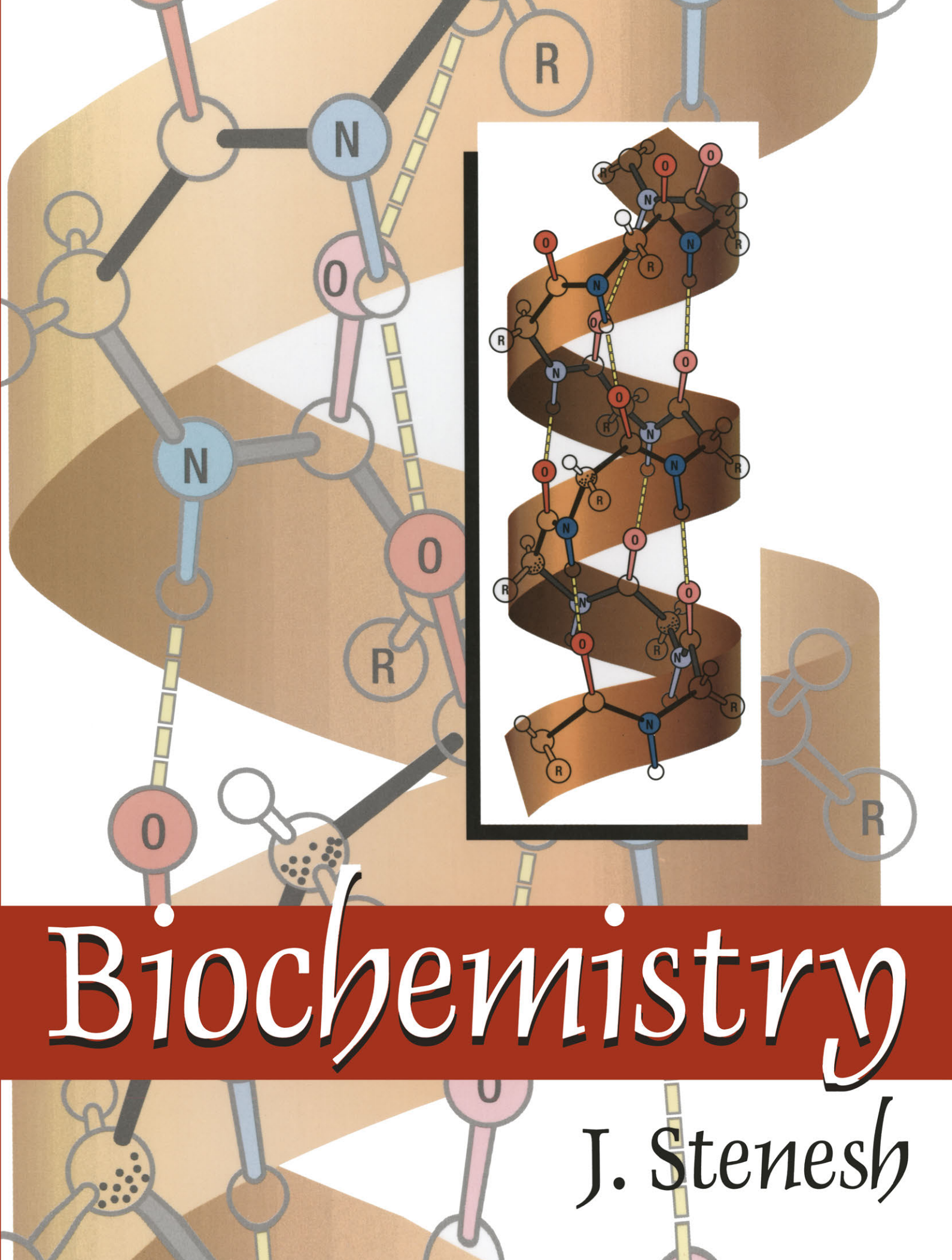


Biochemistry

J. Stenesh

Biochemistry

Biochemistry

J. Stenesh

Western Michigan University
Kalamazoo, Michigan

SPRINGER SCIENCE+BUSINESS MEDIA, LLC

Library of Congress Cataloging-in-Publication Data

Stenesh, Jochanan.
Biochemistry / Jochanan Stenesh.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4757-9429-8 ISBN 978-1-4757-9427-4 (eBook)
DOI 10.1007/978-1-4757-9427-4
1. Biochemistry. I. Title.
QP514.2.S635 1998
572--dc21 97-41014
CIP

ISBN 978-1-4757-9429-8

Originally published by Plenum Press, New York in 1998
Softcover reprint of the hardcover 1st edition 1998

http://www.plenum.com

10 9 8 7 6 5 4 3 2 1

Preface

This text is intended for an introductory course in biochemistry. While such a course draws students from various curricula, all students are presumed to have had at least general chemistry and one semester of organic chemistry.

My main goal in writing this book was to provide students with a basic body of biochemical knowledge and a thorough exposition of fundamental biochemical concepts, including full definitions of key terms. My aim has been to present this material in a reasonably balanced form by neither deluging central topics with excessive detail nor slighting secondary topics by extreme brevity.

Every author of an introductory text struggles with the problem of what to include in the coverage. My guideline has been to make sure first that the essentials are covered in sufficient depth to give students a firm basis on which to build further. Beyond that, material is covered to varying extents. More tangential material is frequently collected in the final sections of a chapter so that it can be omitted at the discretion of the instructor.

Following an introduction, which outlines the scope of biochemistry, the book is organized into four parts along traditional lines. Part I, FOUNDATION OF BIOCHEMISTRY, covers four general "frameworks" of biochemistry—the origin of life, the living cell, water, and noncovalent interactions. Part II, BIOMOLECULES, surveys structures and properties of the different classes of molecules that occur in living systems. Part III, METABOLISM, opens with an introductory chapter, followed by a study of the essential reactions, interconversions, and pathways of biomolecules. For cohesiveness, anabolism (synthesis) and catabolism (degradation) of a given class of biomolecules are covered in a single chapter. The discussion of metabolism concludes with photosynthesis. The last section of the book, Part IV, TRANSFER OF GENETIC INFORMATION, also opens with an introductory chapter and then explores the expression of genetic information. Replication, transcription, and translation are covered in this order. To allow for varying student backgrounds and for possible needed refreshers, a number of topics are included as four appendixes. These cover acid–base calculations, principles of organic chemistry, tools of biochemistry, and oxidation–reduction reactions.

Each chapter includes a summary, a list of selected readings, and a comprehensive study section that consists of three types of review questions and a large number of problems. Asterisks mark more difficult problems, and answers to all problems are given.

A solutions manual, providing step-by-step solutions of the problems, is available separately, as is a set of overhead transparencies.

The text has been extensively reviewed to eliminate all errors in fact. I thank the following reviewers who critically read portions of the manuscript and made valuable comments:

Richard Amasino, University of Wisconsin at Madison
Gary D. Anderson, Marshall University
Linda Atwood, California Polytechnic State University at San Luis Obispo
Katherine A. Bichler, Carthage College
Veronica C. Blasquez, University of Notre Dame
Ronald W. Brosemer, Washington State University
Larry Byers, Tulane University
Stephen W. Carper, University of Nevada

A. J. Clark, Auburn University
Rick H. Cote, University of New Hampshire
Richard C. Crain, University of Connecticut
Dorothy E. Croall, University of Maine
John Cronin, Arizona State University
John R. Edwards, Villanova University
Jeffrey A. Evans, University of Southern Mississippi
William Fordham, Fairleigh Dickinson University
Edward A. Funkhouser, Texas A & M University
Mina L. Hector, California State University at Chico
John L. Hess, Virginia Polytechnic Institute
Ken Hilt, University of California at Davis
Mahendra Jain, University of Delaware
Jerry P. Jasinski, Keene State College
Marjorie A. Jones, Illinois State University
Ivan I. Kaiser, University of Wyoming
Michael R. Kanost, Kansas State University
Herman W. Knoche, University of Nebraska
Roger E. Koeppe II, University of Arkansas
Glenn D. Kuehn, New Mexico State University
Michael Mendenhall, University of Kentucky
Robert Oberlender, University of the Pacific
Robert S. Orr, Delaware Valley College
R. Steven Pappas, Georgia State University
Samuel J. Rogers, Montana State University
D. L. Rohlfing, University of South Carolina
Laura Silberstein, San Jose State University
Thomas W. Sneider, Colorado State University
Ming Tien, Pennsylvania State University
Anthony P. Toste, Southwest Missouri State University
Martin Tuck, Ohio University
Ron Utecht, South Dakota State University
Thomas G. Waddell, University of Tennessee
Edward B. Walker, Weber State University
Daniel Weeks, University of Iowa
James H. Yuan, Old Dominion University

I am very grateful to John Dunigan who was responsible for producing most of the artwork. The excellent quality of the illustrations are a reflection of his tireless effort and careful attention to detail. Last, but not least, I want to express my sincere appreciation to the individuals at Plenum Publishing Corporation involved in this project: Amelia McNamara, Ken Howell, and Jacqueline Sedman, my editors, for their guidance and input, and Mary Curioli and Robert Maged for their supervision of the book's production.

J. Stenesh

Contents

Introduction
The Scope of Biochemistry

Biochemistry—the chemistry of life—deals with the chemical and physical properties of molecules and processes of living organisms. This relatively new science had its origin in the 1700s, emerged as an independent scientific discipline at the beginning of the 20th century, and erupted into unprecedented growth some 50 years ago.

For a long time, many people believed that reactions in living organisms (as distinct from those in nonliving systems) required a special "vital" force. Only when this theory of "vitalism" was discarded could development of biochemistry proceed. One of the earliest scientists whose work helped bring about the downfall of vitalism was Antoine Lavoisier. In 1777, Lavoisier conducted experiments on respiration and combustion and showed that both processes converted organic matter to carbon dioxide and water. He concluded that cellular respiration was slower than combustion but not essentially different from it.

Half a century later, in 1828, Friedrich Wöhler succeeded in synthesizing urea by heating ammonium cyanate in the laboratory. Until then, scientists had assumed that urea, like other organic compounds of living matter, could be synthesized only by and in a living organism. Two other important developments followed in short succession. In 1838, Matthias Schleiden and Theodor Schwann proposed that a membrane-bound structure—the cell—is the fundamental unit of all living organisms. In 1862, Louis Pasteur proved that living organisms arise only from other living organisms and not by "spontaneous generation." Pasteur showed that microorganisms did not form in a sterilized solution of organic matter unless that solution was exposed to air and other microorganisms.

Vitalism was finally rejected as a scientific theory when Eduard Buchner (1896) obtained a cell-free extract from yeast that was capable of carrying out fermentation and when J. B. Sumner (1926) crystallized the enzyme urease from jack beans.

The model of the double helix of DNA, proposed by James Watson and Francis Crick in 1953, revolutionized both biochemistry and biology. The proposal opened up countless avenues for studying life sciences at the molecular level. In particular, the concept of the double helix led to an ever-increasing interweaving of the sciences of biochemistry, cell biology, and genetics. This resulted in formation of a separate discipline called molecular genetics or molecular biology, which has mushroomed into a very active research area with major implications for medicine and other fields.

In the course of its development, at least four central themes have come to characterize biochemistry.

1. *Reactions carried out by living organisms obey the laws of chemistry and physics that describe reactions in the laboratory.* Physical properties of molecules, chemical reaction mechanisms of inorganic and organic compounds, energy relationships of products and reactants—all these apply to biochemical reactions proceeding *in vivo* in precisely the same manner as they apply to *in vitro* reactions carried out in the laboratory. No special forces such as "vitalism," or special processes like "spontaneous generation," play a role in the synthesis, degradation, and interconversions of compounds found in living cells.

2. *Structure and function are interdependent.* Molecules and larger aggregates have particular structures that permit them to perform specific functions. Conversely, in order to be capable of performing specific functions, components must possess particular elements of structure. The structure–function interdependence exists at all levels of

organization. It occurs in low-molecular-weight compounds such as amino acids and fatty acids, in polymers such as proteins and nucleic acids, in supramolecular assemblies such as biological membranes, and in subcellular organelles such as mitochondria. Investigators have achieved many important breakthroughs in biochemistry, especially in molecular genetics, by focusing on the interdependence of structure and function.

3. *The cell is the basic unit of life.* The cell is the fundamental unit of living matter in both single-cell and multicellular organisms. Physical properties, the organization of chemical components, transport of substances in and out, and energy requirements and exchanges—all these play key roles in the functioning of a cell. Living cells are composed of lifeless molecules, but assembly, organization, and interactions of these molecules endow the cell with life. It is the goal of biochemistry to understand how cells accomplish this task.

Living cells exhibit an almost universal hierarchy of molecular organization (Figure 1). All the molecules that occur in living organisms (biomolecules) are formed from a small number of elements, primarily carbon, hydrogen, oxygen, nitrogen, phosphorus, and sulfur. Ultimately, all biomolecules are derived from low-molecular-weight precursors in the environment that contain these and several other elements. Precursors are converted by living matter via metabolic intermediates, or metabolites, into building blocks that become linked covalently to form macromolecules. Building blocks and/or macromolecules associate to form supramolecular assemblies and organelles.

4. *Living organisms exhibit both unity and diversity.* One of the outstanding phenomena of life is the existence of extensive biochemical similarities among diverse cells and organisms, despite significant morphological differences among them. This characteristic, known as the "principle of unity and diversity," describes a stunning paradox: The immense diversity of life at the cellular and organismal level is ultimately reducible to a surprising unity at the molecular level. The unity becomes evident not only in overall organization of cells (Figure 1), but also in properties of cellular components.

Major classes of biomolecules have identical functions in all types of cells. Nucleic acids store genetic information; proteins serve as structural components and as catalysts of metabolic reactions and have many other functions; lipids serve as a storage form of energy and as components of cell membranes; and carbohydrates represent a storage form of energy and, in prokaryotes, serve as components of the cell wall.

The biochemical unity of life, however, extends well beyond similarities among biomolecules. Proteins with the same function in different organisms have similar amino acid sequences. DNAs of closely related bacteria have similar molecular structures. Many of the same metabolic reactions and reaction sequences occur in a wide range of organisms. Fundamental cellular processes such as using nutrient energy, synthesizing proteins, respiring, transporting substances across cell membranes, and replicating genetic material involve identical or very similar mechanisms and components. Likewise, the genetic code used in the genes of nucleic acids is essentially universal. Based on such findings, most scientists today believe that all forms of life arose from a common ancestor and that, in the course of evolution, species diverged much as branches diverge from the trunk of a tree (see Section 1.1.2).

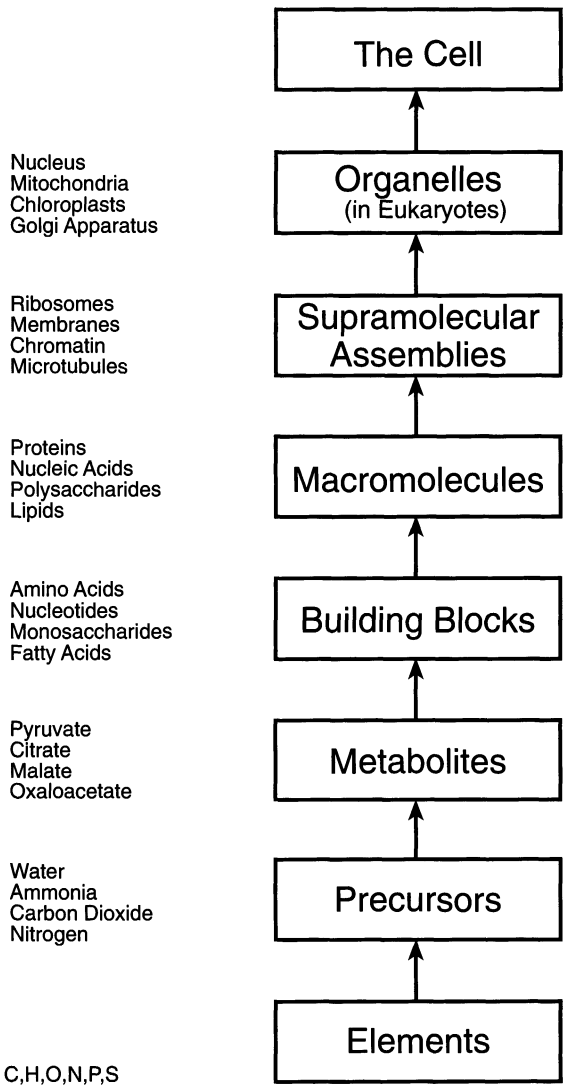


Figure 1. Hierarchy in the molecular organization of cells.

The biochemical unity of life makes it possible to apply information derived from one organism to the study of a different organism. A large body of biochemical knowledge comes from studies of prokaryotic microorganisms and has been found to apply, either directly or with some modifications, to eukaryotes, including humans. Similarly, researchers routinely use studies with laboratory animals to screen for potential hazards that drugs or toxic substances may pose for humans.

While many of life's mysteries still need to be unraveled, a few general characteristics consistently distinguish living organisms from nonliving matter.

All living organisms are complex and highly organized systems. Even the simplest unicellular organism contains many different components and constitutes a marvel of organization. Any single cell incorporates simultaneous control of multiple metabolic pathways and hundreds of different chemical reactions. Multicellular organisms have an even greater complexity of organization, and in higher animals and humans, the number of interrelationships and control systems must be truly staggering.

All living organisms contain many different kinds of biomolecules. Biomolecules can be small or large, simple or complex. They exist singly or form aggregates of varying sizes and intricacies. All forms of biomolecules have unique and specific functions.

Despite the occurrence of numerous types of biomolecules, living systems exhibit an underlying molecular economy. The complexity of biomolecules appears to be no greater than that required for molecular function. Likewise, the number of different types of biomolecules appears to be no greater than that needed to endow the cell with attributes of life in general and characteristics of the species in particular.

All living organisms require enzymes as catalysts for their metabolic reactions. Enzymes are protein molecules that catalyze most of the reactions of metabolism. These biomolecules are specially engineered for their catalytic function; they far exceed any human-made catalyst in both their selectivity and their efficiency. Life processes would not be feasible without the catalytic action of enzymes.

