger RNA. *See* mRNA
Metabolic flux, 232
Metabolic networks, 24
Metabolic pathways, 20
Metabolic syndrome, 626–27
Metabolism, 599–632
aerobic (*see* Aerobic metabolism)
amino acid, 532–33
C4, 515–17, 516*f*
carbohydrate (*see* Carbohydrate metabolism)
cell volume regulation and, 103–4
cholesterol (*see* Cholesterol metabolism)
citrate, 352–53, 352*f*
defined, 3, 282
enzyme kinetics and, 231–32
fatty acid, 468–71, 469*f*
feeding-fasting cycle and, 618–25
fructose, 313–15, 626–27
glycogen (*see* Glycogen metabolism)
hormonal regulation of, 603–17
isoprenoid (*see* Isoprenoid metabolism)
lipid (*see* Lipid metabolism)
lipoprotein, 472
membrane lipid, 472–75
one-carbon, 543–49, 544*t*
overview of, 20, 601–2, 602*f*
phospholipid, 472–75
primary functions of, 13
signal transduction and, 41
triacylglycerol, 446–51
turbo jet engines compared with, 281
Metabolome, 637
Metabolons, 232

Metalloids, P-2
Metalloproteins, 153
Metals, 211–13, P-2
Metastasis, 743
Met-enkephalin, 8*f*
Metformin, 616
Methane, 5*f*, P-13*f*
Methanol, 543, 544*t*
Methanol poisoning, 231
Methicillin-resistant *Staphylococcus aureus* (MRSA), 752
Methionine, 138, 532*t*, 542
  abbreviations, 138*t*
  catabolic pathway of, 586–87, 586*f*
  ionization state, 137*f*
  pK_a values, 142*t*
Methotrexate, 548, 563
Methyl acetate, P-29*f*
Methylation, 779
  cytosine, 662*f*
  DNA, 548, 662
Methyl-CpG-binding proteins (MeCPs), 662
N-Methyl-D-aspartate (NMDA) receptors, 554
Methyl glucoside, 255*f*
Methylmalonic acidemia, 589
Methyltransferases:
  de novo, 662
  maintenance, 634
Micelles, 88, 89*f*
Michaelis constant, 222
Michaelis, Leonor, 220
Michaelis–Menten kinetics, 220–23, 231
  competitive inhibition and, 226*f*, 227
  noncompetitive inhibition and, 227–28, 228*f*
  plot, 224*f*
  uncompetitive inhibition and, 228
Microautophagy, 574
Microbiota, 33–34, 34*f*
Microfibrils, 264, 265*f*
Microfilaments, 63*f*, 64
Microheterogeneity, 273
Microprocessor, 738
Micro RNA. *See* miRNA
Microsatellites, 659
Microsomal triglyceride transfer protein (MTTP), 444
Microsomes, 68
Microtubule-associated proteins (MAPs), 63
Microtubules, 62–64, 63*f*
Migraine headache, 386
Mines, 111, 111*f*
Minisatellites, 659
miRNA, 13, 674, 738–39, 739*f*
Mismatch repair, 700, 702–3
Missense mutations, 642
Mitchell, Peter, 374
Mitochondria:
  aerobic metabolism in, 332*f*

cross section of, 329*f*
eukaryotic, 58–61, 61*f*
fatty acid transport into, 452*f*
Mitochondria-associated membranes (MAM), 60
Mitochondrial DNA. *See* mtDNA
Mitochondrial electron transport chain (ETC). *See* Electron transport system
Mitochondrial fission, 59, 60, 60*f*
Mitochondrial fusion, 59, 60*f*
Mitochondrial respiratory chain (MRC), 59
Mitogen-activated protein kinase (MAPK), 612, 741
Mitogen-activated protein kinase kinase (MAPKK), 612, 741, 742f
Mitogen-activated protein kinase kinase kinase (MAPKKK), 612
Mitotic spindle, 63
Mixed anhydride, defined, 14
Mixed anhydrides, 758
Mixed function oxidases, 485
Mixed noncompetitive inhibition, 228, 229, 229*f*
Mixed terpenoids, 417
Mobile genetic elements, 659
Mobile phase, 184
Modular proteins, 160
Modulators, 164
Modules, 24
Molarity, P-21
Molecular biology, 636, 646
Molecular chaperones, 38, 49, 170–72
mechanisms of action, 173*f*
role of, 170–72
types of, 171–72
Molecular cloning, 712–14, 713*f*
Molecular diseases, 154–56
Molecularity, 220
Molecular structure, P-10–12
Molten globules, protein folding, 170, 170*f*
Molybdenum-iron cofactor (MoFe cofactor; M cluster), 529–31, 529*f*
Monocistronic mRNA, 673
Monod, Jacques, 732
Monosaccharides, 247–60
cyclic structure of, 248–51
derivatives of, 259–60
important, 9*f*, 257–59
properties and functions of, 8
reactions of, 251–57
stereoisomers of, 248
Monoterpenes, 416, 416*t*
Monounsaturated fatty acids, 405
Moonlighting proteins, 152
Mosaic proteins, 160
Motifs, 24–25, 157–58, 158*f*
Motor proteins, 38
mRNA, 12–13, 637, 665
export, 784
monocistronic, 673
polycistronic, 673
pre-, 727
scanning, 768

secondary structure of, 768
stability, 784
structure of, 673
MRSA (methicillin-resistant *S. aureus*), 752
mtDNA, 59, 671–72
mTOR, 575, 612, 625
mTORC1 (mTOR complex 1), 612
mTOR-mediated translational control, 784
Mucin-type carbohydrate units, 269
Mucopolysaccharidoses, 268
Multifunctional proteins, 152
Multisubstrate reactions, 225–26
Multivesicular bodies, 52
Mutarotation, 251
Mutations, 641–45, 702–3
causes of, 643–45
defined, 5, 589
epi-, 664
frameshift, 643, 755
missense, 642
nonsense, 642
point, 642, 755
primary protein structure, diseases, and, 154–56
silent, 5, 642
transition, 642, 642*f*
transversion, 642
types of, 642–43
Myc gene, 356, 740–41
Mycobacterium, 34*f*
Mycobacterium tuberculosis, 570
Mycoplasma, 656
Myoadenylate deaminase deficiency, 592
Myocardial infarction, 384, 395, 406
Myocardial ischemia, 395
Myoglobin:
oxygen dissociation curve of, 178*f*
properties and functions of, 176–78
shape of, 135*f*
structure of, 177*f*
Myristic acid, 404*t*
N-Myristoylation, 778

N

NAD, 19, 213, 214*t*, 335, 336*f*
NADH, 19
in citric acid cycle, 341*t*, 342, 345, 351
formation via alpha-ketoglutarate oxidation, 346
formation via isocitrate oxidation, 345
formation via malate oxidation, 348
pyruvate reduction by, 332–33, 333*f*
recycling during glycolysis, 295*f*
role of, 127
shuttle mechanisms and, 380–82, 381*f*
NADH dehydrogenase complex (complex I), 364–66, 372–73, 372*t*, 385
NADH oxidase, 615
NADP, 213, 214*t*, 335