Enzyme reactions involve both changes of biomolecules and transformations of energy. Many of the hundreds of different enzyme reactions taking place within a single cell are linked into specific sequences of anywhere from 2 to over 20 steps. Intricate control systems regulate the resulting network of metabolic pathways. Enzyme reactions form the basis of the numerous processes characteristic of living organisms, such as vision, growth, nerve-impulse conduction, muscle contraction, and reproduction.

All living organisms require a supply of energy. The Sun is the ultimate source of energy for all life on Earth. Plants harness solar energy during photosynthesis and use it to synthesize carbohydrates. Many animals subsequently obtain their nutrients by feeding on plants. All living organisms extract energy from nutrients in the form of free energy and use it to drive life processes. Organisms return unused energy to the environment in the form of heat. Accumulating waste energy increases the disorder or entropy of the environment. Thus, living organisms maintain their organization at the expense of the environment, which they cause to become more random.

All living organisms have hereditary information encoded in genes. Life processes depend on the information contained within genes, which are structural components of nucleic acids. Because of the properties of their genetic material, living organisms can self-replicate and thereby transfer their genetic information to the next generation. Production of organisms identical in mass, shape, and internal structure constitutes the most extraordinary attribute of living systems.

Based on these characteristics of living matter, biochemistry can be divided into three main areas that deal with respectively, the chemistry of biomolecules, a study of metabolism, and molecular genetics. Biochemistry draws heavily on other life sciences and, in turn, provides key insights for related physical and biological disciplines. Chemistry, biology, nutrition, medicine, microbiology, physiology, agriculture, and biophysics are some of the sciences closely linked with biochemistry.

Foundation of Biochemistry

I

We begin our study of biochemistry by considering four major frameworks within which biochemical systems operate: the origin of life—the postulated chain of events leading to primordial life; the cell—the complex unit of all living organisms; water—the unusual solvent that makes life possible; and noncovalent interactions—the forces that govern many aspects of biochemical reactions.

Frameworks of Biochemistry

1

We can consider biochemical systems to operate within four major frameworks. One of these is the origin of life—the chain of events leading to the first ancient cell from which all forms of life ultimately developed. A second framework deals with the properties of the modern cell, the basic unit of all living organisms and the site of most biochemical reactions. Water constitutes a third framework. This unusual solvent provides the internal aqueous environment of cells without which life, as we know it, would not be possible. Lastly, noncovalent interactions form an essential framework. Biochemical systems require these interactions for stabilization of proteins, nucleic acids, cell membranes, and the like. Noncovalent interactions also participate in the operation of many biochemical systems by affecting the interplay among low-molecular-weight compounds, macromolecules, and other structures.

1.1. THE ORIGIN OF LIFE

1.1.1. Chemical Evolution

Life is believed to have arisen as a result of a long period of **chemical evolution** that consisted of a multitude of simple and complex chemical reactions. Chemical evolution took place over almost one billion years and included the many different processes whereby molecules were formed and altered and reactions arose, underwent change, disappeared, or became established.

These processes are termed "chemical evolution" because scientists believe that they can be described by attributes analogous to those that characterize "biological evolution," namely, *mutation* and *selection.* Investigators think that a very large number of changes occurred at the molecular level, ultimately producing compounds and reactions capable of sustaining life. Such changes constitute the chemical equivalent of biological mutations in living organisms. In the long succession of events, compounds and reactions that, for whatever reason, were more suitable or useful tended to persist. The retention of compounds and reactions constitutes the chemical equivalent of biological selection of one organism or one species over another.

Chemical evolution began some time after formation of the Earth. Scientists consider the Earth to have formed about five billion years ago as a result of a cataclysmic explosion known as the "big bang." According to this theory, the matter of the universe was originally very dense, a large amount of mass contained in a comparatively small volume. This primordial mass exploded with tremendous force, generating extremely high temperatures and a vast expansion. Ever since, the universe has continued to expand and its average temperature has progressively decreased.

The original big bang resulted in formation of the light elements hydrogen, helium, and lithium. Researchers believe that the remaining chemical elements formed by thermonuclear reactions that occurred in stars, by star explosions called supernovas, and by cosmic radiation acting outside stars following the formation of galaxies. Biologically important elements such as carbon, oxygen, nitrogen, phosphorus, and sulfur formed by thermonuclear reactions that occurred in some of the original stars produced after the big bang.

Chemical evolution commenced when some cooling of the Earth occurred, and it continued for the next billion years (Figure 1.1). We commonly divide chemical evolution into three stages called *synthesis, polymerization,* and *self-assembly.*

1.1.1A. Synthesis Stage. Chemical evolution began with the synthesis of low-molecular-weight compounds from primordial gases. Scientists believe that these early **abiotic reactions** ("in the absence of life") occurred in both the primordial atmosphere and the primordial oceans. Synthetic reactions probably involved one or more energy sources such as electric discharge, ultraviolet radiation, and volcanic activity.

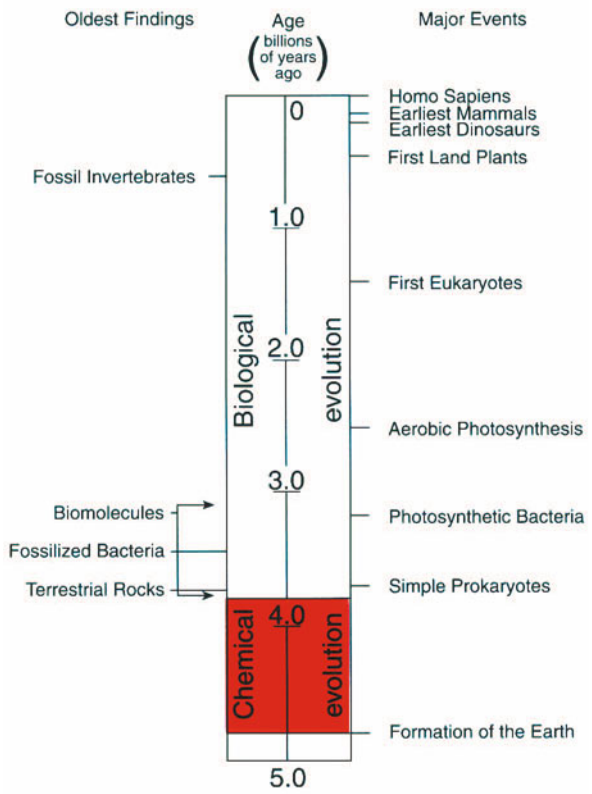


Figure 1.1. The time scale of chemical and biological evolution.

For many years, researchers believed that the primordial atmosphere was largely a reducing one, composed of gases like NH_3, CH_4, H_2, and H_2O. Recent evidence has led us to reevaluate this assumption. Studies of ancient rocks, which contain carbonates, and of volcanic gases, which contain CO_2, have indicated that significant amounts of CO_2 must have existed in the atmosphere at least 3.8 billion years ago. Hence, many investigators currently postulate that the primordial atmosphere was nonreducing, containing such gases as N_2, CO, CO_2, and H_2O.

Numerous laboratory experiments have been conducted to simulate conditions presumed to have existed on the primitive Earth. In these experiments, primordial atmospheres and oceans were simulated by gas mixtures and water, respectively. The temperature was varied, and different energy sources (electric discharge, ultraviolet radiation, etc.) were used to drive the synthetic reactions. Early experiments used primarily reducing atmospheres, whereas later ones used nonreducing atmospheres. These simulations resulted in formation of a great diversity of organic compounds, including many types of low-molecular-weight biomolecules. Amino acids, fatty acids, carbohydrates, and other compounds formed spontaneously under some sets of experimental conditions. Based on experiments of this type, most researchers believe that many different low-molecular-weight biomolecules could have been formed spontaneously under specific primordial conditions.

1.1.1B. Polymerization Stage. In the second stage of chemical evolution, macromolecules formed from the low-molecular-weight compounds produced during the synthesis stage. Low-molecular-weight compounds could have accumulated in primitive oceans or at their edges. Accumulation in small bodies of water, like tide pools, may have resulted in a moderately concentrated aqueous solution, dubbed "primordial soup." Polymerization reactions could have taken place in aqueous solution, possibly via mechanisms resembling those involved in current laboratory procedures.

Some scientists have postulated a thermal polymerization mechanism that occurred outside the aqueous environment. They simulated this process by heating mixtures of dry amino acids and produced protein-like polymers called **proteinoids.** Proteinoids have many properties commonly associated with naturally occurring proteins such as high molecular weight, dependence of solubility on ionic strength, presence of peptide bonds, and possession of catalytic activity.

Other scientists have proposed that polymerization took place on *clay particles* because mineral substances

similar to clay can serve as catalysts for some chemical reactions. According to this theory, patterns of ions on the clay surface served as a code for polymerization of biomolecules, and growth of clay crystals resulted in replication.

1.1.1C. Self-Assembly Stage. During the final stage of chemical evolution, polymeric molecules aggregated to form supramolecular structures that ultimately evolved into forerunners of modern cells called **protocells.** A number of cell-like structures have been considered as possible intermediates in the development of protocells.

The first model system proposed was that of **coacervate droplets.** Coacervation constitutes a phase separation that occurs when a solvent contains two polymers that interact with the solvent but do not interact well with each other. After mixing, one polymer accumulates primarily in one phase while the other polymer concentrates mainly in the second phase. Proponents of this model regard the coacervate droplet as a space where simple chemical reactions could have taken place, ultimately leading to development of a protocell.

A more recently proposed model system is that of **microspheres.** Microspheres are small spherical structures that form upon addition of water to hot proteinoid. Microspheres have many of the properties commonly associated with living cells. Their size compares to that of bacteria, and they can be stained by the Gram stain used in microbiology. Additionally, microspheres exhibit stability to centrifugation and to changes in osmotic pressure and have a boundary that appears to be a double layer.

Yet another model system of protocells is that of **liposomes,** artificial vesicles formed from some lipids under appropriate conditions. Liposomes have a lipid bilayer membrane that encloses an aqueous compartment. The bilayer resembles the lipid bilayer of biological membranes. Liposomes can exchange material with their surroundings.

Clay particles may also have served in the capacity of protocells. According to this view, proteins synthesized on clay particles subsequently aggregated to form microsphere-like structures.

A debate has raged for years as to whether proteins and protein-containing protocells or nucleic acids and nucleic acid-containing protocells arose first. Which came first, the catalytic activity of enzymes, associated with proteins, or the ability to transmit genetic information by coding, associated with nucleic acids? Or did the two types of molecules and processes arise more or less simultaneously, requiring some type of *double-origin theory?*

Several experimental findings tend to support the hypothesis that nucleic acids arose first. Of special note is the finding in the mid-1980s that some *RNA (ribonucleic acid)* molecules can catalyze their own intramolecular transformations and can also catalyze reactions with other substrates. Up until then, scientists had assumed that all biological catalysts were proteins, namely, *enzymes.* In recognition of their catalytic activity, such RNA molecules are called **ribozymes.** The discovery of ribozymes suggests that an RNA-containing protocell may have been able to carry out both catalytic and coding functions. Researchers generally believe that RNA formation preceded the development of *DNA (deoxyribonucleic acid).* So far, however, the question as to whether proteins or nucleic acids came first, or whether they were formed simultaneously, has not been settled in any definitive way.

1.1.2. Biological Evolution

Once protocells were formed, they became subject to the forces of **biological evolution.** Simple cells evolved into more complex unicellular organisms, and these led to multicellular organisms. Concomitantly, the many biochemical processes characteristic of living systems became established. Metabolic pathways, respiration, photosynthesis, energy metabolism, and the like originated and underwent change. The continuous journey along the evolutionary path stretched over a period of some 3.5–4.0 billion years and, ultimately, led to the development of all current forms of life.

Organisms can be classified into five kingdoms (Figure 1.2), based on cellular complexity: *Monera* (unicellular prokaryotes), *Protista* (unicellular eukaryotes), and *Fungi, Plants,* and *Animals* (all multicellular eukaryotes).

1.1.2A. Endosymbiotic Theory. An interesting theory about the evolution of eukaryotes has grown out of accumulated biochemical knowledge. We now know that **mitochondria** and **chloroplasts** (organelles of eukaryotic cells) contain their own genetic material in the form of specific DNA. This DNA differs from the DNA in the cell nucleus and functions in the synthesis of some RNAs and proteins. Mitochondria and chloroplasts also contain **ribosomes,** the sites of protein synthesis, which resemble ribosomes of prokaryotes more than they do those of eukaryotes. Similarities include size, sensitivity to antibiotics, and nucleotide sequences of ribosomal RNA. On the basis of this evidence, theorists have suggested that mitochondria and chloroplasts represent evolutionary descendants of ancient prokaryotes that became incorporated into primitive eukaryotic cells. Specifically, scientists have proposed that *cyanobacteria* took up resi-

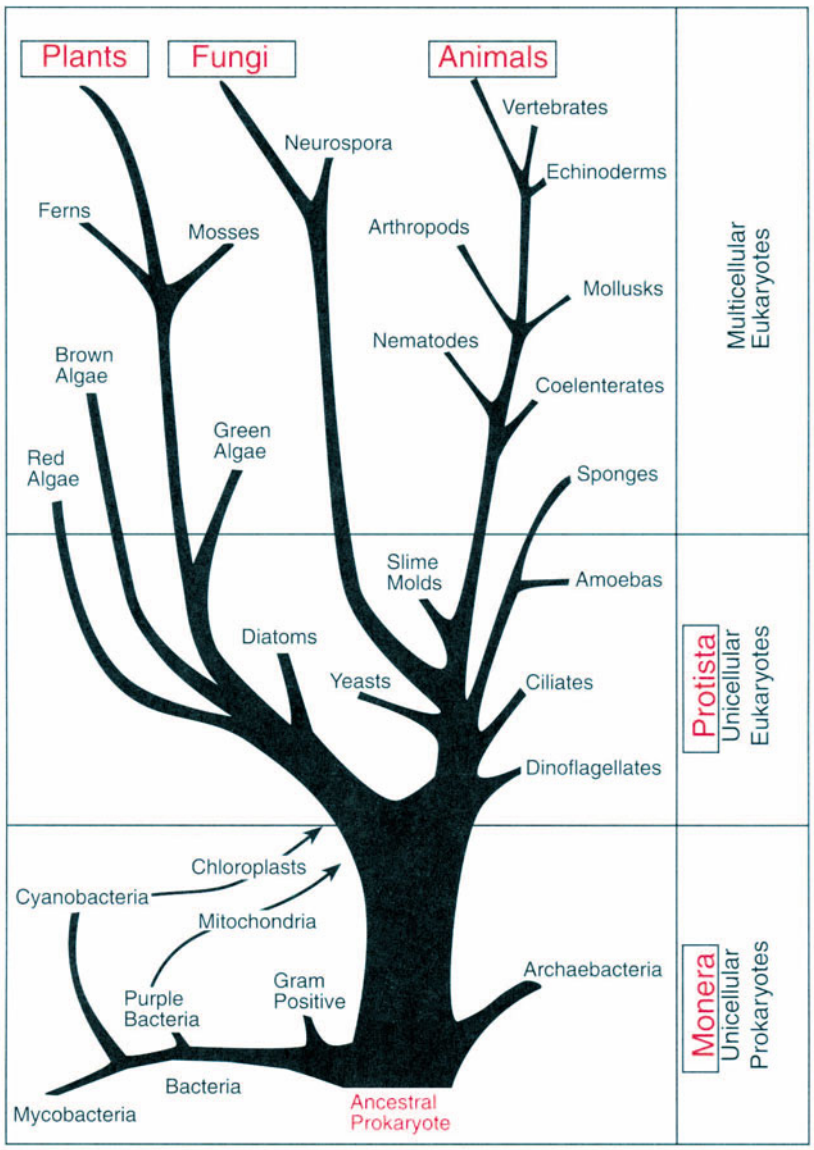


Figure 1.2. An evolutionary tree showing lines of descent from a primordial prokaryote to the five kingdoms of organisms.

dence in ancestral nonphotosynthetic eukaryotes. We term such a relationship **endosymbiosis,** a living together in close association. The symbiotic relationship would have been beneficial to both the eukaryote and the prokaryote, with each contributing metabolically to the other. Cyanobacteria form a heterogeneous group of prokaryotic photosynthetic organisms that contain chlorophyll, evolve oxygen, and use water as an electron donor. Many can also fix nitrogen.

1.1.2B. Archaebacteria. Originally researchers thought that prokaryotes and single-cell eukaryotes were the only forerunners of multicellular organisms. A recently discovered class of organisms, the **Archaebacteria,** has cast doubt on this assumption. Archaebacteria constitute unusual organisms; we consider them to belong to kingdom Monera, but they are phylogenetically neither prokaryotes nor eukaryotes. They have some characteristics of prokaryotes (absence of a nucleus and cell organelles), some characteristics of eukaryotes (initiation of protein synthesis and ribosome sensitivity to antibiotics), and some characteristics unique to themselves (composition of cell wall and types of membrane lipids).

Archaebacteria consist of three groups, all of which grow under extreme environmental conditions. The first group comprises *thermoacidophiles,* which grow in hot, acidic environments. Growth typically occurs at temperatures of 80–90°C and at a pH of about 2. Microbiologists have isolated thermoacidophiles from hot sulfur springs and smoldering piles of coal tailings. The second group, called *halophiles,* consists of organisms that grow only in solutions of high salt concentrations. They have been found in such habitats as the Great Salt Lake and the Dead Sea. The third group comprises *methanogens,* strict anaerobes that generate methane by reduction of CO_2. Methanogens occur in hot springs and the ocean bottom.

Scientists think that these harsh environmental conditions required for growth may reflect the adaptation of ancient archaebacteria to the demanding conditions of the primitive Earth. According to this view, archaebacteria represent one of the earliest forms of living organisms.

1.2. THE LIVING CELL

1.2.1. The Varied Types of Cells

The cell is the unit of structure and function in all living organisms. It represents an incredible miracle of organization. Every cell carries out a large number of simultaneous metabolic reactions and controls the actions of a multitude of molecules. A cell of the bacterium *Escherichia coli,* the colon bacillus, contains about *5000 different types of molecules* and an estimated total of some *500 million molecules.* This number does not include the water molecules, of which there are about four billion (Table 1.1). Orchestrating this huge array of molecules to produce one functional entity is an amazing feat.

Many different types of cells occur in nature. We commonly base their classification on the form in which carbon is required, on the type of energy source used, and on cell structure.

1.2.1A. Carbon Source. Based on the form of carbon they require, we divide cells into two types. **Autotrophs** ("self-feeding") can use CO_2 as their sole source of carbon and construct from it the carbon skeletons of all their organic compounds. **Heterotrophs** ("feeding on others") cannot use CO_2 exclusively and must obtain carbon from compounds produced by other cells (autotrophs). Photosynthetic cells and some bacteria are autotrophs, whereas cells of higher animals and most microorganisms are heterotrophs. Heterotrophs can be subdivided into two groups: **aerobes,** which require oxygen, and **anaerobes,** which use other oxidizing agents.

Table 1.1. Approximate Chemical Composition of a Rapidly Dividing Cell of *Escherichia coli*[a]

Component	Percent of total cell weight	Average molecular weight (MW)	Number of different kinds of molecules per cell	Number of molecules per cell
Water	70	18	1	4.0×10^{10}
Inorganic ions	1	40	20	2.5×10^{8}
Polymers				
Proteins	15	40,000	2000–3000	2×10^{6}–3×10^{6}
DNA	1	1.6×10^{9}	1	2 or 4[b]
RNA	6	2.5×10^{4}–1.0×10^{6}	1000	6.0×10^{5}
Polysaccharides	0.8	1.0×10^{6}	5	1.2×10^{4}
Monomers				
Amino acids	0.4	120	100[c]	3.0×10^{7}
Nucleotides	0.4	300	200[d]	1.2×10^{7}
Carbohydrates	3.0	150	200	2.0×10^{8}
Lipids	2.0	750	50	2.5×10^{7}
Other	0.4	150	200	1.5×10^{7}
Total (excluding H_2O)			~5000	~5×10^{8}

[a]Adapted, with permission, from J. D. Watson, *Molecular Biology of the Gene*, 3rd ed., W. A. Benjamin, Menlo Park, California (1976).
[b]There are two single strands (one double strand) before replication and four single strands after replication. The molecular weight given is that of a single strand.
[c]Includes the 20 amino acids found in proteins plus others functioning in metabolism.
[d]Includes the standard nucleotides of DNA and RNA plus others functioning in metabolism.

1.2.1B. Energy Source. Based on the type of energy source that they use, we divide cells into two groups. **Phototrophs** use light as their source of energy. **Chemotrophs** use chemical compounds as a source of energy in a series of oxidation–reduction reactions.

1.2.1C. Cell Structure. By far the most important classification of cells is based on their structure. Major differences define two large classes of cells and organisms, *prokaryotes* and *eukaryotes* (Figure 1.3). All prokaryotes are unicellular organisms; eukaryotes include all of the multicellular forms of life and some unicellular organisms.

Two major structural differences exist between prokaryotes and eukaryotes. In eukaryotes (from the Greek, meaning "true nucleus"), the genetic material, DNA, occurs in a special subcellular structure, the **nucleus.** The nucleus is bound by a double membrane called the **nuclear envelope.** The material surrounding the nucleus constitutes the **cytoplasm.** Pores pierce the nuclear envelope and permit communication between the nucleus and the cytoplasm. Prokaryotes lack a true nucleus, and their DNA occurs in a **nuclear region,** or **nucleoid,** a section of the cytoplasm that contains the DNA but is not surrounded by a membrane.

The second structural difference relates to the com-

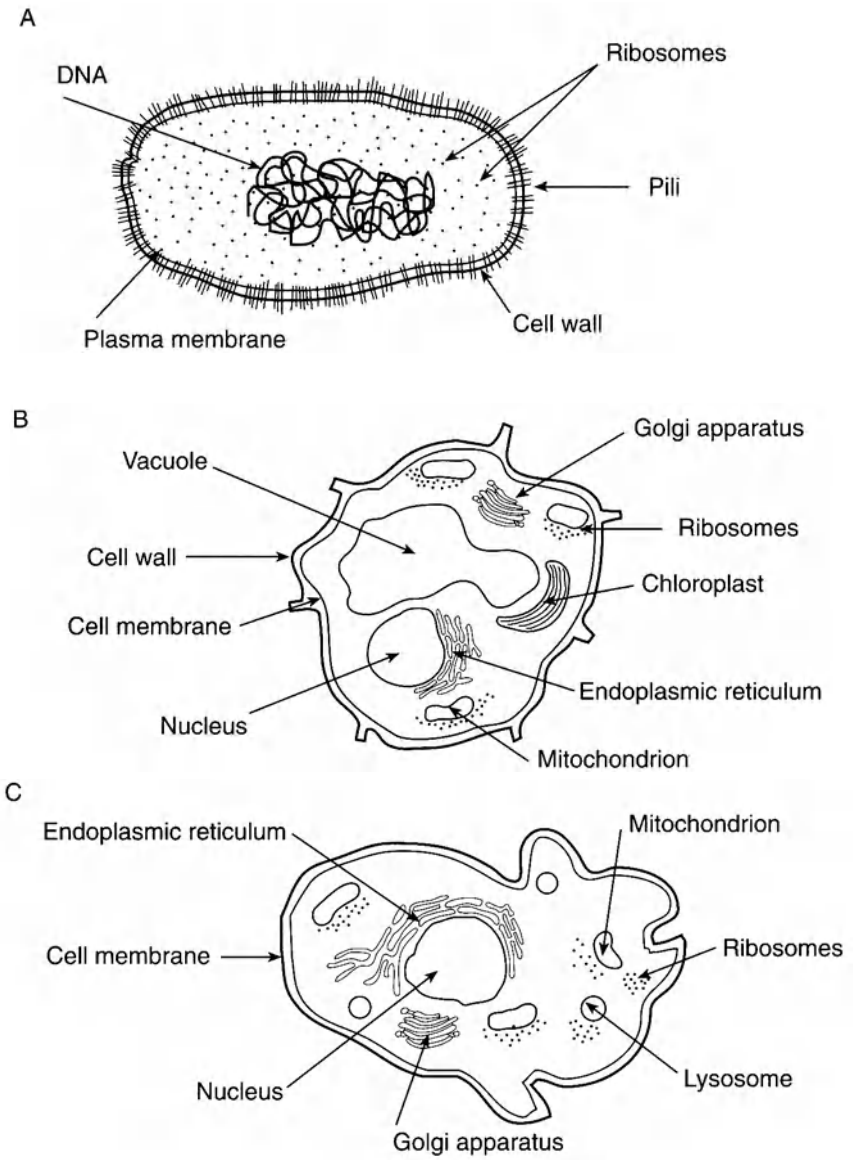


Figure 1.3. Schematic drawings of typical prokaryotic (A), plant (B), and animal cells (C).

Table 1.2. Major Functions of Some Eukaryotic Organelles and Other Cell Compartments

Compartment	Major functions
Plasma or cell membrane	Transport of ions and molecules Receptors for binding molecules Cell morphology and movement Separation of intra- and extracellular environments
Nucleus	DNA synthesis and repair RNA synthesis
Nucleolus	RNA processing
Endoplasmic reticulum	Membrane and lipid synthesis Synthesis of proteins for organelles and for export Detoxification reactions
Golgi apparatus	Modification of proteins Export of proteins
Mitochondria	Cellular respiration Energy conservation Oxidation of carbohydrates and lipids Urea and heme synthesis
Lysosomes	Cellular digestion (hydrolysis) of proteins, carbohydrates, lipids, and nucleic acids
Peroxisomes	Oxidative reactions involving O_2 Metabolism of H_2O_2 and other peroxides
Microtubules and microfilaments	Cytoskeleton structure Cell morphology and motility Intracellular movements
Cytosol	Metabolism of carbohydrates, lipids, amino acids, and nucleotides Protein synthesis other than that of endoplasmic reticulum

position of the cytoplasm. In eukaryotes, the cytoplasm consists of a soluble portion, the **cytosol,** and a particulate portion, the **organelles.** Organelles are specialized subcellular structures that have definite functions. One eukaryotic organelle is the nucleus; several others are listed in Table 1.2. Eukaryotes have many organelles; prokaryotes, by contrast, have few or no organelles and none that are membrane bound.

1.2.2. Comparing Prokaryotes and Eukaryotes

Prokaryotic cells have relatively simple structures. As mentioned, the cytoplasm contains few or no organelles, and the DNA occurs in a cytoplasmic nuclear region. Prokaryotic DNA is a single, double-stranded, circular molecule. It is also called *chromosome,* though we use the term particularly for eukaryotic DNA. Since the DNA contains all of the genes and genetic signals, it constitutes the **genome** of the cell.

Throughout the cytoplasm are scattered *ribosomes,* subcellular particles, composed of RNA and protein, that function as sites of protein synthesis. A **plasma membrane** or **cell membrane** separates the cytoplasm from the outside world, and ions and molecules move across this membrane in both directions. External to the cell membrane in prokaryotes is the **cell wall,** a rigid protective layer containing polysaccharides. Some prokaryotes have an additional outer cell coat of gelatinous polysaccharide, known as a *capsule.*

The outer surface of a prokaryotic cell may be covered with *pili* (from the Latin, meaning "hair") and *flagella* (from the Latin, meaning "whip"). Pili are small filamentous projections that aid in attaching the cell to other cells or surfaces. Pili also play a role in *conjugation,* the genetic recombination in bacteria that resembles sexual reproduction. Flagella are threadlike cellular extensions, longer and less numerous than pili. Their rotation aids in bacterial locomotion.

Eukaryotic cells have greater structural complexity. A nucleus, surrounded by a nuclear envelope, contains the DNA. Eukaryotic DNA occurs in the form of chromosomes that are large enough to be visible with the light microscope when they become condensed. Most cells contain two complete sets of chromosomes (*homologous pairs*). Each chromosome consists of one giant double-stranded DNA molecule associated with basic proteins, named *histones.* DNA and histones form complexes called *nucleosomes.* A dense subcompartment of the nucleus, the **nucleolus,** is rich in RNA and functions in the assembly of ribosomes.

Eukaryotic cells, like prokaryotic cells, possess a plasma membrane. A cell wall, however, does not occur in all eukaryotes. Plant eukaryotes have rigid cell walls, but animal eukaryotes have no cell wall. Many eukaryotic cells have a thick, fuzzy cell coat of complex carbohydrates, called *glycocalyx,* located outside the cell membrane.

The cytoplasm of eukaryotic cells contains a network of cisternae, vesicles, and tubules known as the **endoplasmic reticulum (ER).** This network constitutes part of a single continuous membrane system that extends throughout the cell; it is attached to both the plasma membrane and the nuclear envelope. The endoplasmic reticulum occurs in two forms: one part, the *rough endoplasmic reticulum (RER),* is studded with ribosomes and functions in the synthesis of proteins; the other part, the *smooth endoplasmic reticulum (SER),* does not contain ribosomes and functions in the synthesis of lipids.

We refer to another eukaryotic organelle—the *mitochondrion*—as the "power house" of the cell. As this term implies, mitochondria function in energy metabolism. You have already read that mitochondria are believed to

be descendants of ancient aerobic bacteria that took up symbiotic relationships with primitive eukaryotic cells. Other eukaryotic organelles include the **Golgi apparatus,** a membrane-bound organelle, frequently associated with the endoplasmic reticulum, that functions in protein export from cells; **lysosomes,** sacs rich in enzymes that catalyze hydrolysis reactions as part of metabolite breakdown; and **peroxisomes,** which function in the metabolism of hydrogen peroxide (H_2O_2) and other peroxides.

Throughout the cytosol of eukaryotic cells extends a filamentous, flexible, and dynamic network called the **cytoskeleton.** Its major components—*microtubules, microfilaments,* and *intermediate filaments*—are interconnected by means of the *microtrabecular network,* a lattice of very thin, interlacing filaments. The cytoskeleton gives cells their characteristic shape and is responsible for changes in cell shape during locomotion, cell division, and phagocytosis. It also provides attachment sites for the organelles and allows for communication between different parts of the cell.

Plant cells differ from animal cells in a number of respects. As already mentioned, plant cells have a rigid cell wall whereas animal cells do not. Plant cells also have many large *vacuoles.* These are membrane-bound sacs, filled with water and used for space filling and for storing waste and toxic substances. Animal cells contain few and small vacuoles. Lastly, plant cells possess specialized organelles: *chloroplasts* function in photosynthesis, and **glyoxysomes** contain enzymes that catalyze the reactions of the glyoxylate cycle.

Table 1.3 summarizes major differences between prokaryotes and eukaryotes.

1.3. WATER—THE SOLVENT OF LIFE

Without water, life *as we know it* would not be possible. Water, in the form of primeval oceans, is believed to have been where life first arose, and ever since it has played an essential role in the growth and reproduction of living organisms.

Water constitutes more than 70% of the weight of most forms of life and provides the medium in which metabolic reactions are catalyzed, chemical energy is transferred, and nutrients are transported. In many metabolic reactions, water serves as either a reactant or a product. Cleavage of water by the action of sunlight during photosynthesis makes water a required compound for most photosynthetic organisms. All aspects of cell structure and function intimately relate to the presence of water. In turn, water possesses some unique properties that make it admirably suitable as a major component of the internal environment of living cells.

Table 1.3. Comparison of Prokaryotic and Eukaryotic Organisms

	Prokaryotes	Eukaryotes
Structural organization		
Nucleus	No	Yes
Organelles	Few or none	Many
Nucleolus	No	Yes
Cell membrane	Yes	Yes
Cell wall	Yes	Yes (plants), No (animals)
Cell coat	Yes (capsule in some)	Yes (glycocalyx in many)
Cytoskeleton	No	Yes
Cell diameter	1–10 μm	10–100 μm
Functional attributes		
Metabolism	Anaerobic or aerobic	Aerobic
Cell replication	Binary fission	Mitosis/meiosis
RNA synthesis	Cytoplasm	Nucleus
Protein synthesis (except membrane and secretory proteins)	Cytoplasm	Cytoplasm
Molecular biological characteristics		
Chromosomes	1	>1
Nucleosomes	No	Yes
Ribosome sizes[a]	30S, 50S, 70S	40S, 60S, 80S
Ribosomal RNA sizes[a]	5S, 16S, 23S	5S, 5.8S, 18S, 28S

[a]Sedimentation coefficients in Svedberg (S) units (see Appendix C).

1.3.1. The Structure of Water

Water's properties derive from the structure of the molecule. The two hydrogen atoms are not colinear with the oxygen atom but protrude at an angle of 105° (Figure 1.4). Moreover, the two elements differ significantly in their **electronegativity**—the tendency of an atom to draw electrons toward itself in a chemical bond (Table 1.4). Oxygen is strongly electronegative, but hydrogen has low electronegativity. In most instances, hydrogen does not draw an electron to itself but rather donates its electron to the chemical bond. It then acts as an **electropositive** element.

Because of the differences in electronegativity, electron density is increased around the oxygen atom and decreased around each of the two hydrogen atoms; oxygen becomes slightly negative (δ^-), and each of the hydrogens slightly positive (δ^+). The *partial separation of charges* along each O—H axis makes one end positive and the other negative. Each O—H axis has two electrical poles and constitutes a **dipole.**

Because of the bent structure of the water molecule, these two dipoles do not cancel each other out, so that the molecule as a whole is also a dipole. The part containing the oxygen carries a negative charge, and the part containing the two hydrogens carries a positive charge. Because the dipole of the water molecule persists at all times, it constitutes a *permanent dipole.*

Because of its dipole structure, water provides a highly polar medium that has a high *dielectric constant.* The dielectric constant measures the capacity of solvent molecules to act as insulators of charged particles by becoming interspersed between them. Dispersion of the charged particles results in their being *solubilized* and illustrates the principle that "like dissolves like." Water has a dielectric constant of 80, compared with 1.9 for hexane, a nonpolar solvent.

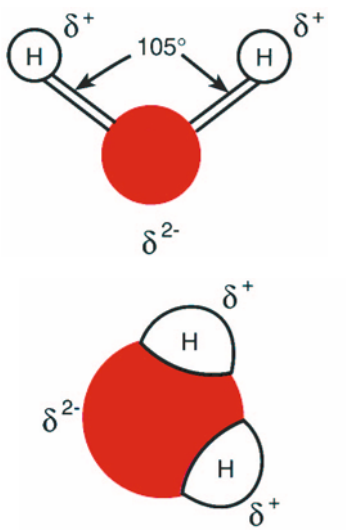


Figure 1.4. Structure of the water molecule. *Top,* Ball-and-stick model; *bottom,* space-filling model.

Table 1.4. Electronegativity of Selected Elements

Element	Electronegativity[a]
Fluorine (F)	4.0
Oxygen (O)	3.5
Nitrogen (N)	3.0
Carbon (C)	2.5
Sulfur (S)	2.5
Phosphorus (P)	2.1
Hydrogen (H)	2.1

[a]Values are derived from bond energies (Pauling scale). Fluorine is the most electronegative element.

Water is an excellent solvent for ionic compounds and for nonionic polar compounds. Dispersion of NaCl illustrates how ionic compounds dissolve in water. When you add NaCl to water, water molecules are attracted to the sodium and chloride ions via the oppositely charged ends of the water dipoles. Surrounding the ions with water prevents them from interacting with each other. The occurrence of solute–solvent interactions that exceed solute–solute interactions constitutes the essence of **solubility.**

1.3.2. Hydrogen Bonding in Water

The partial charges on the H and O atoms in the water molecule also account for water's great capacity to form *hydrogen bonds.* A hydrogen bond forms when a covalently linked hydrogen atom becomes attracted to a neighboring electronegative atom (see Section 1.4.1). In the case of water, hydrogen bonds form when a hydrogen atom of one water molecule becomes attracted to the oxygen of a neighboring water molecule. Extensive hydrogen bonding in water accounts for its high boiling point (100°C) by comparison with the boiling points of related hydrogen compounds such as NH_3 (−33°C), H_2S (−61°C), and CH_4 (−164°C). These compounds exist as gases at room temperature, whereas water is a liquid under the same conditions. Disruption of water's hydrogen-bonded structure and conversion of liquid water to the gas phase requires a great deal more thermal energy than vaporization of the other compounds.