NADPH:
fatty acid biosynthesis, 461, 462f
pentose phosphate pathway and, 283, 309–10, 313
photosystem I and synthesis of, 507–9
role of, 127
NADPH oxidase, 388
NAG. *See* *N*-Acetylglutamate
Na^+-K^+ ATPase transport, 429, 430*f*
Nasal (upper respiratory) tract, 34*f*
Nascent chylomicrons, 445
Nascent polypeptide chains, 153
Nascent protein molecules, 158
Natively unfolded proteins, 165
ncDNA, 13, 658
ncRNA, 13, 637, 665, 673–75, 688
Negative feedback, 25, 25*f*
Negatively supercoiled DNA, 649, 650
Negative-sense RNA genomes, 677
Neisseria, 34*f*
Neisseria gonorrhoeae, 415
Nephrogenic diabetes insipidis, 430–31
Networks, 24
Neurotransmitters, 40, 551
amino acid and amine, 551*t*
defined, 7, 140
excitatory, 551
inhibitory, 551
retrograde, 552
Neutral fats, 408
Neutral solutions, 93
Nicotinamide adenine dinucleotide. *See* NAD
Nicotinamide adenine dinucleotide phosphate. *See* NADP
Nicotinic acid (niacin), 214*t*, 335
Niemann–Pick disease, 415*t*
Niemann–Pick disease type B, 570
Niemann–Pick disease type C, 570
Nif genes, 531
Nirenberg, Marshall, 754
Nitric oxide (NO), 386, 388f, 552
Nitric oxide synthase (NOS), 552*f*, 553, 553*f*
Nitrobacter, 531
Nitrogenase complex, 529, 530*f*, 531
Nitrogen assimilation, 531–32
Nitrogen cycle, 528
Nitrogen fixation, 528–32, P-17
by cyanobacteria, 526*f*, 528
defined, 528
nitrogen assimilation, 531–32
reaction, 529–31
Nitrogen metabolism, 526–68, 569–98. *See also* Amino acids, biosynthesis of; Amino acids, biosynthetic reactions involving; Amino acids, catabolism of; Neurotransmitters, degradation of; Nitrogen fixation; Nucleotides, degradation of; Protein turnover
Nitrogenous bases, 10, 11*f*
Nitrosomonas, 531
N-linked oligosaccharides, 261, 776, 777*f*, 780
NMDA receptors, 554

NMR spectroscopy, 188–89
Nobel Prize winners:
Agre, Peter, 431
Anfinsen, Christian, 167
Brown, Michael, 421
Crick, Francis, 646
Goldstein, Joseph, 421
Kornberg, Arthur, 690
Kornberg, Roger, 725
Pauling, Linus, 146
Watson, James, 646
Wilkins, Maurice, 646
Nonalkylating agents, 645
Noncatalytic amino acid residues, 210
Noncoding DNA. *See* ncDNA
Noncoding RNA. *See* ncRNA
Noncompetitive inhibitors, 226–28
Noncovalent bonding, 80–82, 81*t*
Noncyclic electron transport pathway (photosynthesis), 507, 508*f*
Nonequilibrium thermodynamics, 123
Nonessential amino acids (NAAs), 532, 532*t*
Nonessential fatty acids, 405
Nonhomologous end joining, 700, 703
Nonmetal elements, P-2
Nonoxygenic photosynthesis, 330
Nonpolar amino acids, 136, 137*f*, 138
Nonpolar hydrocarbon tails, 88, 89*f*
Nonreducing ends (glycogen), 263, 263*f*, 317
Nonreplicative transposition, 710, 711*f*
Nonsense-mediated decay (NMD), 729
Nonsense mutations, 642
Nonshivering thermogenesis, 384
Nonspontaneous processes, 117
Nonstandard amino acids, 136
Nonstop-mediated mRNA decay, 771
Non-Watson–Crick base pairs, 756
Norepinephrine, 383–84, 542
Nostoc muscorum, 528
N-terminal residues, 146, 572
Nuclear bodies, 58
Nuclear envelope (NE), 56
Nuclear lamina, 57–58, 57*f*
Nuclear matrix, 58
Nuclear pore complexes (NPC), 56–57, 56*f*
Nuclear pores, 55*f*, 56
Nuclear receptors, 614
Nucleases, 590
Nucleic acids, 633–85. *See also* DNA; RNA; Viruses
methods used to purify and characterize, 666–70
properties and types of, 10–13
Nucleocapsid, 677
Nucleoid, 44, 44*f*, 650
Nucleoid-associated proteins, 650
Nucleolar organizer region, 58
Nucleolus, 58
Nucleolus associated domains (NADs), 655

Nucleophiles, 14, 204, 209, P-24
Nucleophilic acyl substitution reactions, P-28
Nucleophilic addition reactions, 148
Nucleophilic substitution reactions, 14–15, 16*f*
Nucleoplasm, 55
Nucleoporins, 56
Nucleosidases, 590
Nucleoside diphosphate kinase, 347, 557
Nucleosides, 554–55, 556*f*
Nucleosomes, 651
Nucleotidases, 590
Nucleotide excision repair (NER), 700, 701
Nucleotides, 554–63
  catabolism of, 591*f*
  degradation of, 590–92
  functions of, 6*t*, 10–13, 554
  structure of, 11*f*
Nucleus, eukaryotic, 55–58, 55*f*
Nucleus tractus solitarius (NTS), 624
Nutrients, antioxidant, 392–93

O

OAA. *See* Oxaloacetate
Obesity, 402, 626–27
Obligate aerobes, 331
Obligate anaerobes, 331
O blood group, 272
Octet rule, P-5, P-9, P-10
Odd-chain fatty acids, 458–59
Oils:
  essential, 416
  properties of, 408
Okazaki fragments, 693–95, 697, 699
Okazaki, Reiji, 693–94
Oleic acid, 10*f*, 404*t*
Oleomargarine, 408
Oleoyl-CoA, 458*f*
Oligomers, 162
Oligonucleotides, 590
Oligosaccharides:
  asparagine-linked, 261, 268–69
  N-linked, 261, 776, 777*f*, 780
  O-linked, 261
  polypeptide-linked, 261
  role in biological recognition, 272*f*
O-linked oligosaccharides, 261
Omega-3 fatty acids, 405–6
Omega-6 fatty acids, 405
Omega-oxidation, 451
Oncogenes, 743, 743*t*
One-carbon metabolism, 543–49, 544*t*
Open reading frames (ORFs), 673, 718, 755
Open thermodynamic systems, 113–14
Operators, 732
Operons, 657, 732
Optical isomers, 141

Optimum temperature (enzymes), 214
Orbital hybridization, P-13–14
Orbitals, P-4, P-4*f*, P-13*f*
Order (enzyme kinetics), 219
Organelle DNA, 656
Organelles, 36
Organic chemistry, P-22–34
Organic reactions, 204–5
Organic solvents, 168
Orientation effects (enzyme catalysis), 207
Origin of replication complex (ORC), 697
Ornithine, 140, 140*f*
Orotic aciduria, 560
Oseltamivir, 227
Osmolarity, 90, 92
Osmolytes, 104
Osmometer, 90, 90*f*
Osmosis, 89
Osmotic diuresis, 615
Osmotic pressure, 89–91, 90*f*, 103*f*
Osteoclasts, 484
Osteoporosis, 484
Ouabain, 418
Outer membrane (Gram-negative bacteria), 42–43
Outer mitochondrial membrane, 59
Outer nuclear membrane (ONM), 56
Overwound DNA molecules, 649
Oxalate, 246
Oxaloacetate (OAA), 302–3, 541
  amino acids forming, 588
  malate oxidation forming, 348
Oxaloacetate/aspartate pair, 533
Oxidation, 252–53
  alpha-, 459, 460*f*
  alpha-ketoglutarate, 346
  beta- (*see* Beta-oxidation)
  fatty acid, 455–60, 467
  glucose, 380–83, 381*t*
  glyceraldehyde-3-phosphate, 288
  isocitrate, 345
  malate, 348
  omega-, 451
  succinate, 347–48
  water, 505–6
Oxidation-reduction (redox) reactions, 332–39, P-20
  coenzymes in, 335–36
  described, 18–19
  photosynthesis and, 502
Oxidation state (atoms), P-5
Oxidative damage, 363, 384
Oxidative deamination, 575, 576
Oxidative phosphorylation, 59, 374–84, 454
  chemiosmotic coupling theory of, 374–76, 375*f*
  complete oxidation of glucose, 380–83
  control of, 378–80
  defined, 364, 374

overview of, 385*f*
uncoupled electron transport and, 383–84
Oxidative stress, 39, 331, 384
Oxidizing agents, 18, P-20
Oxidoreductases, 201, 202*t*
Oxonium ions, P-25
Oxyanion hole, 216–18
Oxyanions, 211
Oxygen:
cell function and, 384–94
evolution and, 330
paradoxical nature of, 363
properties of, 364
waste product of photosynthesis, 505–06
Oxygen-evolving complex (OEC), 330, 497, 505, 506
Oxygenic photosynthesis, 330, 491*f*. *See also* Photosynthesis
5-Oxyproline, 550, 551*f*
Oxytocin:
amino acid sequence of, 148*t*
functions of, 150
structure of, 150

P

p53, 740, 743, 786
glycolysis and, 356–57
properties and structure of, 165–66
Pain, 196–97
Palindromes, 648, 709
Palmitic acid, 10*f*, 404*t*
Palmitoleic acid, 404*t*
Pantothenic acid, 214*t*
Paracrine signaling, 603
Parallel *β*-pleated sheets, 158
Paramagnetic atoms, P-4
Paraquat, 510
Parkinson's disease, 40, 384, 386, 570, 786
Pasha, 738
Passenger strand (miRNA), 738
Passive transport, 428
Passive transporters, 429
Pasteur effect, 297, 301, 356
Pasteur, Louis, 301
Pauli exclusion principle, P-4
Pauling, Linus, 145–47, 155, 646
P-block elements, P-8
P bodies, 730
P cluster, 529–31
Pentose phosphate pathway, 283, 309–13
nonoxidative phase, 309, 311*f*, 312
oxidative phase, 309–10, 310*f*, 313
Pentoses, 247
PEP. *See* Phosphoenolpyruvate
Pepsin, 135*f*
Peptide bonds:
defined, 7
formation of, 145–47, 146*f*–47*f*, 759, 761*f*, 765

Peptide motifs, 572
Peptides:
defined, 136
functions and structure of, 7, 148–50
signal, 766
Peptidoglycan, 42
Peptidyl transferase center (PTC), 762
Perinuclear space, 56
Periodic law, P-5
Periodic table, P-5–8, P-5*t*
Peripheral proteins, 37, 38*f*, 425–26, 425*f*
Periplasmic space, 43
Pernicious anemia, 545
Peroxiredoxins (PRX), 389
Peroxiredoxin/thioredoxin reductase, 389
Peroxisome proliferator-activated receptors (PPAR), 471
Peroxisomes, 61–62, 456
PEST sequences, 572
PFK. *See* Phosphofructokinase
pH, 93–97
acid-base catalysis and, 207–9
acid-base equilibria and, P-17–18
acids, bases, and, 93–94
buffers and, 94–97, 100–101
catalysis effects of, 214–15, 215*f*
of common fluids, 95*f*
hemoglobin and, 102
photosynthesis and, 517
Phagocytosis, 570
Phase contrast microscopy, 69–70
Phase II reactions (biotransformation), 485
Phase I reactions (biotransformation), 485
Phenolic antioxidants, 392
Phenylalanine, 138, 532*t*, 542, 589
abbreviations, 138*t*
catabolic pathway of, 583*f*, 584
conversion to tyrosine, 584*f*
ionization state, 137*f*
pK_a values, 142*t*
structural formula for, 7*f*
Phenylalanyl-t-RNA synthetase structure,135*f*
Phenylketonuria (PKU), 589
Pheochromocytoma, 353
Pheophytin a, 502, 505
pH optimum (enzymes), 215
Phosphate buffer, 102
Phosphate translocase, 378, 379*f*
Phosphatidic acid, 411*t*
Phosphatidylcholine, 10*f*, 410, 411*t*
Phosphatidylethanolamine, 411*t*
Phosphatidylglycerol, 411*t*
Phosphatidylinositol, 411*t*
Phosphatidylinositol cycle, 411, 609–10, 609*f*
Phosphatidylserine, 411*t*
Phosphocreatine, 379–80
Phosphodiesterases, 590

Phosphoenolpyruvate (PEP), 290–91
elimination reaction, 17*f*
synthesis of, 302–3
Phosphofructokinase-1 (PFK-1), 287, 295, 298–99
Phosphofructokinase-2 (PFK-2), 298–99
2-Phosphoglycerate elimination reaction, 15, 17*f*
Phosphoglycerides, 10, 410
Phospholipases:
properties and types of, 412–13
structure of, 412*f*
toxic, 413
Phospholipids:
in aqueous solutions, 410, 410*f*
important structural features of, 37
major classes of, 411*t*
metabolism of, 472–75
properties and types of, 409–11
synthesis of, 474*f*
Phospholipid turnover, 472
Phosphoprotein phosphatase-2A (PP-2A), 464, 464f, 471, 482f
Phosphoproteins, 153
Phosphoric acid:
dissociation constants and pK_a values for, 94*t*
titration with NaOH, 100
Phosphorylation, 778
fructose-6-phosphate, 287
overview of, 124
oxidative (*see* Oxidative phosphorylation)
photo, 509–10
substrate-level, 290, 346–47
Phosphoryl group transfer, 290
Phosphoryl group transfer potential, 125–26, 126*f*
Phosphoserine, 141*f*
5-Phospho-α-D-ribosyl-1-pyrophosphate. *See* PRPP
Photoautotrophs, 19
Photoheterotrophs, 19
Photomorphogenesis, 741
Photophosphorylation, 509–10
Photoreactivation repair, 699, 700*f*
Photorespiration, 514–15, 514*f*
Photosynthesis, 18–19, 62, 113, 283, 330, 339, 491–525
biomimetic, 520–21
chlorophyll in, 494–99, 501
chloroplasts in (*see* Chloroplasts)
climate change, renewable energy, and, 492
defined, 43, 493
light in, 500–503, 517–18
light-independent reactions in, 510–17
light reactions in, 503–10
nonoxygenic, 330
oxygenic, 491*f*
overview, 493*f*
pigment molecules used in, 494*f*
regulation of, 517–19
working units of, 496*f*
Photosynthetic carbon reduction (PCR) cycle, 510