Hydrogen bonding of water involves an interaction of one water molecule with four neighbors. Each water

molecule acts as an *H-bond acceptor* for two neighbors and as an *H-bond donor* for two others. Together, the five water molecules form a tetrahedral structure (Figure 1.5). This arrangement becomes most clearly defined, and has greatest regularity, when liquid water freezes into ice. In ice, the tetrahedral lattice of H-bonded water molecules is continuous and forms a crystal-like structure with many open spaces. In liquid water, the ice lattice has been partially dismantled, with remnants of the structure continually breaking apart and re-forming as molecules move about. Since H-bonds break and form in ice-like clusters, the structure has been dubbed the "flickering cluster" model of liquid water.

Because of its many open spaces, ice is less dense than water and floats on it. If ice were to sink in water, as most solids do in their own liquid form, lakes and rivers would continue to freeze throughout the winter until they became solid blocks of ice, incapable of supporting life.

Two other properties result from the molecular structure of water and provide certain biological advantages. One of these is the large *heat of vaporization*—the number of joules (see Table 1.9) required to change one gram of a liquid to a gas at its boiling point. Because of the large amount of energy involved, evaporation of water, as by sweating, provides living organisms with an effective mechanism for dissipating heat. Water also has a large *specific heat*—the number of joules required to raise the temperature of one gram of a substance by 1°C. Because of its high specific heat, water absorbs heat well and thereby helps organisms minimize changes in temperature. You can see that the large heat of vaporization and specific heat of water aid living organisms in maintaining a constant internal environment.

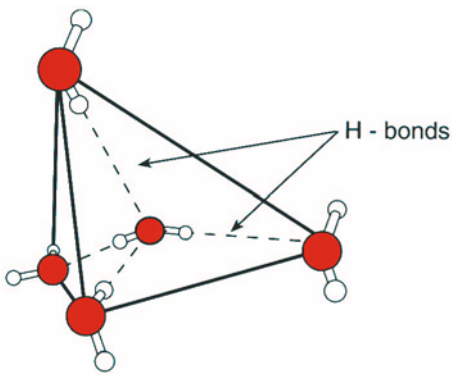


Figure 1.5. Hydrogen bonding in water. Each water molecule is capable of forming hydrogen bonds with four neighboring molecules, located at the corners of a tetrahedron.

1.3.3. Ionization of Water

In addition to its capacity for hydrogen bonding, water's tendency to ionize is also critical for the life of the cell. Water constitutes a *weak electrolyte* that dissociates into ions only to a very small extent. At 25°C, approximately one out of every 10 million water molecules dissociates into hydrogen and hydroxide ions:

$$H_2O \rightleftarrows H^+ + OH^-$$

Actually, a "bare" hydrogen ion (H^+)—a proton—does not exist as such in solution. Instead, the proton becomes hydrated to form a *hydronium ion* (H_3O^+) as described by the equation

$$2H_2O \rightleftarrows H_3O^+ + OH^-$$

In this reaction, one water molecule loses an H^+ to become OH^- while the other gains an H^+ to become H_3O^+. According to the Brønsted–Lowry theory of acids and bases, a **Brønsted acid** is a proton donor and a **Brønsted base** is a proton acceptor. Thus, water, as it ionizes, functions as both an acid and a base; it is **amphoteric.** For simplicity, we will use H^+ throughout this book to describe hydrogen ions in solution.

The extent of water dissociation has been determined by measurements of electrical conductivity. The *equilibrium constant* (K'_{eq}) for this reaction is given by

$$K'_{eq} = \frac{[H^+]\,[OH^-]}{[H_2O]} \qquad (1.1)$$

where brackets indicate molar concentrations. At 25°C, K'_{eq} has a value of 1.8×10^{-16}. Since we base K'_{eq} on molar concentrations rather than activities, we refer to it as an *apparent equilibrium constant.* K'_{eq}, like all equilibrium constants, does not have units.

Because water dissociates into ions to such a small extent, the concentration of undissociated water may be taken to be essentially constant and equal to the total water concentration. We can calculate total water concentration by considering 1.00 liter of water at 25°C. Since the density of water is 1.00 g/ml and its mass 18.0 g/mol, one liter contains 55.6 moles of H_2O. In other words, water is a 55.6*M* solution. Accordingly, we can rewrite Eq. (1.1) as

$$K'_{eq}(55.6) = [H^+][OH^-]$$

The product of the hydrogen and hydroxide ion concen-

trations is called the **ion product of water** (K_w). At 25°C, K_w has a value of 1.0×10^{-14}:

$$K_w = [H^+][OH^-] = 1.0 \times 10^{-14} \qquad (1.2)$$

Provided that we are dealing with ordinary dilute solutions, we can make the approximation that *in any aqueous solution* at 25°C, the product of the molar H^+ and OH^- concentrations is a constant and equal to 1.0×10^{-14}.

1.3.4. pH

In addition to water, many biomolecules ionize in biochemical systems as well. Since both the intracellular and extracellular environments are aqueous solutions, ionizations can occur in either location. The state and extent of ionization depend on the existing pH and on the pK'_a values of ionizable groups.

1.3.4A. Definition and Scale.

The **pH** constitutes a measure of the hydrogen-ion concentration in solution. S. P. L. Sørensen introduced the term in 1909 and defined it as

$$pH = -\log[H^+] = \log \frac{1}{[H^+]} \qquad (1.3)$$

Since dissociation of every molecule of water produces one proton and one hydroxide ion, the concentrations of these two ions in pure water are identical. It follows from Eq.(1.2) that

$$[H^+] = [OH^-] = 10^{-7} M$$

so that

$$pH = -\log(10^{-7}) = \log(1/10^7) = 7$$

The Sørensen definition of pH converts cumbersome exponential values of concentration to a set of small positive numbers and allows us to construct a simple *pH scale*. On that scale, which typically runs from 0 to 14 (Figure 1.6), a *neutral* solution has a pH of 7, *acidic* solutions have pH values below 7, and *basic* solutions have pH values above 7. Two simple pH calculations are illustrated in Appendix A.

Three corollaries follow from the definition of pH. First, the *pH scale is not a linear scale* in terms of hydrogen-ion concentration (Table 1.5). For example, a change of 0.1 pH unit represents different changes in proton concentration ($[H^+]$), depending on the location within the scale. Second, *an increase in $[H^+]$ is equivalent to a decrease in pH.* Lastly, *a small change in pH represents a sizable change in hydrogen-ion concentration.* Thus, a decrease of the pH by one unit represents a 10-fold increase in the hydrogen-ion concentration.

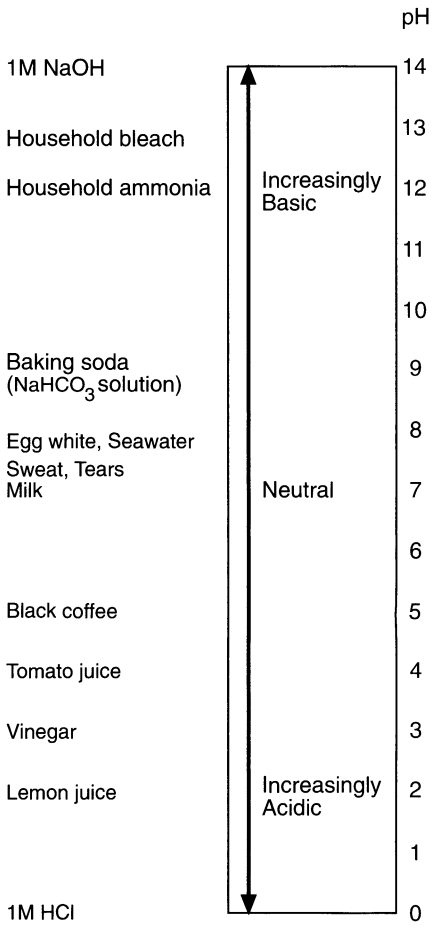


Figure 1.6. The pH scale.

Living systems generally cannot tolerate appreciable changes in pH. In humans, the normal pH of blood is about 7.4; pH values of several other biological fluids are listed in Table 1.6. A small decrease in pH, *acidosis,* or a small

Table 1.5. The Relationship between pH and $[H^+]$

pH	$[H^+]$
0	$1 \times 10^0 M$
6.9	$130 \times 10^{-9} M$
7.0	$100 \times 10^{-9} M$
7.1	$80 \times 10^{-9} M$
7.9	$13 \times 10^{-9} M$
8.0	$10 \times 10^{-9} M$
8.1	$8 \times 10^{-9} M$
14	$1 \times 10^{-14} M$

Table 1.6. pH Values of Some Human Fluids

Fluid	pH
Pancreatic juice	7.8–8.0
Blood plasma	7.4
Interstitial fluid	7.4
Cytoplasm, liver cell	6.9
Saliva	6.4–7.0
Cytoplasm, muscle cell	6.1
Urine	5.0–8.0
Gastric juice	1.5–3.0

increase in pH, *alkalosis,* can be tolerated, but larger changes are deleterious. A drop in the pH of blood below 6.8 or a rise above 7.8 is usually lethal. Thus, human life can function normally only within a blood pH range of about one pH unit (7.3 ± 0.5).

1.3.4B. Brønsted Acids and Bases. Many of the acids and bases of importance in biochemistry are weak acids and bases. These include amino acids that serve as building blocks of proteins, purines and pyrimidines that constitute structural components of nucleic acids, and numerous metabolites. Like water, such weak acids and bases dissociate only slightly into ions in solution. When a weak *monoprotic acid* ("containing one dissociable proton") dissociates into ions

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

HA constitutes a *Brønsted acid* or proton donor, and A^- a *Brønsted base* or proton acceptor. Together, HA and A^- represent a **conjugate acid–base pair.** Whenever a compound acts as a Brønsted acid (the protonated form), it must form a Brønsted base (the deprotonated form), and vice versa. For the above reaction:

HA = Brønsted acid; proton donor; conjugate acid

A^- = Brønsted base; proton acceptor; conjugate base

We must stress two points regarding Brønsted acids and bases in biochemical systems. First, note that a compound may act as an acid in one reaction and as a base in another, in which case it is amphoteric. As an illustration, consider the ionization of phosphoric acid, a *triprotic acid* ("containing three dissociable protons"):

$$H_3PO_4 \text{ (acid)} \rightleftarrows H_2PO_4^- \text{ (base)} + H^+ \qquad (1.4)$$
$$pK'_{a_1} = 2.12$$
$$H_2PO_4^- \text{ (acid)} \rightleftarrows HPO_4^{2-} \text{ (base)} + H^+ \qquad (1.5)$$
$$pK'_{a_2} = 7.21$$
$$HPO_4^{2-} \text{ (acid)} \rightleftarrows PO_4^{3-} \text{ (base)} + H^+ \qquad (1.6)$$
$$pK'_{a_3} = 12.4$$

In reaction (1.4), $H_2PO_4^-$ functions as a base, while in reaction (1.5) the same ion functions as an acid. Likewise, HPO_4^{2-} acts as a base in reaction (1.5), but as an acid in reaction (1.6).

The second point deals with an important convention: *In biochemistry, we consider all dissociations involving protons in terms of a loss of protons.* Thus, all dissociation constants used in biochemistry represent **acid dissociation constants (K'_a)** and the corresponding pK' values represent **acid pK' values,** or **pK'_a** ($pK'_a = -\log K'_a$). This means that not only weak acids, like the phosphoric acid forms shown above, dissociate to yield protons, but weak bases are likewise considered to be in their protonated forms and to dissociate to yield protons:

Glycine: $R\text{-}NH_3^+ \rightleftarrows R\text{-}NH_2 + H^+$
$pK'_a = 9.60$

Ammonia: $NH_4^+ \rightleftarrows NH_3 + H^+$
$pK'_a = 9.25$

1.3.4C. Henderson–Hasselbalch Equation. An important relationship exists between the acid dissociation constant of a weak acid and the pH of a solution containing both that acid and its conjugate base. This relationship is described by the **Henderson–Hasselbalch equation.** The equation has wide application in biochemical research, particularly when we need to control pH for optimum reaction conditions.

We can readily derive the Henderson–Hasselbalch equation from the dissociation reaction of a weak acid ($HA \rightleftarrows H^+ + A^-$):

$$K'_a = \frac{[H^+]\,[A^-]}{[HA]} \qquad (1.7)$$

By taking logarithms and rearranging, we obtain

$$\log K'_a = \log [H^+] + \log \frac{[A^-]}{[HA]}$$

or

$$-\log[H^+] = -\log K'_a + \log \frac{[A^-]}{[HA]}$$

Since $pK'_a = -\log K'_a$ (analogous to pH = $-\log [H^+]$), we replace the logarithmic terms with pH and pK'_a to obtain the Henderson–Hasselbalch equation:

$$pH = pK'_a + \log \frac{[A^-]}{[HA]} \qquad (1.8)$$

Because we consider both weak acids and weak bases to behave as *Brønsted acids* with regard to proton dissociation, the HA of the Henderson–Hasselbalch equation denotes *any* Brønsted acid: an undissociated weak acid *or* a protonated weak base. Likewise, A^- denotes *any* Bronsted base: a dissociated weak acid *or* an unprotonated weak base. It follows that the symbols HA and A^- do not necessarily indicate the correct charges of the Brønsted base and acid. Some examples of $[A^-]/[HA]$ illustrate this point:

$$[CH_3COO^-]/[CH_3COOH]$$
$$[HPO_4^{2-}]/[H_2PO_4^-]$$
$$[NH_3]/[NH_4^+]$$

Based on the above, we can write the Henderson–Hasselbalch equation in its most general form as

$$pH = pK'_a + \log \frac{[\text{proton acceptor or Brønsted base}]}{[\text{proton donor or Brønsted acid}]} \quad (1.9)$$

To practice using the Henderson–Hasselbalch equation, refer to the examples in Appendix A. In one example, we calculate the percent of a compound present in its HA and A^- forms at certain pH values. This calculation leads to an important general conclusion: *At a pH that is one pH unit above the pK'_a of a Brønsted acid, 91% of the compound will be present in the dissociated form (A^-).* If the pH is two units above the pK'_a, *99% of the compound will be present in its A^- form. Conversely, if the pH is one (or two) units below the pK'_a of a Brønsted acid, 91% (or 99%) of the compound will be present in the undissociated form (HA).* Note that these percentages are determined by the difference between pH and pK'_a, not by their absolute values. You must keep these results in mind when considering charges of biomolecules at various pH values. When the pH differs from the pK'_a by two units or more, you can assume that, for all practical purposes, the compound exists entirely in either one form or the other.

1.3.5. pK'_a

1.3.5A. The Meaning of pK'_a.

In the previous section we saw that $pK'_a = -\log K'_a$. This constitutes a *mathematical definition.* By referring to the Henderson–Hasselbalch equation we can also provide a *conceptual definition* of pK'_a. Note that, when $[A^-] = [HA]$, the logarithmic term becomes log 1, which is equal to zero so that $pH = pK'_a$. Thus, pK'_a represents a pH value. Specifically, pK'_a *is the pH of a solution that contains equal concentrations of the dissociated and undissociated forms of a conjugate acid–base pair.*

We can illustrate the conceptual definition of pK'_a with a hypothetical solution of lactic acid ($HA = CH_3—CHOH—COOH$; $A^- = CH_3—CHOH—COO^-$), which has a pK'_a of 3.86. If we dissolve, for example, 4×10^{20} molecules of undissociated lactic acid (HA, the protonated form) and 4×10^{20} "molecules" of sodium lactate (A^-, the deprotonated form) in 250 ml of water, the pH of the resulting solution will be exactly 3.86. The number of molecules and the number of milliliters chosen here are completely arbitrary. As long as the amounts of lactic acid and sodium lactate are identical (in terms of moles, millimoles, micromoles, and the like, not in terms of grams), they can be dissolved in any volume and the resulting pH will be 3.86.

The condition $[HA] = [A^-]$ occurs at the inflection point of the titration curve of the weak acid HA (Figure 1.7). At that point in the curve, exactly one-half of the total number of molecules have been titrated and converted from the HA form to the A^- form. Hence, we can also define pK'_a as the pH corresponding to the inflection point of such a titration curve.

1.3.5B. The Effects of Variables on pK'_a.

At the same molar concentration, a stronger acid dissociates more and yields greater concentrations of A^- and H^+ than a weaker acid. Therefore, the stronger the acid is, the greater the dissociation constant and the smaller the pK'_a. Lactic acid, with a pK'_a of 3.86 ($K'_a = 1.38 \times 10^{-4}$), is a stronger acid than acetic acid, which has a pK'_a of 4.76 ($K'_a = 1.76 \times 10^{-5}$). Table 1.7 lists the pK'_a values of some acids encountered in biological systems.

Several other variables affect pK'_a. Because an equilibrium constant varies with temperature according to the van't Hoff equation, pK'_a must vary likewise. Additionally, pK'_a decreases as the ionic strength (see Appendix A)

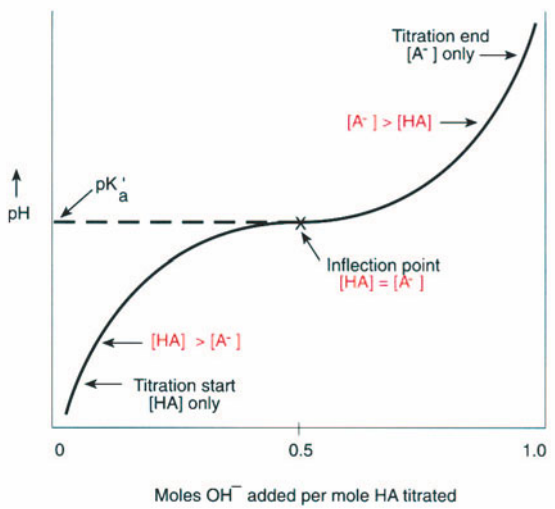


Figure 1.7. Titration of a weak monoprotic acid, HA, with base: $HA + OH^- \rightleftarrows A^- + H_2O$. The relatively flat portion of the curve, on either side of the inflection point, constitutes the buffering region (Section 1.3.6).

Table 1.7. pK_a' Values of Some Brønsted Acids

Acid	Conjugate acid/conjugate base (HA) (A⁻)	pK_a'
Phosphoric acid (1)[a]	H_3PO_4 / $H_2PO_4^-$	2.12
Pyruvic acid	CH_3—CO—COOH / CH_3—CO—COO^-	2.50
Malic acid (1)[a]	HOOC—CH_2—CHOH—COOH / HOOC—CH_2—CHOH—COO^-	3.40
Formic acid	HCOOH / $HCOO^-$	3.75
Lactic acid	CH_3—CHOH—COOH / CH_3—CHOH—COO^-	3.86
Succinic acid (1)[a]	HOOC—CH_2—CH_2—COOH / HOOC—CH_2—CH_2—COO^-	4.21
Acetic acid	CH_3—COOH / CH_3—COO^-	4.76
Malic acid (2)[b]	HOOC—CH_2—CHOH—COO^-/$^-$OOC—CH_2—CHOH—COO^-	5.26
Succinic acid (2)[b]	HOOC—CH_2—CH_2—COO^-/$^-$OOC—CH_2—CH_2—COO^-	5.64
Carbonic acid (1)[a]	H_2CO_3 / HCO_3^-	6.37
Phosphoric acid (2)[b]	$H_2PO_4^-$ / HPO_4^{2-}	7.21
Ammonium ion	NH_4^+ / NH_3	9.25
Carbonic acid (2)[b]	HCO_3^- / CO_3^{2-}	10.25
Phosphoric acid (3)[c]	HPO_4^{2-} / PO_4^{3-}	12.40

[a]Refers to dissociation of the first proton.
[b]Refers to dissociation of the second proton.
[c]Refers to dissociation of the third proton.

of a solution increases. Lastly, the pK_a' of an ionizable group is affected by other charges in the molecule. We can illustrate this by referring once more to the ionization of phosphoric acid (Eqs. 1.4–1.6). The progression in pK_a' values for dissociation of the first (2.12), second (7.21), and third (12.4) protons proceeds in the order of decreasing acid strength or increasing pH because of the changes that occur in the molecule as it ionizes.

The first pK_a' reflects the strength of base, $[OH^-]$, required to abstract a proton from H_3PO_4, a neutral molecule that carries no charge. Subsequent protons must be abstracted from an anion having a charge of -1 ($H_2PO_4^-$) or -2 (HPO_4^{2-}). Abstracting a proton *against* the electrical attraction of the molecular anions becomes progressively more difficult and requires increasing base strength so that $pK_{a_3}' > pK_{a_2}' > pK_{a_1}'$. Such pK_a' relationships hold for all acids containing more than one dissociable proton.

1.3.6. Buffers

Proper functioning of biochemical systems generally requires careful control of pH. Both *in vivo* and *in vitro,* this is accomplished by means of buffers. A **buffer** is a solution containing a mixture of a Brønsted acid, HA, and its conjugate base, A^-, that has the capacity of resisting substantial changes in pH upon addition of small amounts of acid or base (see Figure 1.7).

1.3.6A. Buffers in Biochemistry.

Table 1.8 lists some examples of buffer systems useful in the laboratory. *In vivo,* buffers that control pH belong primarily to three classes: proteins, phosphoric acid systems, and the carbonic acid system. Proteins constitute very effective buffers, on a molar basis, because each molecule contains a large number of weakly acidic and weakly basic functional groups. The phosphate buffer of greatest importance in biological systems is the $H_2PO_4^-/HPO_4^{2-}$ acid–base pair because its pK_{a_2}' (7.21) is close to physiological pH (7.00). Contributions from the other two phosphate systems, $H_3PO_4/H_2PO_4^-$ and HPO_4^{2-}/PO_4^{3-}, can usually be neglected because their pK_a' values differ from pK_{a_2}' by five units. The carbonic acid/bicarbonate system ($pK_{a_1}' = 6.37$) also constitutes an effective buffer under physiological conditions. Its pK_{a_1}' is based both on the ionization of H_2CO_3 to HCO_3^- and on the equilibrium between H_2CO_3 and dissolved CO_2:

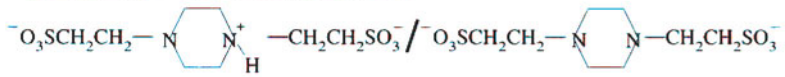
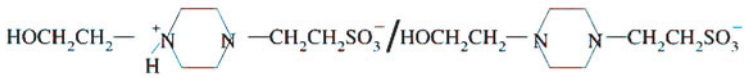
$$H_2CO_3 \rightleftarrows CO_2 + H_2O$$

Laboratory buffers may consist of inorganic compounds (e.g., NaH_2PO_4/Na_2HPO_4) or organic compounds (e.g., citric acid/sodium citrate). Certain buffers, called *biological buffers* or *Good's buffers* (e.g., TRIS, HEPES, and PIPES), are especially useful for biological systems because they:

- have a pK_a' in the range of 6–8,
- are nontoxic and noninhibitory,
- are enzymatically and hydrolytically stable,
- have good solubility in water, and
- have a pK_a' that varies little with temperature.

You can prepare a buffer in one of two ways. You may weigh out specific amounts of the conjugate acid and base forms and dissolve them in a given volume of water. Alter-

Table 1.8. Some Useful Biochemical Buffer Systems

Buffer system HA/A⁻	pK_a'
Cacodylate $(CH_3)_2AsO_2H/(CH_3)_2AsO_2^-$	6.27
PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] 	6.8
Phosphate $H_2PO_4^-/HPO_4^{2-}$	7.21
HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) 	7.6
TRIS [tris(hydroxymethyl)aminomethane] $(HOCH_2)_3CNH_3^+/(HOCH_2)_3CNH_2$	8.3

natively, you can begin with one buffer component and convert some of it to the conjugate form by adding either strong acid or base. The two methods are illustrated in Appendix A.

1.3.6B. Mechanism of Action. Controlling pH with a buffer involves an interconversion of the two buffer components, the Brønsted acid and its conjugate base. Adding acid (H^+) to a buffer converts some of the A^- to the HA form; adding base (OH^-) converts some of the HA to the A^- form. As a result, adding *either* H^+ or OH^- changes the $[A^-]/[HA]$ ratio. Because that ratio appears as a logarithmic function in the Henderson–Hasselbalch equation, the change in pH is relatively minor.

The interconversion of buffer components can be illustrated by the phosphate buffer system, $H_2PO_4^-$ / HPO_4^{2-}, for which:

$$H_2PO_4^- \rightleftarrows HPO_4^{2-} + H^+$$

Adding protons drives this reaction from right to left. Added protons combine with HPO_4^{2-}, thereby converting some HPO_4^{2-} to $H_2PO_4^-$; the equilibrium shifts from right to left according to *Le Chatelier's principle.* Adding hydroxide ions removes protons from the reaction as they combine with the added OH^- to form water. Consequently, some $H_2PO_4^-$ dissociates to form HPO_4^{2-}, and the equilibrium shifts from left to right.

We can characterize a buffer by its working range and by its capacity. The *working range* is determined by the buffer's pK_a'; it falls within ±1 pH unit from pK_a' so that a buffer functions within the range of $[A^-]/[HA] = 10$ to $[A^-]/[HA] = 0.1$ (Figure 1.8). Beyond that range, not enough of *both* buffer forms is present to allow the buffer to function effectively when *either* acid or base is added. The system can, however, still accommodate addition of acid *or* base, as the case might be.

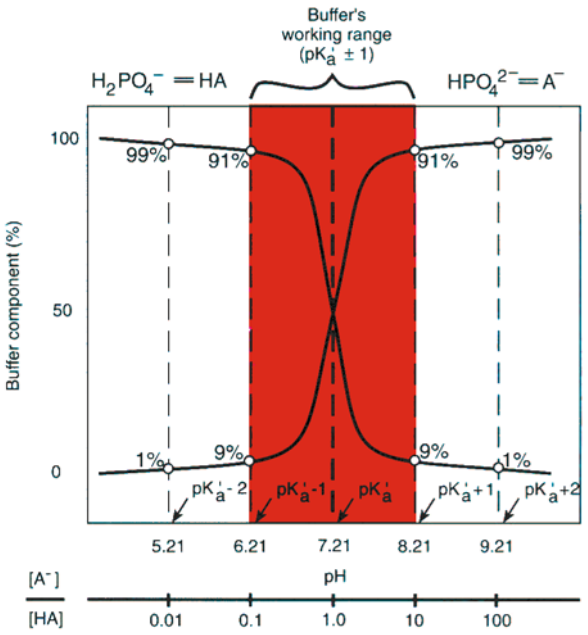


Figure 1.8. Changes in the relative amounts of the two components of a phosphate buffer as a function of pH.

The *capacity* of a buffer is determined by the concentrations of its two components. The greater is the concentration of buffer components, the greater the buffer capacity. Note that buffer concentration always refers to the *total, combined concentration of both buffer components.* Thus, in a 0.5*M* phosphate buffer, made from $H_2PO_4^-$ and HPO_4^{2-}, the sum of the concentrations of $H_2PO_4^-$ and HPO_4^{2-} is 0.5 mole per liter. A buffer has maximum capacity for addition of *either* H^+ or OH^- when buffer pH equals buffer pK'_a since at that point $[A^-] = [HA]$.

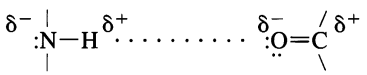
1.4. NONCOVALENT INTERACTIONS

We have already noted the role of one noncovalent interaction—the hydrogen bond—in determining the properties of water. In addition, H-bonds, as well as other types of noncovalent interactions, serve to stabilize proteins and nucleic acids and participate in various binding reactions.

1.4.1. Hydrogen Bonds

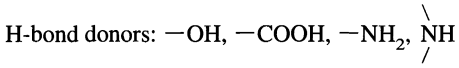
We define a **hydrogen bond** as *the attractive interaction between a covalently linked hydrogen atom and a neighboring electronegative atom.* The atom to which the hydrogen is linked covalently is the **hydrogen bond donor;** the electronegative atom to which the hydrogen becomes attracted is the **hydrogen bond acceptor.** Hydrogen bonds can form between groups within the same molecule, resulting in *intramolecular* H-bonds. Alternatively, groups in different molecules may interact, forming *intermolecular* H-bonds.

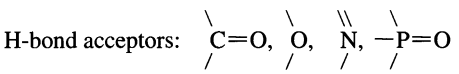
A hydrogen bond results from the interaction between two *permanent dipoles,* with the hydrogen being the positive end of one of these dipoles, as in the following:



Hydrogen bond donor Hydrogen bond acceptor

In biological systems, the electronegative atoms serving as donors and acceptors are usually those of oxygen and nitrogen, and the distance between them varies from 0.26 to 0.31 nm. Typical donor and acceptor groups include the following:





Hydrogen bonds are weaker and have smaller *bond energies* (the standard enthalpy change for breaking a bond) than covalent bonds (Table 1.9). In fact, all noncovalent interactions are weaker than typical covalent bonds; because of this, we frequently refer to them as "weak interactions."

Strengths of hydrogen bonds depend on the electronegativities of the bonded atoms and on their relative alignment. The bond has greatest strength if the two interacting dipoles are both coplanar and coaxial, and it has least strength if the dipoles are neither coplanar nor coaxial (Figure 1.9). Even though each individual bond is weak, numerous H-bonds taken together represent a significant bonding strength. The multiple hydrogen bonds that occur in proteins and nucleic acids play a major role in stabilizing these biopolymers.

1.4.2. Hydrophobic Interactions

We refer to *the attractive forces between nonpolar groups, which lead to their association in an aqueous environment,* as **hydrophobic interactions.** The nonpolar groups can be intact molecules or parts of molecules.

Introducing nonpolar groups into water restricts the

Table 1.9. Some Typical Values of Bond Energies

Bond type	Bond energy[a] (ΔH°; kJ/mol)
Covalent bonds	
S—S	213.4
C—N	292.9
S—H	338.9
C—C	368.2
N—H	393.3
C—H	414.2
H—H	435.1
O—H	460.2
C=N	615.0
C=C	682.0
C=O	711.3
N≡N	945.6
Noncovalent interactions	
Hydrogen bonds	10–40
Hydrophobic interactions	−10 to −30
Ionic interactions	
Ion–ion	40–400
Ion–permanent dipole	3–10
Ion–induced dipole	0.4–3
Van der Waals interactions	
Permanent dipole–permanent dipole	0.5–3
Permanent dipole–induced dipole	0.4–3
Induced dipole–induced dipole	0.1–10

[a]The preferred units of energy are the **joule (J)** or **kilojoule (kJ)** rather than the **calorie (cal)** or **kilocalorie (kcal).** A joule, the SI unit of energy, is the product of force ($kg \cdot m/s^2$) and distance (m). Thus, $1\ J = 1\ kg \cdot m^2/s^2$. One kilojoule is equal to 1000 joules. A calorie, or *small calorie,* is the amount of heat required to raise the temperature of one gram of water by one degree Celsius (from 14.5 to 15.5°C). The calorie used in nutrition is the *large calorie* (Cal; kcal), equal to 1000 small calories. The conversion between calories and joules is 1 cal = 4.184 J.

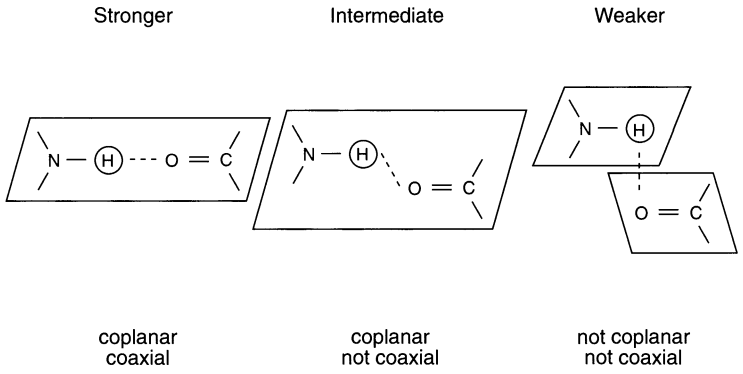


Figure 1.9. The strength of hydrogen bonds.

water molecules in their movement and in their capacity to form hydrogen bonds with each other. When these nonpolar groups associate, water molecules can move about more freely and can form hydrogen bonds more extensively. Because aggregation of the nonpolar groups results in a more random water structure, water's *entropy* (a measure of randomness) increases. The increase in entropy constitutes the basic driving force of hydrophobic interactions. In a sense, the term *hydrophobic* (from the Greek, meaning "fear of water") is a misnomer. It implies that the dissolved substance "dislikes" water, whereas in reality it is the water that dislikes the dissolved substance.

We can illustrate the principles of hydrophobic interactions by some data (Table 1.10) on the transfer of small nonpolar hydrocarbons from the polar environment of water to the nonpolar environment of an organic solvent. Such transfers are analogous to forming hydrophobic interactions in which nonpolar groups must also leave their polar aqueous environment to interact with other like groups in a nonpolar environment composed of the associated groups.

On the basis of these data, we can characterize hydrophobic interactions by an increase in entropy and a decrease in free energy. Consequently, forming hydrophobic interactions is thermodynamically favored. The interactions represent unusual bonding because they stabilize as the temperature increases (ΔH° is positive); the equilibrium constant for forming the interactions increases with temperature over the range of about 0–60°C.

To illustrate the role of hydrophobic interactions in biochemical systems, let us consider the formation of a soap *micelle.* Soap molecules are salts of long-chain carboxylic acids. As an example, the sodium salt of *stearic acid,* sodium stearate, has the structure

$$CH_3-(CH_2)_{16}-COO^-Na^+$$

which can be represented as consisting of two parts:

Hydrocarbon	Ion
Organic	Inorganic
Nonpolar	Polar
Water-insoluble	Water-soluble

The long hydrocarbon chain has nonpolar character, typical of organic compounds. If it existed by itself, it would be water-insoluble. By contrast, the carboxyl group with its associated sodium ion has ionic character, typical of inorganic compounds. If this polar group existed by itself, it would be water-soluble. We call such molecules, composed of two distinctly different regions, one polar or

Table 1.10. Thermodynamic Parameters for the Transfer of Hydrocarbons from Polar to Nonpolar Solvents at 25°C[a]

Process	ΔH° (kJ / mol^{-1})	ΔS° (J deg^{-1} mol^{-1})	ΔG° (kJ / mol^{-1})
CH_4 (in H_2O) → CH_4 (in benzene)	+11.7	+75.3	−10.9
CH_4 (in H_2O) → CH_4 (in CCl_4)	+10.5	+75.3	−12.1
C_2H_6 (in H_2O) → C_2H_6 (in benzene)	+ 9.2	+83.7	−15.9
C_2H_6 (in H_2O) → C_2H_6 (in CCl_4)	+ 7.1	+75.3	−15.5

[a]Changes in free energy (ΔG°), enthalpy (ΔH°), and entropy (ΔS°) are related as follows: $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$, where T is the absolute temperature.

hydrophilic, the other nonpolar or *hydrophobic,* **amphipathic molecules.**

When we dissolve soap molecules in water at a concentration above a *critical micelle concentration,* they spontaneously form spherical aggregates called **micelles** (Figure 1.10). In each micelle, the nonpolar hydrocarbon chains are arranged in radial fashion on the inside of a sphere and are held together by hydrophobic interactions. The carboxyl groups with their associated sodium ions constitute the surface of this sphere. By forming a micelle, nonpolar hydrocarbon chains move from the polar environment of water to a nonpolar environment, composed of other hydrocarbon chains.

Note that each micelle becomes coated with a negative charge and is thus repelled by, and kept separate from, other micelles. Because of their surface charge, micelles can be hydrated and "solubilized" by water like ordinary salt ions.

As we shall see later, the principle of the soap micelle also applies to the structure of globular proteins and biological membranes. In the case of globular proteins, nonpolar amino acids tend to be clustered inside the molecule while polar amino acids are spread over its surface. In the case of biological membranes, their basic structure is a lipid bilayer, a double layer of amphipathic lipids. The lipids have their nonpolar parts in the interior of the bilayer and their polar parts on its two surfaces.