Photosystem I (PSI), 496*f*, 502–5
components of, 493
functions of, 496–97
NADPH synthesis and, 507–9
Photosystem II (PSII), 330, 502–4
components of, 493
functions of, 497
light damage and, 506
monomer structure of, 497*f*
water oxidation and, 505–6
Photosystems, 493, 503*f*
pH scale, 94, 95*f*
Phylloquinone, 496
Physiological buffers, 101–2
Phytanic acid, 459, 460*f*
Phytanic acid storage syndrome, 459
Phytochrome, 518
Phytol, 416*t*, 459
Phytoplankton, 527
Phytosphingosine, 413*f*
Pili, 44
Ping-pong reactions. *See* Double-displacement reactions
Piwi-interacting RNAs (piRNAs), 674
pK_a values:
for ionizing groups of amino acids, 142*t*
for weak acids, 94*t*
PKA. *See* Protein kinase A
PKB. *See* Protein kinase B
PKC. *See* Protein kinase C
Planck's constant, 501
Plaque, 257, 395, 473
Plasma membranes, 782*f*
animal cell, 47*f*
eukaryotic, 45–48
prokaryotic, 43, 43*f*
Plasmids, 44, 658
Plastocyanin (PC), 498, 785
Plastoquinone (PQ), 417, 497, 507
Platelet-derived growth factor (PDGF), 67, 611
Pneumonia, 415
Point mutations, 642, 755
Polar amino acids, 137*f*, 138–39
Polar bonds, 79, P-12
Polar covalent bonds, P-9
Polar head groups, 409
Polar heads, 88, 89*f*
Polyamines, 548, 651
Poly(A)-binding protein (PABP), 769
Polycistronic mRNA, 673
Polycystic kidney disease, 67
Polymerase chain reaction (PCR), 671, 714–15, 714*f*
Polynucleotide phosphorylase, 767
Polypeptides:
defined, 136
homologous, 153
oligosaccharides linked to, 261

properties and functions of, 7
structure of, 9*f*, 153–66
Polypeptide Signal Molecules, 603–12
Polysaccharides, 8, 255, 261–66
Polysomes, 759, 774*f*
Poly(A) tail, 727
Polyterpenes, 416, 416*t*
Polyunsaturated fatty acids, 405
Porins, 43, 135*f*, 297
Positional isomers, P-24
Positive feedback, 25, 25*f*
Positively supercoiled DNA, 649, 650
Positive-sense RNA genomes, 677
Postabsorptive state, 619, 622*f*
Postprandial state, 619, 621*f*
Posttranslational modifications, 136, 271, 761
defined, 753
in eukaryotes, 776–79
in prokaryotes, 766
Posttranslational translocation, 780
Posttranslocation state, 765
PPAR, 471
PPARγ, 612
Prader–Willi syndrome, 626
Prednisone, 407
Pre-eclampsia, 660
Preinitiation complex (PIC), 725, 727*f*, 768
43S Preinitiation complex, 768, 769*f*
Preinitiation replication complex (preRC), 697, 697*f*
pre-mRNA, 727
Prenylation, 417, 417*f*
Prenyl groups, 416–17
Preproteins, 776
Preservatives, BHT as, 394
Pretranslocation state, 765
Pribnow box, 720
Primaquine, 156
Primary active transport, 429
Primary cilia, 64, 67
Primary miRNA (pri-miRNA), 738
Primary structure, protein, 153–56
Primase, 690
Primers, 690
Primer synthesis, 690
pri-miRNA, 738
Primosomes, 690
Principal quantum number, P-3
Processivity (DNA synthesis), 690
Proenzymes, 233
Progesterone, 419*f*
Progression (carcinogenesis), 744
Prokaryotes:
cell wall, 42–43
characteristics of, 36
chromosomes, 650
cytoplasm, 44, 44*f*

DNA recombination in, 707–8
DNA replication in, 692–95, 693*f*
DNA synthesis in, 689–92
flagella, 44
gene expression in, 657–58, 732–35
genomes of, 657–58, 657*f*, 736
pili, 44
plasma membrane, 43, 43*f*
protein synthesis in, 762–67
ribosomes of, 672–73
structure of, 41–44, 42*f*
transcription in, 719–23, 767, 767*f*
Proline, 136, 138, 532*t*, 537
abbreviations, 138*t*
biosynthesis from glutamate, 538*f*
catabolic pathway of, 585*f*, 586
hydroxylation of, 777–78, 778*f*
ionization state, 137*f*
pK_a values, 142*t*
Promoters, 658, 660, 720, 724–25
Proofreading site (aminoacyl-tRNA synthetases), 759
Prophage, 677
Propionyl-CoA, 459, 459*f*
Proproteins, 776
Prostaglandin E_2, 407*f*
Prostaglandins, 406
Prosthetic groups, 153
Protease inhibitors, 226
Proteasomes, 572–74, 574*f*
Protein buffer, 102
Protein families, 152
Protein folding, 40, 49, 160, 395, 777, 785–86
defined, 158
disease and, 786
energy landscape for, 169*f*
problem of, 168–72
schematic of, 170*f*
Protein kinase A (PKA), 607
Protein kinase B (PKB), 356, 612
Protein kinase C (PKC), 609
Protein microarrays, 787
Proteins, 134–36, 151–80
actin-binding, 87
acylated, 406
acyl carrier, 461, 463*f*
carrier, 37
channel, 37
client, 171
conjugated, 153
core, 266
defined, 136
diversity of, 135, 135*f*, 136
DNA-binding, 688–90, 688*f*, 694
docking, 780
envelope, 676
fatty acid-binding, 451

Fe, 529–31, 529*f*
fibrous, 152, 172–75
flavo-, 335–36
functions of, 6*t*, 7, 151–52
G, 605–7, 606*f*
globular, 152, 176–80
GTPase-activating, 741
half-lives of, 552*f*, 572
heat shock (*see* Heat shock proteins)
holo-, 153
integral (*see* Integral proteins)
intrinsically unstructured, 165
loss of structure, 167–68
matrix, 783, 783*f*
membrane, 425–26
metallo-, 153
methyl-CpG-binding, 662
microsomal triglyceride transfer, 444
microtubule-associated, 63
modified amino acids in, 140–41
modular (mosaic), 160
motor, 38
multifunction (moonlighting), 153
natively unfolded, 165
peripheral, 37, 38*f*, 425–26, 425*f*
phospho-, 153
poly(A)-binding, 769
primary structure of, 153–56
purification of, 184
quaternary structure of, 153, 162–64
secondary structure of, 153, 156–58
simple, 153
SREBP cleavage-activating, 483–84
SRP receptor, 780
sterol carrier, 478
stress response, 152
structure of, 153–68
supersecondary structure, 157–58
technology used in analysis, 184–89
telomere end-binding, 699
tertiary structure of, 153, 158–62
uncoupling, 383–84
unstructured, 165
Protein sequence-based function prediction, 188
Protein superfamilies, 152
Protein synthesis, 759–85
antibiotic inhibitors of, 767*t*
defined, 753
elongation phase, 759, 764–65, 771*f*, 772
eukaryotic, 768–85
initiation phase, 759, 763, 763*t*, 768–69, 769*f*, 771*f*
process of, 760*f*
prokaryotic, 762–67
termination phase, 759, 765–66, 773, 773*f*
Protein turnover, 571–75
autophagy-lysosomal system, 573*f*, 574–75

ubiquitin-proteasomal system, 572–74, 573*f*
Proteoforms, 166
Proteoglycans, 266–68, 268*f*
Proteolytic cleavage, 776
Proteomes, 39, 166, 637
Proteomics, 26, 753, 787–88
Proteosomes, 37, 672-74, 673f, 674f
Proteostasis, 39–40, 785-86
Proteostasis network (PN), 39–40, 785–86, 785*f*
Proteotoxic stress, 39–40
Proteus, 34*f*
Protocols (systems biology), 24
Protocruciforms, 648
Protomers, 162
Protonmotive force, 374
Proton shuttle mechanism, 764*f*, 765
Protooncogenes, 741, 743, 743*t*
Proximity effects (enzyme catalysis), 207
PRPP (5-phospho-α−D-ribosyl-1-pyrophosphate), 555, 557, 557*f*, 559
Pseudo-first-order reactions, 220
Pseudogenes, 659
Pseudomembranes, 774
Pseudopodium, 87
Pseudouridine, 674, 674*f*
Psoriasis, 548
PTENP1, 659
Pulsed-field electrophoresis (PFGE), 666
Purine nucleoside phosphorylase deficiency, 592
Purine nucleotide cycle, 590, 591*f*
Purines, 11, 554
biosynthesis of, 555–59
catabolism of, 590–92
most common naturally occurring, 555*f*
origin of ring atoms, 557*f*
structure of, 11*t*
Pyramids, Mayan, 78*f*
Pyran, 250, 251*f*
Pyranoses, 251
Pyranosides, 255
Pyridoxal phosphate, 213, 214*t*, 533–34, 535*f*
Pyridoxine. *See* Vitamin B_6
Pyrimidines, 11–12, 554
biosynthesis of, 559–60, 560*f*, 561*f*
catabolism of, 592
degradation of, 593*f*
most common naturally occurring, 555*f*
origin of ring atoms, 561*f*
structure of, 11*f*
Pyruvate:
conversion to acetyl-CoA, 341–42
glycolysis and, 283, 291–94, 294*f*
NADH reduction of, 332–33, 333*f*
synthesis of, 291
Pyruvate/alanine pair (transmission), 533–34
Pyruvate carboxylase deficiency, 349
Pyruvate dehydrogenase complex (PDHC), 341–42, 342*f*, 342*t*, 343*f*

Pyruvate family, 536*f*, 542
Pyruvate kinase (PK), 295, 298–99

Q

Q cycle, 369, 385
Quantum theory, P-3
Quantum tunneling, 207
Quaternary structure, protein, 153, 162–64
Quinine, 246

R

Racemization, 534
Rad 51, 709
Rad 52, 709
Radiation, 643–44, 744
Radiation damage, 391
Radiationless decay, 502
Radical chain reaction, 386, 387*f*
Radicals, 385. *See also* Free radicals
Radioactive decay, P-3
Radioactivity, P-3
Random sequential mechanisms, 225
Rapamycin, 575. *See also* mTOR
Ras proteins, 417, 741, 742f
Reaction center, 493
Reaction equilibrium, 198–99
Reaction kinetics, P-15–16
Reaction mechanism, 204–5
Reaction rate, P-15–16
Reactions, chemical
  products, P-14
  reactants, P-14, P-15–16
Reactive nitrogen species (RNS), 386
Reactive oxygen species (ROS), 331, 364, 384–86, 394, 395
  overview of formation, 385*f*
  properties and behavior of, 385–86
  respiratory burst and, 386, 387*f*
Readers (epigenetic), defined, 663
Receptor-mediated endocytosis. *See* Clathrin-dependent endocytosis
Receptors:
  G-protein-coupled-, 605–7, 606*f*
  guanylate cyclase, 610–11
  insulin, 269, 611, 611*f*
  LDL, 434–35
  membrane, 37, 434
  NMDA, 554
  peroxisome proliferator-activated, 471
  tyrosine kinase, 41
Receptor tyrosine kinases (RTKs), 611–12
Recombinant DNA technology, 181, 186, 712, 712*f*
Recombination, DNA, 660, 704–11
  defined, 689
  general (homologous), 700, 704–7, 705*f*
  site-specific, 704, 709–11
Red blood cells (erythrocytes):
  band 3 anion exchanger macrocomplex, 426–27

free energy changes during glycolysis, 294–95, 295*f*
integral proteins of, 426*f*
Redox code, 391–92
Redox proteome, 391–92
Redox reactions. *See* Oxidation-reduction reactions
overview of, 124
Reducing agents, 18, 168, P-7, P-20
Reducing ends, 262, 263*f*
Reducing sugars, 253
Reduction, 18, 253
Reductionism, 22
Reduction potential, 333–34
Reductive pentose phosphate (RPP) cycle, 510
Redundancy, 23
Refsum's disease, 459
Regulatory networks, 24
Relaxed state. *See* R state
Release factors, 765
Renaturation, 666, 667*f*
Renewable energy, 492
Renin, 150
Repair, DNA, 699–704
Reperfusion, 395
Repetitive DNA, 659–60
Replication, DNA, 636, 689–99
in eukaryotes, 697–99, 697*f*, 698*f*
in prokaryotes, 692–95, 693*f*
semiconservative, 689, 689*f*
Replication factor C (RFC), 699
Replication factories, 689
Replication fork, 693, 694, 694*f*, 698*f*
Replication licensing factors, 697
Replication protein A (RPA), 698, 698*f*
Replication slippage, 702
Replicative transposition, 710–11, 711*f*
Replicons, 693, 696–97
Replisomes, 690
Repolarization, 428
Repressors, 732
Resistance, torsional, 134
Resonance energy transfer, 502
Resonance hybrids, 127, 127*f*, P-10, P-26, P-30
Resonance stabilization, 127
Respirasome, 372
Respiration, 43
Respiratory burst, 386, 388*f*, 588
Respiratory control, 378
Respiratory distress syndrome, 412
Response elements, 13, 165, 299
Restriction enzymes, 667, 668*f*, 669
Restriction fragment polymorphisms (RFLPs), 671
Retinitis pigmentosa (RP), 67
Retinoblastoma, 740
Retroelements, 711
Retrograde neurotransmitters, 552
Retroposons, 711

Retrotransposons, 660
Retroviruses, 676
Reverse genetics, 675
Reverse transcriptase, 638
Reversible inhibition, 226
R groups (amino acids), 7, 136
R groups (organic molecules), P-28
Rheumatoid arthritis, 407, 548
Rhizobium, 528
Rho-dependent termination, 721, 722*f*
Rhodopsin, 135*f*
Rho-independent termination, 721
Riboflavin (Vitamin B_2), 214*t*, 335–36, 337*f*
Ribonuclease. *See* RNase
Ribonucleic acid. *See* RNA
Ribonucleotide reductase, 560–63, 562*f*
Ribonucleotides, 556*f*
Ribose, 8, 554–55
  diastereomers, 250*f*
  structure of, 9*f*, 11*f*
Ribose-5-phosphate, 283, 309
Ribosomal RNA. *See* rRNA
Ribosome-associated chaperones, 171
Ribosome recycling factor, 766
Ribosomes, 672–73, 753
  defined, 37, 44
  functional, 762*f*
  structure of, 751*f*
  trapped, rescue of, 770–71, 770*f*
Riboswitches, 732, 735, 735*f*
Ribozymes, 665
Ribulose-1,5-bisphosphate carboxylase. *See* Rubisco
Rieske-Fe-S protein, 498
Ring strain (cyclic hydrocarbons), P-25
RISC. *See* RNA-induced silencing complex
RNA, 271, 590, 636–37, 661–75, 688
  DNA compared with, 664–65
  functions of, 6*t*, 635, 661–65
  information flow, 637*f*, 638
  messenger (*see* mRNA)
  micro (*see* miRNA)
  noncoding (*see* ncRNA)
  ribosomal (*see* rRNA)
  secondary structure of, 665*f*
  small (short) interfering (*see* siRNA)
  small nuclear (*see* snRNA)
  small nucleolar (*see* snoRNA)
  structure of, 12, 664–65
  tm, 770–71
  transcription in (*see* Transcription)
  transfer (*see* tRNA)
  transport of, 739
  types of, 12–13, 637
RNA editing, 737–38
RNA-induced silencing complex (RISC), 673, 738
RNA interference (RNAi), 13, 674, 738