1.4.3. Ionic Interactions

The electrostatic forces between charged entities, at least one of which is an ion, are called **ionic interactions.** Three types of ionic interactions occur: ion–ion, ion–permanent dipole, and ion–induced dipole.

The energy of interaction between charged entities decreases rapidly with increasing distance (r) between them. It varies with $1/r$, $1/r^2$, and $1/r^4$ for the ion–ion, ion–permanent dipole, and ion–induced dipole interaction, respectively. The first two types of interactions may be both attractive, formed between oppositely charged entities, or repulsive, formed between charged entities of the same sign. Ion–induced dipole interactions are strictly attractive in nature.

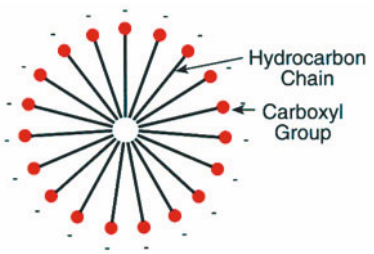


Figure 1.10. Structure of a soap micelle in water. Sodium ions or other cations (not shown) surround the micelle, so that the solution is electrically neutral.

Ion–ion interactions constitute the strongest of the three types of ionic interactions (Table 1.9). We call the attractive force of such an interaction an **electrostatic bond** and describe it by *Coulomb's law:*

$$F = \frac{q_1 q_2}{Dr^2} \tag{1.10}$$

where F is the force of attraction or repulsion between two ions, q_1 and q_2 are the charges of the two ions, D is the dielectric constant of the medium, and r is the distance between the two ions.

Note that the force of interaction varies with $1/r^2$ whereas the energy of interaction, as noted above, varies with $1/r$. Electrostatic bonds occur in proteins between positively and negatively charged functional groups in the side chains of neighboring amino acids. For example, a bond could form between the β-carboxyl group of aspartic acid and the ϵ-amino group of lysine:

$$\text{———}\ CO_2^- \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot\ {}^+H_3N\text{———}$$

We can use Coulomb's law to get a rough idea of the interaction between charged groups in a protein, provided that these groups are far enough removed from other charged groups that the effect of the latter can be ignored.

In an ion–permanent dipole interaction, an ion is attracted to, or repelled by, one of the ends of a permanent dipole. Examples of such dipoles in biochemical systems are the potential donors and acceptors of hydrogen bonds:

$$-COO^- \cdot \cdot \cdot \cdot \overset{\delta^+}{H}-\overset{\delta^-}{N}:-$$

Ion — Permanent dipole (H-bond donor)

Hydration of salt ions, like Na^+ and Cl^-, by water molecules is another example of ion–permanent dipole interactions.

An ion–induced dipole interaction represents the attractive force between an ion and an induced dipole. In order to see how this type of interaction can occur, imagine a positive ion and a nonpolar molecule moving toward each other (Figure 1.11). As the cation approaches the molecule, some of the electrons in the molecule shift toward the side closest to the ion, producing a slight polarity in the molecule. When the cation is next to the molecule, it induces a definite polarity so that the molecule becomes an *induced dipole.* The attractive force that now

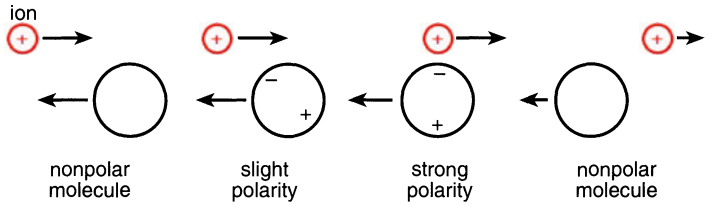


Figure 1.11. Ion–induced dipole interaction (left to right). Induced polarization in the ion has been neglected for simplicity.

exists between the ion and the induced dipole also produces a slight polarization (a dipole) in the cation. As the ion passes the molecule, the molecule loses its polarization and returns to its nonpolar state.

1.4.4. Van der Waals Interactions

The electrostatic forces of attraction between two dipoles are called **van der Waals interactions.** Because dipoles may be either permanent or induced, three types of interactions occur:

1. Permanent dipole–permanent dipole: Orientation effect
2. Permanent dipole–induced dipole: Induction effect
3. Induced dipole–induced dipole: London dispersion forces

Van der Waals interactions represent short-range forces. The energy of interaction drops off rapidly with increasing distance (r) between the dipoles; it varies with $1/r^6$. Although van der Waals interactions are weak, their contribution can add up to a considerable force as the number of interactions between molecules increases. The interaction between two permanent dipoles resembles that between two ions. A permanent dipole–induced dipole interaction is analogous to the ion–induced dipole interaction described above.

Other types of induced dipoles occur because, at any given moment, the distribution of electrons around an atom is not perfectly symmetrical. The atom functions as a small instantaneous dipole that can induce a small instantaneous dipole in a neighboring atom. The interaction between these two transient dipoles constitutes a van der Waals induced dipole–induced dipole interaction.

SUMMARY

Scientists believe that life evolved over a long period of both chemical and biological evolution, the former amounting to about one billion years followed by about four billion years of the latter. Chemical evolution consisted of three stages: synthesis of low-molecular-weight compounds; their polymerization to yield macromolecules; and self-assembly of macromolecules to produce forerunners of modern cells.

Living cells are classified on the basis of the form of carbon they require, the type of energy source they use, and their overall structure. Eukaryotes differ from prokaryotes in having DNA located in a membrane-bound nucleus and in possessing cytoplasmic organelles.

Water molecules form dipoles that are held together by hydrogen bonds. This structure is responsible for the polarity of water, its large specific heat, and large heat of vaporization, all properties that make water an ideal solvent for living systems. Water ionizes to a slight extent, yielding protons (H^+) and hydroxide (OH^-) ions.

Proton concentration is expressed as $pH = -\log [H^+]$ and is equal to 7.0 in pure water. Acidic solutions have pH values below 7.0, and basic solutions have pH values above 7.0.

In biochemistry, we use the Brønsted definition of acids and bases. An acid (HA) is a proton donor, and a base (A^-) is a proton acceptor. The two form a conjugate acid–base pair. We consider proton equilibria of both weak acids and weak bases in terms of loss of protons. The pK'_a of an

acid–base pair is the pH of a solution containing equal concentrations of the HA and A^- forms. Calculations for such systems involve the Henderson–Hasselbalch equation.

Proper functioning of living systems requires careful control of pH by means of buffers—conjugate acid–base systems that minimize changes in pH upon addition of acid or base. A buffer functions by interconversion of its two components. Important buffers in biochemistry include proteins, the $H_2PO_4^-/HPO_4^{2-}$ pair, and the H_2CO_3/HCO_3^- system. A buffer functions effectively within one pH unit of its pK'_a ($pH = pK'_a \pm 1$).

Four types of noncovalent interactions occur in biochemical systems. A hydrogen bond forms when a covalently linked hydrogen becomes attracted to another electronegative atom. Hydrophobic interactions comprise the attractive forces that cause nonpolar molecules or groups to associate together in an aqueous environment. Ionic interactions consist of attractive and repulsive forces between ions or between an ion and a permanent or induced dipole. Van der Waals interactions consist of attractive forces produced by different combinations of permanent and induced dipoles. Noncovalent interactions serve to stabilize the structure of proteins and nucleic acids and participate in binding reactions. We refer to these interactions as "weak interactions" because their strength is considerably less than that of common covalent bonds.

SELECTED READINGS

Alberts, B., *et al., Molecular Biology of the Cell,* 3rd ed., Garland Publishing, New York (1994).

Calvin, M., *Chemical Evolution,* Oxford University Press, New York (1969).

Cassé, M., Lehoucq, R., and Vangioni-Flam, E., Production and evolution of light elements in active star-forming regions, *Nature (London)* 373:318–319 (1995).

de Duve, C., The birth of complex cells, *Sci. Am.* 274:50–57 (1996).

Fenchel, T., and Finlay, B. J., The evolution of life without oxygen, *Am. Sci.* 82:22–29 (1994).

Ferris, J. P., Hill, A. R., Jr., Liu, R., and Orgel, L. E., Synthesis of long prebiotic polymers on mineral surfaces, *Nature (London)* 381: 59–61 (1996).

Gould, S. J., The evolution of life on the earth, *Sci. Am.* 271:84–91 (1994).

Herrmann, B., and Hummel, S. (eds.), *Ancient DNA,* Springer-Verlag, New York (1994).

Kleinsmith, L. J., and Kish, V. M., *Principles of Cell and Molecular Biology,* HarperCollins, New York (1995).

Lohse, P. A., and Szostak, J. W., Ribozyme-catalyzed amino acid transfer reactions, *Nature (London)* 381:442–444 (1996).

McMillan, G. K., *pH Measurement and Control,* 2nd ed., Instrument Society of America, Research Triangle Park, North Carolina (1994).

Orgel, L. E., The origin of life on the earth, *Sci. Am.* 271:76–83 (1994).

Robertson, M. P., and Miller, S. L., An efficient prebiotic synthesis of cytosine and uracil, *Nature (London)* 375:772–774 (1995).

Roger, J., Angel, P., and Woolf, N. J., Searching for life on other planets, *Sci. Am.* 274:60–66 (1996).

Stenesh, J., *Core Topics in Biochemistry,* Cogno Press, Kalamazoo, Michigan (1993).

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

Woese, C. R., Archaebacteria, *Sci. Am.* 244:98–122 (1981).

REVIEW QUESTIONS

A. Define each of the following terms:

Hydrophobic interactions	pH
Micelle	Protocell
Proteinoid	Chemotroph
H-bond donor	Dipole
Electronegativity	Buffer
Amphoteric	Ion product of water
Conjugate acid–base pair	Plasma (cell) membrane

B. Differentiate between the two terms in each of the following pairs:

Cytoplasm/cytosol	Amphoteric/amphipathic
Autotroph/heterotroph	Lysosome/liposome
Brønsted acid/Brønsted base	Phototroph/chemotroph
Aerobe/anaerobe	
Nucleus/nucleolus	

C. (1) Why is the term *evolution* applied to the chemical reactions believed to have occurred during the first billion years after formation of the Earth? What are the three stages into which we commonly divide chemical evolution?
(2) How can cells be classified? What are the major differences between prokaryotes and eukaryotes? How do plant cells differ from animal cells? What are archaebacteria?
(3) What are some of the biochemical implications of the molecular structure of water?
(4) Why are pK'_a values of triprotic acids related as follows: $pK'_{a_1} < pK'_{a_2} < pK'_{a_3}$?
(5) Write the proton dissociation reactions, as customary in biochemistry, for propionic acid (CH_3CH_2COOH) and methylamine (CH_3NH_2). Identify the Brønsted acids and bases.
(6) List the major types of noncovalent interactions and explain how they arise.

PROBLEMS

1.1. Why are soap micelles and NaCl soluble in water?

1.2. A researcher is attempting to prepare "inverted" micelles of soap molecules by dispersing them in benzene together with a small amount of water. Could inverted micelles be formed under these conditions, and, if so, what would their structure be?

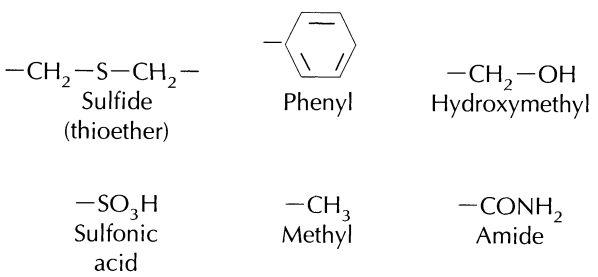
1.3. Consider two interacting charged structures, separated by a distance r. By what factor will their energy of interaction decrease when r is increased to $2r$ if the interactions are of the type: (a) ion–ion; (b) ion–permanent dipole; (c) ion–induced dipole; (d) van der Waals?

1.4. What is the free energy change ($\Delta G°$) accompanying the transfer of 32.0 g of methane (CH_4; MW = 16.0) from water to benzene?

1.5. The compounds *urea* ($NH_2-CO-NH_2$), *sodium dodecyl sulfate* [*SDS*, a detergent; $CH_3-(CH_2)_{11}-OSO_3^-Na^+$], and NaCl can be used to disrupt noncovalent interactions. Which of these compounds do you expect to disrupt electrostatic bonds, hydrogen bonds, and hydrophobic interactions, respectively? What is the mechanism of action for each compound?

1.6. The three atoms comprising carbon dioxide are colinear (O=C=O). Is there a charge separation along each C=O axis? Does the molecule, as a whole, constitute a dipole?

1.7. Which of the following functional groups are polar and which are capable of hydrogen bonding?



1.8. Use the data in Table 1.1 to calculate the mass, in grams, of water, protein, DNA, RNA, and polysaccharide in a cell of *E. coli*. Assume that the cell contains two single strands of DNA, and use 1.0×10^6 for the molecular weight of RNA and 2.0×10^6 for the number of protein molecules. Recall that one mole contains Avogadro's number (6.02×10^{23}) of molecules.

1.9. How do you explain the fact that ammonia (NH_3) and ethanol (CH_3CH_2OH) are very soluble in water?

1.10.* How would you prepare 2.00 liter of a 0.250*M* acetate buffer at pH 4.50 from concentrated acetic acid (CH_3COOH; 17.4*M*) and 1.00*M* NaOH? (pK'_a of acetic acid = 4.76)

1.11. In the acidification of urine, plasma at pH 7.40 enters the kidney; a urine at pH 5.60 is excreted. HPO_4^{2-} and $H_2PO_4^-$ are present in both plasma and urine. Calculate the factor by which the ratio $[HPO_4^{2-}]/[H_2PO_4^-]$ changes from what it is in plasma to what it is in urine.

1.12. What is the pH of pure water at 37°C, given that the ion product of water (K_w) at that temperature has a value of 2.4×10^{-14}?

1.13. An enzyme is assayed at pH 6.8. During the assay, hydroxide ions are released. Which of the following would you choose to prepare a buffer for the enzyme assay? Why?

CH_3COOH/CH_3COO^-	$pK'_a = 4.76$
$H_2PO_4^-/HPO_4^{2-}$	$pK'_a = 7.21$
HCO_3^-/CO_3^{2-}	$pK'_a = 10.25$

1.14.* How would you prepare 250 ml of a 0.100*M* phosphate buffer at pH 7.00 from solid KH_2PO_4 (MW = 136) and K_2HPO_4 (MW = 174)?

1.15.* Calculate: (a) the ionic strength of the buffer prepared in the previous problem (ignore contributions of H^+ and OH^- from water); (b) the number of milliliters of 1.00*M* HCl that can be added to the buffer in Problem 1.14 before its capacity for addition of acid becomes exhausted.

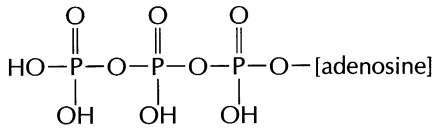
1.16.* A buffer is prepared by mixing 20.0 mmol of solid sodium acetate ($CH_3COO^-Na^+$), 16.0 ml of 1.00*M* HCl, and 750 ml of water. After preparation of the

buffer, 3.00 ml of 2.00*M* NaOH is added. What is the final pH? (pK'_a of acetic acid = 4.76)

1.17. By what factor does the hydrogen-ion concentration change when the pH is changed from 4 to 8?

1.18. Why does the pH scale ordinarily run from 0 to 14? Is it possible for the scale to extend beyond 0 and 14? If so, how?

1.19. Part of adenosine triphosphate (ATP), a key compound in energy metabolism, consists of three phosphate groups linked as shown below. Assume that the ionization of each phosphate group is not affected by the state of ionization of the other two groups. On that basis, indicate whether you would expect the first proton dissociating from each phosphate group to have a pK'_a *approximately* equal to pK'_{a_1}, pK'_{a_2}, or pK'_{a_3} of H_3PO_4. Explain your answer.



1.20.* What is the pK'_a of a monoprotic acid (HA) if, at pH 6.00, 20.0% of the acid is in the ionized form (A^-)?

1.21.* What is the molarity of a monoprotic acid (HA) if a 15.0-ml sample, when titrated to the inflection point, requires precisely 15.0 ml of 0.100*M* KOH?

1.22. Refer to Table 1.7 to decide which is the stronger acid in each of the following pairs: pyruvic/lactic; acetic/formic; succinic (2)/phosphoric (3).

1.23. Explain why it is possible to prepare a number of phosphate buffers that have identical pH values and buffer components but differ in the concentrations of these components, such as the following:

0.010*M* phosphate buffer at pH 7.0
0.20*M* phosphate buffer at pH 7.0
1.0*M* phosphate buffer at pH 7.0

1.24. What is the range of hydrogen-ion concentrations in urine?

1.25. What is the percent ionization of 0.100*M* pyruvic acid at pH 3.00? (see Table 1.7)

Biomolecules

II

Living organisms contain many different kinds of both small and large molecules. Some of these, such as proteins, polysaccharides, and nucleic acids are polymers, composed of hundreds or thousands of low-molecular-weight building blocks. In this part of the book, we examine the structures and properties of the major types of biomolecules. In so doing, we must discuss them not only in light of classical chemical principles, but also by considering their size, shape, and charge—properties integral to their function in living systems.

Amino Acids and Peptides

2

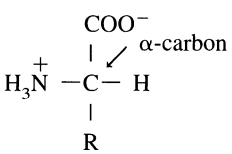
Amino acids are low-molecular-weight biomolecules that serve as building blocks of peptides and proteins. In forming these larger compounds, amino acids become linked covalently, producing chains of varying lengths. We call short chains, consisting of two or more amino acids, **peptides.** Long chains, often containing hundreds of amino acids, are **polypeptides** or **proteins.** Each chain has a unique sequence of amino acids.

Twenty amino acids occur commonly as structural components of proteins. Of these, *glycine* was the first and *threonine* the last to be identified. H. Braconnot isolated glycine from a gelatin hydrolysate in 1820, and W. C. Rose isolated threonine from a fibrin hydrolysate in 1935. In addition to the amino acids found in proteins, other amino acids function as intermediates in the reactions of metabolism.