RNA polymerase (RNAP):
eukaryotic transcription and, 723–24
prokaryotic transcription and, 719–23, 721*f*
RNA polymerase I (RNAP I), 723
RNA polymerase II (RNAP II), 723–27, 724*f*, 725*f*
RNA polymerase III (RNAP III), 723
RNA processing:
in gene expression, 736–38, 737*f*
in transcription, 727–30
RNase, 269, 590
RNase II, 767
RNA splicing, 727–30, 729*f*
RNA transposons, 660
Robustness (systems biology), 23, 25
Rossman fold, 159
Rotenone, 373*f*
Rough endoplasmic reticulum (RER), 48, 48*f*, 49
rRNA, 12–13, 637, 665
processing reactions in *E. coli*, 722, 722*f*
structure of, 672–73, 672*f*
R state, 178–79, 234, 235
Rubber, 416*t*
Rubisco, 510–12, 514, 515, 517
carbamoylated active site, 519*f*
carboxylation mechanism, 512*f*
control of, 518–19
Ruv proteins, 707–8, 708*f*

S

Saccharomyces cerevisiae, 237, 237*f*, 296–97, 705, 717
Salicin, 256
Salivary amylase, 634–35
Salt bridges, 81, 160
Salt concentration, denaturation and, 168
Salting out, 168, 184
SAM. *See S*-Adenosylmethionine
Sanger, Frederick, 184, 669
Saponification, 409
Sarcoma, 743
Sarcomeres, 37
Sarcoplasmic reticulum (SR), 49
Satellite DNA, 659
Saturated fatty acids, 10, 404, 404*t*
Saturated hydrocarbons, P-23
Scanning electron microscopy (SEM), 69
Schiff base, 148, 256, 534, 534*f*, 535*f*
Scramblase, 423
Scurvy, 778
SDS-polyacrylamide gel electrophoresis (SDS-PAGE), 186–87
SECIS element, 775
Secondary active transport, 429
Secondary structure, protein, 153, 156–58
Second messengers, 604, 607*f*, 608*f*
Secretory process, 51
Selectin ligands, 273
Selectins, 269, 273

Selectively permeable membranes, 90, 90*f*, 424–25
Selenium, 391
Selenocysteine, 136, 775, 775*f*
Semiconservative replication, 689, 689*f*
Semi-essential amino acids, 532*t*
Sense strand (DNA), 675
Septic shock, 386
Sequential model (allosteric regulation), 234, 234*f*, 235
Sequential reactions, 225
Serine, 138–39, 532*t*, 540
  abbreviations, 138*t*
  biosynthesis of, 540*f*
  catabolic pathway of, 582, 582*f*
  ionization state, 137*f*
  pK_a values, 142*t*
  structural formula for, 7*f*
Serine dehydratases, 576
Serine family, 536*f*, 540
Serotonin, 140
Sesquiterpenes, 416, 416*t*
Shigella, 709
Shikimate pathway, 543
Shine–Dalgarno sequence, 763, 767, 768, 786
Shivering thermogenesis, 384
Short-chain fatty acids (SCFAs), 34
Short interspersed nuclear elements, 654, 660
Short tandem repeats (STR), 671, 671*f*, 672
Shotgun cloning, 715, 715*f*
Sickle-cell anemia, 155–56, 155*f*, 432
Sigma bonds, P-13
Signal hypothesis, 780
Signaling cascade, 40
Signaling networks, 24
Signal peptides, 766
Signal recognition particle (SPR), 780
Signal transduction:
  cytoskeleton role in, 66
  defined, 40
  gene expression and, 740–41
  mechanisms of, 40–41, 604*f*
  metabolism and, 41
  pathways, 21
  phospholipases and, 413
Silencers, 658, 724
Silent mutations, 5, 642
Silk:
  fibroin, 172, 174*f*
  spider (*see* Spider silk)
Simian virus 40, 135*f*
Simple diffusion, 428
Simple proteins, 153
SINEs (short interspersed nuclear elements), 654, 660
Singer–Nicholson model, 422
Single nucleotide polymorphisms (SNPs), 642, 671, 672
Single-strand DNA repairs, 700–702
Single-stranded DNA-binding protein, 694

siRNA, 13, 674, 738, 739*f*
Site-specific recombination, 704, 709–11
Skin, microbiota of, 34*f*
Slicer, 739
Slime layers, 43
Slipped strand mispairing, 643
Small dense LDL (sdLDL), 421
Small interfering RNA. *See* siRNA
Small ncRNA (sncRNA), 674
Small nuclear ribonucleoproteins (snRNPs), 674
Small nuclear RNA. *See* snRNA
Small nucleolar RNA. *See* snoRNA
Smooth endoplasmic reticulum (SER), 48, 48*f*, 49
SNAREs, 418–19
sncRNA, 673
snoRNA, 13, 674
S_N1 reactions, P-31, P-32*f*
S_N2 reactions, P-31, P-32
snRNA, 13, 674
snRNPs, 674
Soap, 409, 409*f*
Sodium channels, chemically gated, 428
Sodium-potassium (Na^+-K^+) ATPase transport, 429, 430*f*
Sol-gel transitions, 86–87
Solid state model (mitochondrial electron transport), 372–73
Solutes, 90
Solvation spheres, 85–86, 85*f*
Solvent properties of water, 85–92
Sorbitol, 253, 254*f*, 615
SOS response, 703–4
Southern blotting, 666, 667, 668*f*
Space-filling models:
  of adenylate kinase, 154*f*
  of cholesterol, 418*f*
  of *E. coli* chaperonin, 172*f*
  of fatty acids, 405*f*
  of monosaccharides, 251, 258*f*
  of sphingomyelin, 414*f*
  of triacylglycerols, 408*f*
  of water molecules, 80*f*
Special pair (photosynthesis), 496
Specific activity, 223
Specificity constant, 222
Speckles, 58
Spectrin, 426
Spermidine, 651
Spermine, 651
S phase (cell cycle), 696, 696*f*
Sphingolipidoses (sphingolipid storage diseases), 415
Sphingolipids, 413–15
  properties and types of, 413–14
Sphingomyelin, 410, 413, 414*f*
Sphingosine, 413*f*
Spider silk:
  assembly of, 181–83

biomimetics and, 181–83
diagrammatic view of, 182*f*
processing of, 182*f*
properties of, 134
spinning spigot of spider spinneret, 183*f*
structure of, 181
Spin quantum number, P-4
Spliceosomes, 674, 728, 730*f*
Spontaneous changes, 114, 117, 118
Spontaneous hydrolytic reactions, 643
Squalene, 416*t*, 476, 478, 479*f*, 480*f*
SREBP, 434, 471
SREBP-1, 299, 612
SREBP2, 483–84, 483*f*
SREBP cleavage-activating protein (SCAP), 483–84
SRP receptor protein, 780
Stahl, Franklin, 689
Standard amino acids, 136, 137*f*, 138*t*
Standard reduction potential, 333–34, 334*t*
Standard state (thermodynamics), 119
Staphylococcus, 34*f*
Starch, 261, 262
Start codons, 718, 754
Start signal, 780
State functions (thermodynamics), 114, 115
Statins, 484
Stationary phase, 184
Stator, 376
Steady state, 601
Steady state assumption, 221
Stearic acid, 404*t*, 405*f*
Stearoyl-CoA, 467*f*
Stereoisomers:
amino acid, 141
monosaccharide, 248
Steric hindrance, P-32
Steric repulsion, 39, 40*f*
Steroids, 415
model of within a target cell, 618*f*
properties and types of, 417–18
structure of selected, 419*f*
Sterol carrier protein, 478
Sterols, 418
Stokes, Alex, 646
Stomach ulcers, 272
Stomata, 515, 518
Stop codons, 718, 732, 754, 755
Stop transfer signal, 780
Streptococcus, 34*f*
Streptococcus mutans, 261
Streptococcus pneumoniae, 415
Streptokinase, 395
Streptomyces, 376
Stress response proteins, 152
Stroke, 386
Stroma, 62, 495

Stromal lamellae, 62, 495
Strong acids, 168
Strong bases, 102*f*, 168
Strong promoters, 720
Structured water, 86, 86*f*
Strychnine, 246
Submitochondrial particles, 376, 376*f*
Subshell-filling sequence, P-4, P-5*f*
Substituted hydrocarbons, P-22, P-26–31
Substitution reactions, P-19
  aliphatic, P-31–32
  electrophilic aromatic, P-26
  nucleophilic, 14–15, 16*f*
  nucleophilic acyl, P-28
Substrate-accelerated death, 302
Substrate cycle, 305
Substrate-level phosphorylation, 290, 346–47
Substrates, 200
  concentration of and catalysis rate, 234*f*
  defined, 198
  gluconeogenesis, 306
  multisubstrate reactions, 225–26
Subunits, 162–63
Succinate dehydrogenase, 336
Succinate dehydrogenase complex (complex II), 367–68, 372–73, 372*t*
Succinate oxidation, 347–48
Succinyl-CoA, 346–47
  amino acids forming, 586–88
  conversion of propionyl-CoA to, 459, 459*f*
Sucrose, 8, 260, 261*f*
Sugar code, 270–73
Sugar-phosphate backbone, 641
Sugars:
  amino, 259, 259*f*
  cells and, 245*f*
  deoxy, 259–60, 259*f*
  functions of, 6*t*, 8–9
  reducing, 253
  taste and, 246
Sulfatides, 414
Superbugs, 752
Supercoiling. *See* DNA, supercoiling
Supernatant, 68
Superoxide dismutase (SOD), 389, 643–44
Supersecondary motifs/structures, 157–58, 158*f*
Supraspliceosomes, 730, 730*f*
Surface active agents, 409
Surfactant, 412
Svedberg unit, 673
Sweet tastes, 246
Symbiotic bacteria, 528
Symmetry model. *See* Concerted model
Symporters, 378, 434
Synaptic plasticity, 554, 634
Synaptic vesicles, 551
Syn conformations, 555

Syndecans, 267, 267*f*
Synthesis-dependent strand annealing (SDSA) model, 707, 707*f*
Synthetic reactions, P-18
Synuclein gene, 744
Systematic names, 201
Systems, 24
  closed, 114
  dissipative, 123
  open, 113–14
Systems biology, 22–26

T

T_3 (triiodothyronine), 602, 603*f*, 604
T_4 (thyroxine), 140, 602, 603*f*, 604
Tamiflu, 227
Tandem mass spectrometry (MS/MS), 787
Tandem repeats, 659
Target cells, 601, 618*f*
Targeting, 753, 761, 779–84
Taste, 246
TATA box, 724, 727*f*
Taurine, 588
Taurine monochloramine, 588
Tautomeric shifts, 642*f*, 643
Tautomerization, 291
Tautomers, 291, 556*f*
Taut state. *See* T state
Taxol, 66
Tay–Sachs disease, 54, 268, 414*f*, 415, 415*t*, 660
T cells (T lymphocytes), 592
Telomerase, 699, 699*f*
Telomere end-binding proteins (TEBPs), 699
Telomere repeat-binding factors (TRFs), 699
Telomeres, 659, 699
Temperate phage, 677
Temperature:
  catalysis and, 214, 215*f*
  denaturation and, 168
  reaction rate and, P-16
Tensile strength, 134
Termination:
  protein synthesis (*see* Protein synthesis, termination phase)
  signal transduction, 40
  transcription (*see* Transcription, termination phase)
Terpenes:
  examples of, 416*t*
  properties and types of, 416–17
Tertiary structure, protein, 153, 158–62
Testosterone, 419*f*
Tetanus, 415
Tetragametic chimeras, 687
Tetrahydrobiopterin (BH_4), 213, 214*t*, 584*f*
Tetrahydrofolate (THF), 213, 545
  biosynthesis of, 545*f*
  pathway of, 548*f*

structures and enzymatic interconversions of coenzymes, 546*f*
Tetramers, 162
Tetraterpenes, 416, 416*t*
Thermal properties of water, 83–84
Thermodynamically favorable processes, 114
Thermodynamics, 113–18
defined, 113
first law of, 113–15
nonequilibrium, 123
second law of, 113, 116–18
third law of, 113
universe of, 113, 113*f*
Thermodynamic systems, 117
Thermogenesis, 383–84
Thermus aquaticus, 714
Thermus thermophilus, 366*f*
Thiamine (Vitamin B_1), 214*t*, 312
Thiamine pyrophosphate (TPP), 213, 214*t*, 341*t*, 342
Thiols:
properties of, P-30–31
structure and significance of, 6*t*
Thiolytic cleavage, 454
Thioredoxin (TRX), 389, 560
Thioredoxin-centered system, 389, 390*f*
Threonine, 138–39, 532*t*, 542
abbreviations, 138*t*
catabolic pathway of, 582*f*, 583
ionization state, 137*f*
pK_a values, 142*t*
Threonine dehydratases, 576
Thromboxane A_2, 407*f*
Thromboxanes, 406
Thylakoid lumen, 62, 495
Thylakoid membrane, 62, 495
Thymine, 11
structure of, 11*f*
tautomers of, 556*f*
Thymine dimers, 644*f*, 700*f*, 702*f*
Thyroid hormones, 602, 604
Thyroxine. *See* T_4
TIM complex, 783
Titration:
of acetic acid with NaOH, 96*f*
of dihydrogen phosphate, 102*f*
of phosphoric acid with NaOH, 100
tmRNA, 770–71
Tocopherol. *See* Vitamin E
TOM complex, 783
Topoisomerases, 649, 692
Topoisomers, 692
Topologically associated domains (TADs), 653–56, 655*f*
Topology, 692
Torque, 377, 649, 692
Torsional resistance, 134
Toughness, 134

Toxic phospholipases, 413
Toxins, protein, 152
TPS1 gene, 302
Transaldolase, 309, 312
Transaminases, 533
Transamination, 306, 533–34, 535*f*, 575
Transcription, 12, 636, 637*f*, 718–30, 754
  activation by activator proteins, 726*f*
  elongation phase, 719, 721, 725, 727
  in eukaryotes, 723–30
  initiation phase, 719–20, 725
  PIC assembly phase, 725
  in prokaryotes, 719–23, 767, 767*f*
  termination phase, 719, 721–22, 725
Transcription bubble, 721, 725, 727
Transcription coupled repair (TC-NER), 701
Transcription factories, 58, 723
Transcription factors, 13, 299, 471, 637
Transcription start site (TSS), 724
Transcript localization, 780
Transcriptomes, 636
Transcripts, 636
Transduction, 40, 677, 708. *See also* Signal transduction
Transfection, 713
Transferases, 201, 202*t*
Transfer RNA. *See* tRNA
Transformation, 708, 743
Trans-Golgi network (TGN), 50
Trans-isomers, 404
Transitional endoplasmic reticulum (tER), 49
Transition elements, P-7
Transition mutations, 642, 642*f*
Transition state, 204–5
Transition state analogs, 227
Transition state stabilization, 205–6
Transition state theory, 205
Transketolase, 309, 312
Translation, 637, 754, 767, 767*f*
Translational control, 739–40, 766–67, 784
Translesion polymerases, 703
Translocase of the mitochondrial inner membrane (TIM) complex, 783
Translocase of the mitochondrial outer membrane (TOM) complex, 783
Translocation, 643, 759, 765
Translocon, 780
Transmethylation acceptors and products, 547*t*
Transmission electron microscopy (TEM), 69
Transpeptidation, 759
Transposable elements, 659, 704
Transposition, 659, 704, 710–11, 711*f*
Transposons, 659–60, 709–10
Transsulfuration pathway, 586, 587*f*
Transversion mutations, 642
Traveler's disease, 611
Triacylglycerol cycle, 447–48, 447*f*
Triacylglycerols:
  biosynthesis of, 448, 449*f*