Amino acids were among the first products identified in experiments on the origin of life. S. L. Miller performed a classic experiment in 1953. He subjected a gaseous mixture of NH_3, H_2O, CH_4, and H_2—components then believed to have constituted the primordial atmosphere—to an electric discharge. After an extended period, significant quantities of organic compounds formed in the reaction mixture, including several amino acids. Similar experiments have since been performed using an atmosphere of different composition in line with current thinking, discussed in Section 1.1.

2.1. AMINO ACID STRUCTURE

As the name implies, amino acids are *bifunctional compounds,* being both *acids* and *amines.* All but one of the amino acids found in proteins have the structure



in which a hydrogen, a carboxyl group, and an amino group are all attached to the same α-carbon atom. The remaining part of the molecule, the *R group,* constitutes the **amino acid side chain.** Under physiological conditions (pH 7.0), both the **α-carboxyl** and the **α-amino** groups are ionized as shown above. By reference to glyceraldehyde, the above structure is that of an L-amino acid (see Appendix B). One amino acid, *proline,* has a different, cyclic structure:

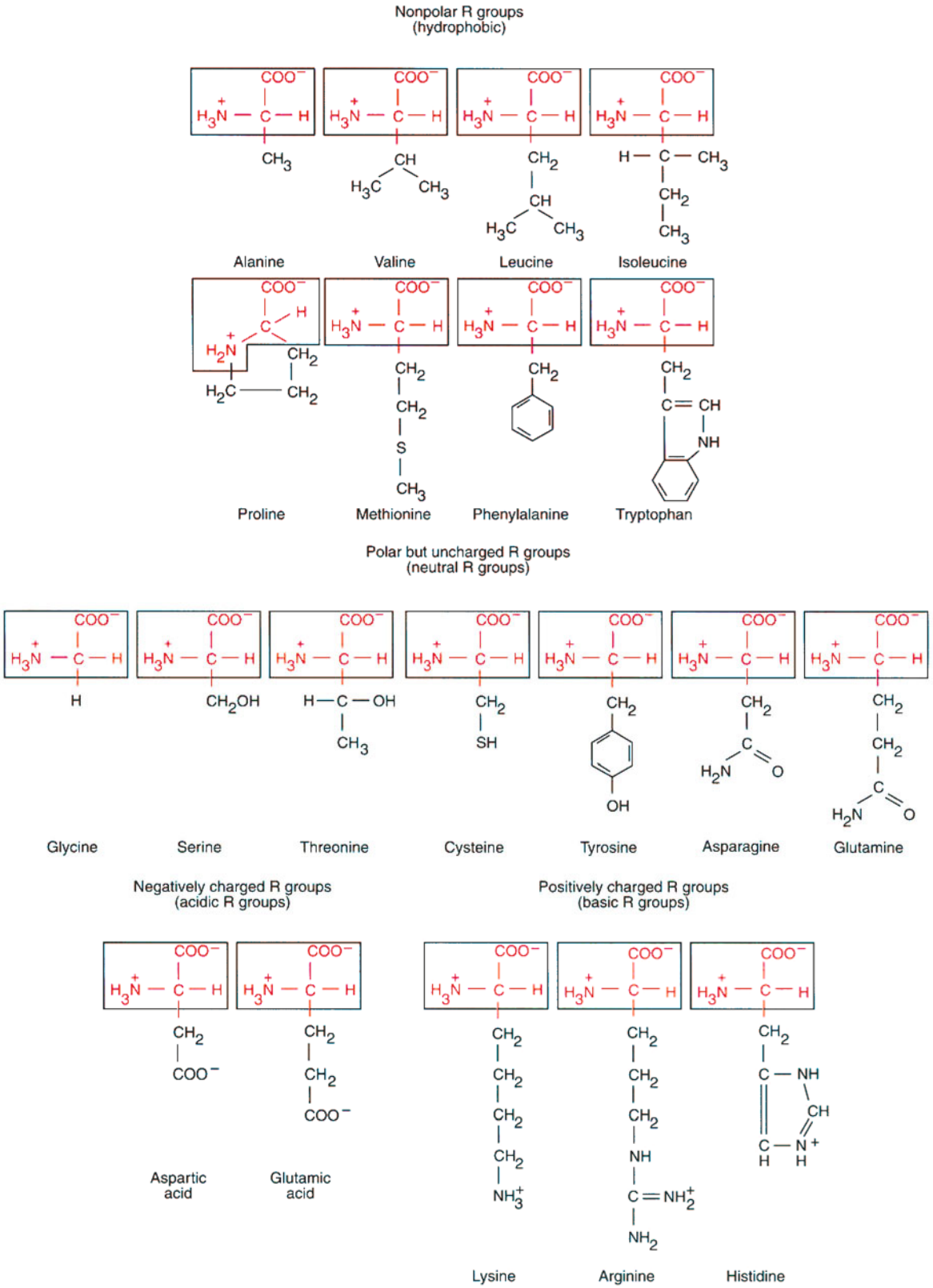
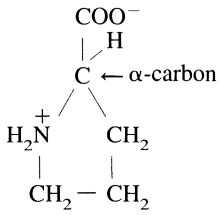


Figure 2.1. Structures of the 20 amino acids found in proteins. At pH 7.0, amino acids exist in the ionized forms shown. Portions outside the boxes represent the side chains or R groups.



2.1.1. Side-Chain Functional Groups

Figure 2.1 gives the structures of the amino acids found in proteins. Each amino acid is abbreviated by a three-letter symbol that, in most cases, consists of the first three letters of its name (Table 2.1). Alternatively, we can designate amino acids by one-letter symbols, which is useful for depicting amino acid sequences of larger peptides and proteins. Of the amino acids found in proteins, all but glycine have one or more chiral carbons and exist as optical isomers. Generally, amino acids of peptides and proteins have the L-configuration.

All of the amino acids have at least two functional groups, an α-carboxyl and an α-amino group ($-NH_2^+-$ for proline). Additionally, many have a third functional group as part of their side chain. The chemical nature of the side chain varies greatly. Some amino acids have an aliphatic side chain; others have one that is aromatic. The side chains of two amino acids contain sulfur, and other amino acids have side chains that contain an alcoholic hydroxyl, a second acidic carboxyl, or a basic nitrogen-containing group. Based on their side-chain structure, amino acids found in proteins include five **aliphatic amino acids** (alanine, glycine, isoleucine, leucine, and valine), three **aromatic amino acids** (phenylalanine, tryptophan, and tyrosine), two **hydroxy amino acids** (serine and threonine), two **sulfur amino acids** (cysteine and methionine), two **acidic amino acids** (aspartic acid and glutamic acid), and three **basic amino acids** (arginine, histidine, and lysine). The remaining amino acids comprise proline, a cyclic amino acid, and asparagine and glutamine, *amide* derivatives of aspartic acid and glutamic acid, respectively.

Some of the side-chain functional groups can ionize (Table 2.2). Recall from Section 1.3 that biochemists consider proton ionizations for both acidic and basic groups in terms of a *loss of protons.* The state of ionization of each functional group depends on its pK_a' (Table 2.1) and on the pH. The nonionizable functional groups of the amino acids are listed in Table 2.3.

2.1.2. Classifying Amino Acids

We divide amino acids into two large groups, **polar** and **nonpolar amino acids** (Figure 2.1). These terms refer specifically to the *nature of the side chain.* Overall, all of the amino acids are polar compounds because each contains at least two ionizable functional groups. We classify amino acids in this fashion because α-COOH and $\alpha\text{-}NH_2$ groups link amino acids together in a peptide or protein, thereby forming uncharged amide groups ($-CO-NH-$). This leaves the side-chain functional groups to

Table 2.1. Amino Acid Abbreviations and pK_a' Values (at 25°C)

			pK_a' value		
Amino acid	Abbreviation		α-COOH	$\alpha\text{-}NH_3^+$	Side chain
Alanine	Ala	A	2.34	9.69	
Arginine	Arg	R	2.17	9.04	12.48
Asparagine	Asn	N	2.02	8.80	
Aspartic acid	Asp	D	2.09	9.82	3.86
Cysteine	Cys	C	1.71	10.78	8.33
Glutamic acid	Glu	E	2.19	9.67	4.25
Glutamine	Gln	Q	2.17	9.13	
Glycine	Gly	G	2.34	9.60	
Histidine	His	H	1.82	9.17	6.04
Isoleucine	Ile	I	2.36	9.68	
Leucine	Leu	L	2.36	9.68	
Lysine	Lys	K	2.18	8.95	10.53
Methionine	Met	M	2.28	9.21	
Phenylalanine	Phe	F	1.83	9.13	
Proline	Pro	P	1.99	10.60	
Serine	Ser	S	2.21	9.15	
Threonine	Thr	T	2.63	10.43	
Tryptophan	Trp	W	2.38	9.39	
Tyrosine	Tyr	Y	2.20	9.11	10.07
Valine	Val	V	2.32	9.62	

Table 2.2. Ionizable Functional Groups of the Amino Acids

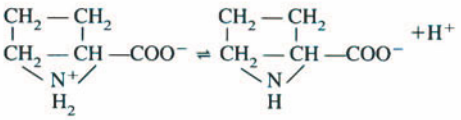
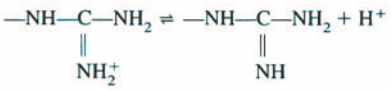
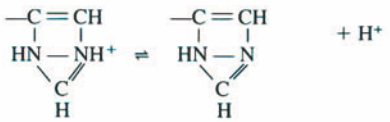
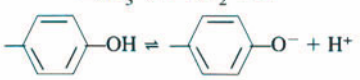
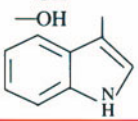
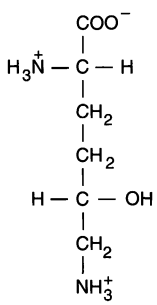
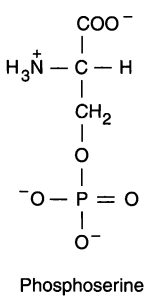
Amino acid	Name	Functional group: Ionization — Low pH ⟵ / High pH ⟶
		Functional groups in basic structure
All	α-Carboxyl	$-COOH \rightleftharpoons -COO^- + H^+$
All but proline	α-Amino	$-NH_3^+ \rightleftharpoons -NH_2 + H^+$
Proline	Pyrrolidine ring	
		Side-chain functional groups
Arginine	Guanido	
Aspartic acid	β-Carboxyl	$-COOH \rightleftharpoons -COO^- + H^+$
Cysteine	Sulfhydryl	$-SH \rightleftharpoons -S^- + H^+$
Glutamic acid	γ-Carboxyl	$-COOH \rightleftharpoons -COO^- + H^+$
Histidine	Imidazole	
Lysine	ϵ-Amino	$-NH_3^+ \rightleftharpoons -NH_2 + H^+$
Tyrosine	Phenolic hydroxyl	

Table 2.3. Nonionizable Functional Groups of the Amino Acids

Amino acid	Functional group: Name	Structure
Asparagine	β-Amide	$-CONH_2$
Glutamine	γ-Amide	$-CONH_2$
Methionine	Thioether	$-CH_2-S-CH_3$
Phenylalanine	Phenyl	
Serine	Alcoholic hydroxyl	$-OH$
Threonine	Alcoholic hydroxyl	$-OH$
Tryptophan	Indole ring	


Hydroxylysine


Phosphoserine

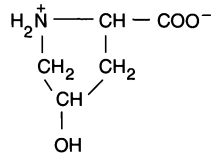

Hydroxyproline

Figure 2.2. Structures of some derived amino acids.

Table 2.4. Human Dietary Amino Acid Requirements

Essential amino acids	Nonessential amino acids
Arginine[a]	Alanine
Histidine	Asparagine
Isoleucine	Aspartic acid
Leucine	Cysteine
Lysine	Glutamic acid
Methionine	Glutamine
Phenylalanine	Glycine
Threonine	Proline
Tryptophan	Serine
Valine	Tyrosine

[a]Essential in young organisms only; arginine is synthesized by humans, but the rate is insufficient to meet the need during growth.

provide most of the polar or nonpolar character of the molecule. Polar amino acids are subdivided into two groups, those having charged and those having uncharged side chains, respectively, at neutral pH.

Polar amino acids with charged side chains consist of two types: *acidic amino acids* have negatively charged groups at pH 7.0, and *basic amino acids* have positively charged groups at the same pH. The imidazole group of the amino acid histidine generally carries a partial positive charge because its pK_a' increases to about 7.0 when histidine becomes incorporated into protein.

At times, some amino acids undergo modification after their incorporation into protein. Figure 2.2 shows three examples of such derived amino acids. Lysine and proline can be hydroxylated to form *hydroxylysine* and *hydroxyproline,* respectively. Serine readily forms *phosphoserine,* in which a phosphate group is esterified at the hydroxyl group of the amino acid.

On the basis of their nutritional value, we categorize amino acids as **essential** or **nonessential amino acids** (Table 2.4). The two terms relate strictly to dietary requirements. All amino acids are "essential" in the sense that they are needed as building blocks of proteins. As used here, the term essential amino acids refers to those that an organism cannot synthesize, or cannot synthesize in sufficient quantity, and that must be obtained through the diet. Amino acids that are essential for one organism must be produced by some other organism(s) for which they are nonessential.

2.2. ACID–BASE PROPERTIES

2.2.1. Dipolar Ions

Because all amino acids contain ionizable functional groups, the net charge of the molecule is a function of pH. Simple amino acids—like glycine, alanine, and valine—have only two ionizable functional groups, an α-COOH and an α-NH_2 group. At low pH (e.g., pH 1), both of these groups are protonated, giving the molecule a net charge of +1. At high pH (e.g., pH 11), both groups lose a proton, resulting in a molecule with a net charge of −1 (Figure 2.3). How does the charge of such a simple amino acid change from +1 to −1 as the pH increases from 1 to 11?

Since the carboxyl group is a stronger acid (has a lower pK_a') than the amino group, it loses its proton first as the pH increases from 1 to 11. This means that an intermediate form of the amino acid is produced that carries *both* a negative ($-COO^-$) and a positive ($-NH_3^+$) charge. We call this form a **dipolar ion** or **zwitterion** (from the German, meaning "hybrid ion"). The dipolar ion is *amphoteric;* it can act as either an acid or a base. When the dipolar ion acts as an acid, or *proton donor,* it is converted to the negatively charged amino acid present at high pH. When it acts as a base, or *proton acceptor,* it is converted to the positively charged amino acid present at low pH. Dipolar ions, like other ions, are *electrolytes.* Accordingly, we also refer to zwitterions as *amphoteric electrolytes,* or *ampholytes.*

We can extend the acid–base relationships discussed for glycine and other simple amino acids to more complex amino acids that carry an additional ionizable group in the side chain. When these groups are protonated, they become either uncharged (e.g., β-COOH, −SH) or positively charged (e.g., ϵ-NH_3^+). Upon dissociation of a proton, the groups become either negatively charged (e.g., β-COO^-, $-S^-$) or uncharged (e.g., ϵ-NH_2). As the pH increases, the fate of these groups parallels that of α-COOH and α-NH_2 groups.

As a result, all amino acids exist as positively

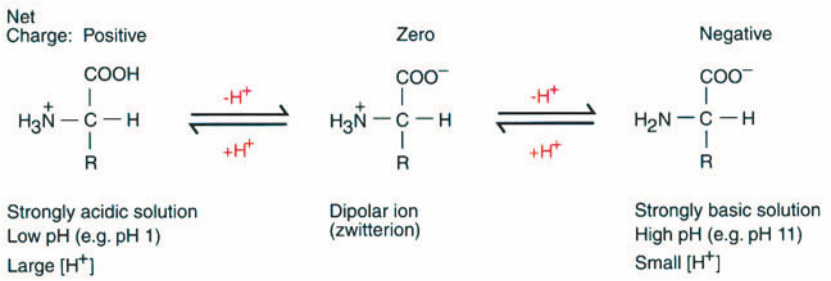


Figure 2.3. Transformation of a simple amino acid, containing only an α-carboxyl and an α-amino group, as a function of pH.

charged molecules at low pH and as negatively charged ones at high pH. Moreover, because peptides and proteins contain the same ionizable groups (only in larger numbers), they are subject to the same pH–charge relationship. Thus, *all amino acids, peptides, and proteins have a net positive charge when the pH is suffiently low and a net negative charge when the pH is sufficiently high.* This generalization applies regardless of the types of amino acids that occur in the molecule or the order in which they are linked. However, the actual pH values at which these states are attained, as well as the precise number of charges, depend on the structure of the molecule.

2.2.2. The Isoelectric Point

The dipolar ion depicted in Figure 2.3, in addition to being amphoteric, has another important property. It has a net charge of zero; the negative and positive charges cancel each other out. The pH at which a molecule has a net charge of zero is termed the **isoelectric point.** We designate this pH as **pI.** All amino acids, peptides, and proteins have isoelectric points. As the pH increases from low to high, these molecules change from cationic forms having net positive charges to anionic forms having net negative charges. Accordingly, they must pass through a state at which they have a net zero charge. That state is the isoelectric point.

In terms of pI, the pH–charge relationship can be stated as follows: *All amino acids, peptides, and proteins have a net positive charge at a pH below the pI and a net negative charge at a pH above the pI* (Figure 2.4).

The isoelectric point of an amino acid always falls halfway between two pK'_a values (see Problem 2.3). Specifically, for

neutral amino acids: $pI = \frac{1}{2}(pK'_{a_1} + pK'_{a_2})$

acidic amino acids: $pI = \frac{1}{2}(pK'_{a_1} + pK'_{a_2})$

basic amino acids: $pI = \frac{1}{2}(pK'_{a_2} + pK'_{a_3})$

We can determine the isoelectric point experimentally by subjecting the compound of interest to *electrophoresis* (see Appendix C) in buffers of varying pH. As long as the pH is above or below the pI, the compound moves either toward the anode or the cathode. The buffer pH at which the compound does not migrate in the electric field constitutes the experimentally determined isoelectric point. The isoelectric point can also be determined by *isoelectric focusing,* a modified form of electrophoresis. In this technique, buffers are used to create a pH gradient in a gel column. A given protein bands in the gradient at a point where the pH of the gradient equals the isoelectric point of the protein.