digestion and absorption of, 446*f*
hydrolysis of, 448–51
metabolism in adipocytes, 446–51
properties and types of, 408
in soapmaking, 409
structure of, 10, 10*f*, 408*f*
Trigger factor (TF), 766
Trigonal planar geometry, P-12
Trigonal pyramidal geometry, P-12
Triiodothyronine. *See* T_3
Triose phosphate isomerase (TPI), 205–6, 206*f*
Trioses, 247
Triple bonds, P-13, P-24
Triskelion, 52, 54*f*
Triterpenes, 416, 416*t*
tRNA, 12–13, 637, 665, 672
aminoacyl- (*see* Aminoacyl-tRNA synthetase)
initiator, 759
structure of, 665, 670*f*
Tropical sprue, 545
Tropomyosin, 736, 737*f*
Tryptophan, 138, 532*t*, 542
abbreviations, 138*t*
catabolic pathway of, 583*f*, 584
ionization state, 137*f*
pK_a values, 142*t*
T state, 178–79, 234, 235
Tuberculosis, 236, 570
Tubulin, 62
Tumor promoters, 614, 744
Tumor suppressor genes, 740, 743, 744
Turbo design dangers, 302
Turbo jet engines, metabolism compared with, 281
Turnover:
phospholipid, 472
protein (*see* Protein turnover)
Turnover number, 222
Tus, 695, 696*f*
Twist (DNA supercoiling), 649
Two-hybrid screening, 787
Type 1 diabetes, 615
Type 2 diabetes, 40, 615–17, 786
Type 3 diabetes, 617
Tyrosine, 138–39, 532*t*, 542, 589
abbreviations, 138*t*
catabolic pathway of, 583*f*, 584
ionization state, 137*f*
phenylalanine conversion to, 584*f*
pK_a values, 142*t*
Tyrosine kinase receptors, 41
TψC loop (tRNA), 665

U

Ubiquitin, 569
Ubiquitin-activating enzyme, 572

Ubiquitination, 572–74
Ubiquitin-conjugating enzyme, 572
Ubiquitin ligase, 572
Ubiquitin-proteasomal system (UPS), 40, 572–74, 573*f*
UDP-glucose, 316–17
Ulcers, stomach, 272
Ultracentrifuge, 68
Uncoated vesicles, 434
Uncompetitive inhibitors, 226, 228, 229, 229*f*
Uncouplers, 374
Uncoupling proteins, 383–84
Underwound DNA molecules, 649
Unfolded protein response (UPR), 40, 49, 395
Unimolecular reactions, 220
Uniporters, 434
Universal solvent, water as, 86
Unsaturated fatty acids, 10, 404–6
  examples of, 404*t*
  oxidation of, 458
Unsaturated hydrocarbons, P-23, P-24
Unstructured proteins, 165
Unwinding helicases, 690
Uracil, 11*f*
Urate oxidase, 592
Urea cycle, 22, 533, 577*f*
  control of, 580
  urea synthesis, 576–80
Ureotelic organisms, 571
Uric acid, 559, 571
  catabolism of, 590, 592, 592*f*
  hypertension and, 600
Uridine, 674, 674*f*
Urinary tract infections, 415
Urogenital tract, 34*f*
Uronic acids, 252, 259, 259*f*

V

Vagus nerve, 624
Valence bond theory, P-13–14
Valence electrons, P-4, P-9
Valence shell, P-5
Valence shell electron pair repulsion (VSEPR), P-11
Valine, 138, 532*t*, 542
  abbreviations, 138*t*
  catabolic pathway of, 586*f*, 587
  degradation of, 588*f*
  ionization state, 137*f*
  pK_a values, 142*t*
Vancomycin, 752
Van der Waals forces, 81*t*, 82, 88
Van't Hoff's rule, 248
Variable loops, 665
Variable substitutions, 154, 155
Vasopressin, 604
  amino acid sequence of, 148*t*
  functions of, 149–50

structure of, 150
Vectors (recombinant DNA), 712–13
Velocity, 218–19
Very-low-density lipoproteins (VLDL), 421, 422
Vesicles, 45
Viruses, 638, 675–77
cancer and, 744
endogenous retro, 660
properties and structures of, 676
retro, 676
Vitamin A, 214*t*, 392, 481
Vitamin B_1. *See* Thiamine
Vitamin B_2. *See* Riboflavin
Vitamin B_6 (pyridoxine), 214*t*, 534, 534*f*
Vitamin B_{12}, 214*t*, 545
Vitamin C. *See* Ascorbic acid
Vitamin D, 214*t*, 418, 481
Vitamin E (α-tocopherol), 392, 417, 481
coenzyme forms and process promoted, 214*t*
properties and functions of, 392
regeneration by ascorbate, 392, 393*f*
structure of, 393*f*
Vitamin K, 214*t*, 417, 417*f*, 779
Vitamin-like molecules, 213, 214*t*
Vitamins, 213, 214*t*
Volcanoes, 112
Voltage gated potassium (K^+) channels, 429
Voltage gated sodium (Na^+) channels, 428
Von Gierke's disease, 305

W

Warburg effect, 356
Warburg, Otto, 356
Waste removal, 21–22
Water, 76–109, P-12. *See also* Hydration
deforestation impact on, 77–78
flow through aquaporin channels, 430, 431*f*
ionization of, 92–102
melting and boiling points of, 83, 83*t*
molecular structure of, 79–80
noncovalent bonding in, 80–82, 81*t*
as percentage of body, 77
as percentage of cell, 5
photosystem II and oxidation of, 505–6
properties and importance of, 36–37
solvent properties of, 85–92
structured, 86, 86*f*
tetrahedral structure of, 79, 79*f*, 81*f*
thermal properties of, 83–84
Water-oxidizing clock, 506, 506*f*
Water-soluble hormones, 604
Water-soluble vitamins, 213
Water-splitting complex. *See* Oxygen-evolving complex
Watson–Crick base pairs, 756, 762
Watson, James, 635–36, 635*f*, 645–46

Waxes, 409
Wax esters, 409
Weak acids, 93, P-17–18, P-27, P-28
dissociation constants and pK_a values for, 94*t*
with more than one ionizable group, 100–101
Weak bases, 93, P-17–18
Weak noncovalent interactions, 37, 38*f*
Weak promoters, 720
Wilkins, Maurice, 646
Wilson's disease, 212
Wine, 237, 280, 280*f*
Wnt signaling pathway, 67
Wobble hypothesis, 756–57, 757*f*
Work, energy as capacity for, 112, 114
World Health Organization, 402
Writers (epigenetic), defined, 663
Writhe (DNA supercoiling), 649
WRN gene, 744

X

Xanthomas, 435
Xanthphylls, 416
Xenobiotics, 644–45
Xeroderma pigmentosum, 701
X-ray crystallography, 188, 189*f*
X-ray diffraction studies, 646, 646*f*

Y

Yeast, 237, 416, 725*f*
Yeast artificial chromosomes (YAC), 713
Yeast two-hybrid screening, 787–88
Yucatan peninsula, 77–78

Z

Z-DNA, 647–48, 647*f*, 647*t*
Zea mays, 651
α / β Zinc-binding motif, 159*f*
Z scheme, 504*f*, 505
Zwitterions, 136
Zymogens, 233