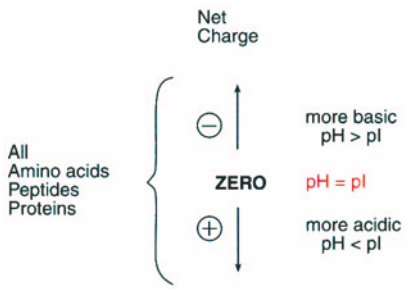


Figure 2.4. Net charge of biomolecules as a function of pH.

2.2.3. Titration Curves

Side-chain ionizable functional groups of amino acids undergo acid–base equilibria like those illustrated in Figure 2.3. Like α-NH_2 and α-COOH groups, side-chain functional groups are amphoteric and can act as either a Brønsted acid or a Brønsted base. Hence, all ionizable functional groups of amino acids can be titrated with either acid or base. Titrating the zwitterionic form of alanine results in the following two reactions:

$$\underset{\text{Base (A}^-\text{)}}{H_3\overset{+}{N}-\underset{\displaystyle CH_3}{\underset{|}{CH}}-COO^-} + HCl \rightleftarrows \underset{\text{Acid (HA)}}{H_3\overset{+}{N}-\underset{\displaystyle CH_3}{\underset{|}{CH}}-COOH} + Cl^-$$

$$\underset{\text{Acid (HA)}}{H_3\overset{+}{N}-\underset{\displaystyle CH_3}{\underset{|}{CH}}-COO^-} + NaOH \rightleftarrows \underset{\text{Base (A}^-\text{)}}{H_2N-\underset{\displaystyle CH_3}{\underset{|}{CH}}-COO^-} + Na^+ + H_2O$$

Similar equations apply to other amino acids and other functional groups. For any particular equilibrium, adding one millimole of acid produces one millimole of the HA form; adding one millimole of base produces one millimole of the A^- form. Each group represents a typical buffer system.

Amino acid titration curves resemble those of any weak acid (see Figure 1.7). Each curve has two or more buffering regions, depending on the number of functional groups in the molecule (Figure 2.5). For any functional group, we calculate the fractions of [HA] and [A^-] present at a particular pH from the Henderson–Hasselbalch equation (Section 1.3).

To titrate an amino acid with base, the starting solution must be acidic (low pH) so that the amino acid is in its fully protonated form and has a net positive charge. As titration proceeds, the charge becomes less positive, changes to zero at the isoelectric point, and subsequently becomes increasingly negative. A case in point is the titra-

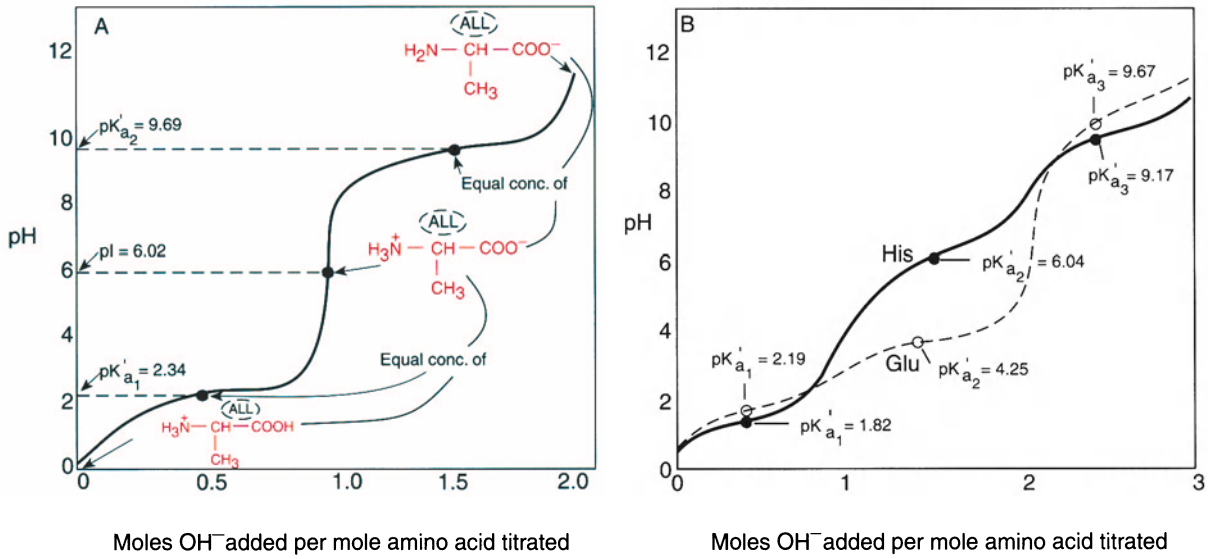


Figure 2.5. Titration curves of amino acids. Fully protonated forms are titrated with NaOH. (A) Alanine; (B) histidine (●) and glutamic acid (○).

tion of *cystine,* the compound formed by linking two cysteine residues via a disulfide bond:

				Net charge
HOOC—CH—CH_2—S—S—CH_2—CH—COOH				
NH_3^+	$\downarrow OH^-$	NH_3^+	pK'_{a_1}(α-COOH) = 1.04	+2
HOOC—CH—CH_2—S—S—CH_2—CH—COO^-				
NH_3^+	$\downarrow OH^-$	NH_3^+	pK'_{a_2}(α-COOH) = 2.05	+1
$^-$OOC—CH—CH_2—S—S—CH_2—CH—COO^-				
NH_3^+	$\downarrow OH^-$	NH_3^+	pK'_{a_3}(α-NH_2) = 8.00	0
$^-$OOC—CH—CH_2—S—S—CH_2—CH—COO^-				
NH_2	$\downarrow OH^-$	NH_3^+	pK'_{a_4}(α-NH_2) = 10.25	−1
$^-$OOC—CH—CH_2—S—S—CH_2—CH—COO^-				
NH_2		NH_2		−2

Note that the two carboxyl groups constitute identical functional groups and have identical locations in a symmetrical molecule. Yet they have different pK'_a values. The same applies to the two amino groups. Why is this so?

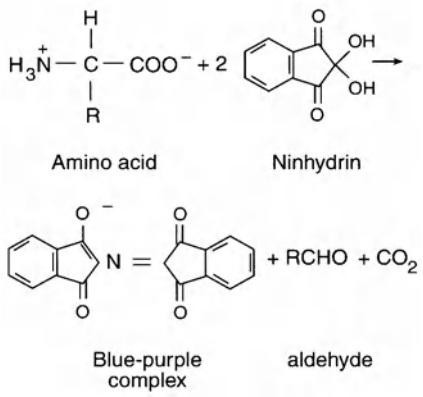
The answer lies in the electrostatic interactions between the proton and the molecule from which it dissociates. We have already seen how such interactions account for the three different pK'_a values of phosphoric acid (Section 1.3). In the case of cystine, the first carboxyl proton dissociates from a molecule that has a net charge of +2. The second carboxyl proton dissociates from a molecule that has a net charge of +1. The second proton is *repelled less* by cystine than the first proton and is more difficult to remove. Hence, $pK'_{a_2} > pK'_{a_1}$. Likewise, the protons dissociating from the first and second amino groups are removed from a molecule having a net charge of 0 and −1, respectively. The second proton is *attracted more* to cystine than the first proton and is more difficult to remove. Hence, $pK'_{a_4} > pK'_{a_3}$.

In nonsymmetrical compounds, larger peptides, and proteins, such electrostatic interactions are more complex. Identical groups may be located in disparate electronic environments and may have quite different pK'_a values. Accordingly, pK'_a values of free amino acids can serve only as a rough guide to those of the same groups in peptides and proteins.

Charge properties of amino acids form the basis for their separation by *ion-exchange chromatography* (see Appendix C). To illustrate the method, assume that we apply a mixture of arginine, glycine, and glutamic acid to a cation-exchange resin at a pH at which the amino acids have a net positive charge but the resin groups will still be negatively charged. Typically, this occurs at a pH of about 2.5. We now proceed to elute the amino acids with buffers of increasing pH. As the pH increases, each amino acid changes from its cationic to its isoelectric, and then to its anionic form. Glutamic acid, having the lowest pI, will be the first to change in this fashion. Arginine will be the last, and glycine will fall in between. Hence, the order of elution will be glutamic acid (first), glycine, and arginine (last).

Analyzing amino acid mixtures by ion-exchange chromatography usually involves reacting the amino

acids in the eluted fractions with the compound ninhydrin. The colored complex formed absorbs strongly in the visible range (at 440 nm for proline, at 540 nm for all other amino acids):



By measuring the absorbance of the eluted fractions, we can make both a qualitative and a quantitative determination of amino acids (Figure 2.6). We identify amino acids by the elution volumes or the fraction numbers of the peaks and determine their amounts by measuring the absorbance of the peak fractions. The entire analysis can be carried out automatically by means of an apparatus called an *amino acid analyzer.*

2.2.5. Estimating Net Charge

As you can see from Table 2.2, protons dissociate from the functional groups of amino acids in one of two ways:

- An ionizable group that has no charge can lose a proton to acquire a charge of -1; this applies to the carboxyl groups (α, β, and γ), the sulfhydryl group, and the phenolic hydroxyl group.
- An ionizable group that has a charge of $+1$ can lose a proton to become uncharged; this applies to the amino groups (α and ϵ), the pyrrolidine ring, the imidazole group, and the guanido group.

With that in mind, you can estimate the net charge of an amino acid as a function of pH by considering the charges of the functional groups in the molecule. The method involves three steps:

1. Write the structure of the amino acid in schematic form, showing the ionizable functional groups in uncharged form, regardless of the pH. Indicate pK_a' values in parentheses.

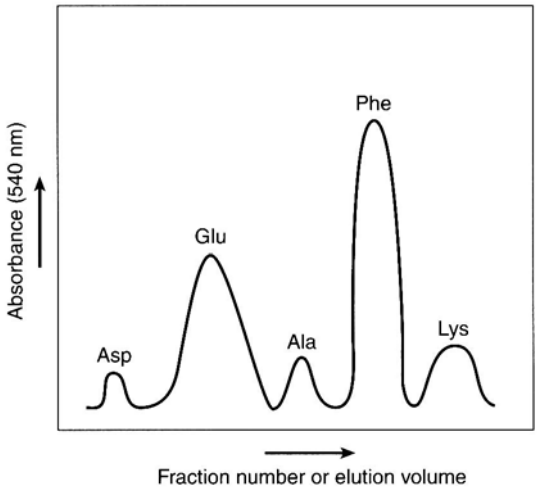
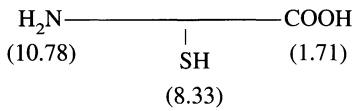


Figure 2.6. Ion-exchange chromatography of amino acids on a cation-exchange resin. Amino acids are applied to the column at pH 2.5 and eluted with buffers of increasing pH.

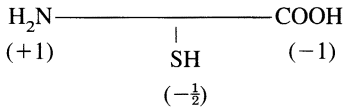
2. Assign charges of -1, $-\frac{1}{2}$, 0, $+\frac{1}{2}$, $+1$ to each group by comparing its pK_a' with the given pH. Since each ionizable group is either a weak acid or a weak base, you can treat its dissociation by means of the Henderson–Hasselbalch equation. Based on that equation, when the pH is 1–2 units removed from the pK_a', an ionizable group exists essentially in one form only, and you can assign it a full positive or a full negative charge. If the pH is near the pK_a', concentrations of undissociated (HA) and dissociated (A^-) forms are essentially equal, and you assign the group a charge of $+\frac{1}{2}$ or $-\frac{1}{2}$.
3. Estimate the net charge of the molecule to the nearest $\frac{1}{2}$ unit by adding algebraically all of the positive and all of the negative charges.

Let us illustrate the method by determining the net charge of cysteine at pH 8.0. We first write the structure of cysteine in schematic form (step 1). In this case, $pK_{a_1}' = 1.71$, $pK_{a_2}' = 8.33$, and $pK_{a_3}' = 10.78$:



In order to assign charges (step 2), we note that pH 8.0 is about two pH units below pK_{a_3}'. Therefore, the amino group exists essentially entirely in its protonated (HA) form and can be assigned a charge of $+1$. In contrast, pH 8.0 is about six pH units above pK_{a_1}'. It follows

that the carboxyl group exists essentially entirely in its dissociated form (A^-) and can be assigned a charge of -1. Lastly, pH 8.0 is practically at the pK_a' of the sulfhydryl group, so that in essentially half of the molecules, that group exists in the dissociated form, and in half of the molecules it exists in the undissociated form. Consequently, $[HA] \approx [A^-]$, and we can assign the group a charge of $-\frac{1}{2}$. On this basis, we rewrite the schematic structure as

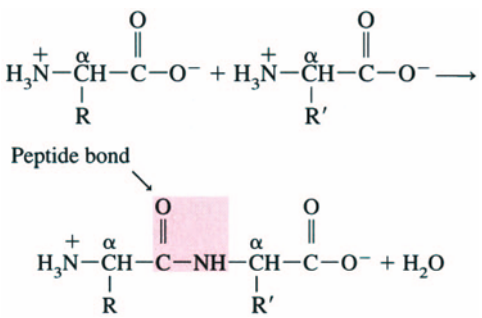


and conclude (step 3) that the net charge of cysteine at pH 8.0 is $-\frac{1}{2}$.

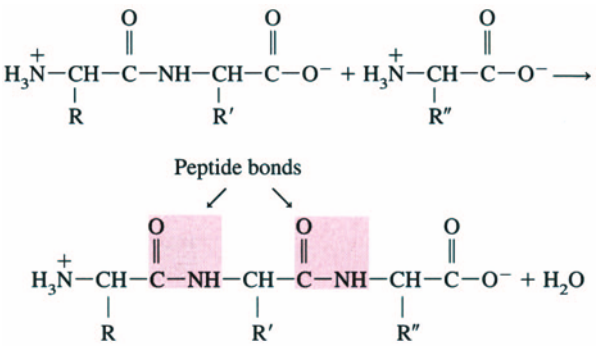
You can use the same method to estimate the net charge of a peptide, assuming that the pK_a' values of its amino acids are identical to those of the free amino acids. Lastly, you can estimate the pI of a peptide by determining the pH range within which the pI must fall. For example, the tetrapeptide Tyr-Ser-Glu-Lys has a net charge of $-\frac{1}{2}$ at pH 9 and a net charge of $+\frac{1}{2}$ at pH 4. Hence, the pI must lie between pH 4 and pH 9, probably near the midpoint of pH 6.5.

2.3. THE PEPTIDE BOND

Peptides and proteins result when amino acids become linked together by means of covalent **peptide bonds.** In living systems, peptide bond formation is part of a multistep process. A peptide bond forms when two amino acids are joined, cleaving out a molecule of water between the carboxyl group of one amino acid and the amino group of a second amino acid:



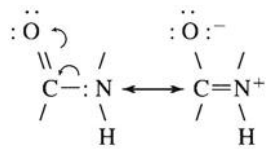
A peptide bond represents an amide bond formed between two amino acids. Linking two amino acids together by means of a peptide bond produces a *dipeptide* (*two* amino acids, *one* peptide bond). When a third amino acid becomes linked to the dipeptide, a *tripeptide* is formed (*three* amino acids, *two* peptide bonds):



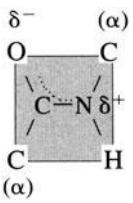
Continuing in this fashion leads to formation of larger peptides (designated with Greek prefixes *di, tri, tetra, penta,* and so on) and ultimately to polypeptides and proteins. We call peptides containing from 2–10 amino acids *oligopeptides,* and those containing over 10 amino acids *polypeptides.* Large polypeptides, if naturally occurring, are *proteins.* The dividing line between polypeptides and proteins is arbitrary. Typically, we consider a polypeptide having a molecular weight of several thousands to be a protein. One of the smallest known proteins is the hormone *insulin* (Figure 2.7), which has a molecular weight (MW) of 5733 and consists of 51 amino acids.

Linking amino acids together by means of peptide bonds produces a **polypeptide chain.** Because of the tetrahedral nature of the α-carbon, the carbon backbone of this chain exists as a zigzag-type structure (Figure 2.8). A protein may consist of one or several polypeptide chains, and these may be linked covalently or noncovalently. Insulin consists of two polypeptide chains, linked covalently via two disulfide bonds.

The four atoms of the peptide bond form two resonance structures



so that the bond is a *resonance hybrid* of these two structures. As a result, the carbon–nitrogen bond has a *partial double bond character,* and the four atoms of the peptide bond, plus the two linked α-carbon atoms, lie in one plane:



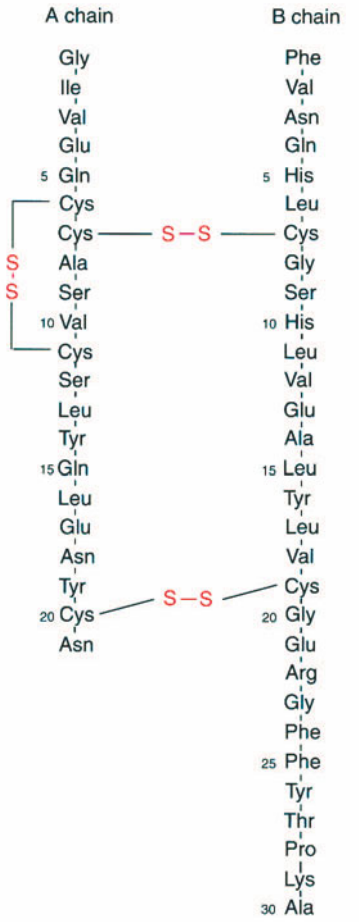
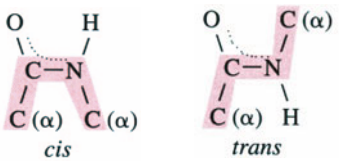


Figure 2.7. Amino acid sequence of the A and B chains of bovine insulin.

Moreover, since configurations of atoms about a double bond yield two *geometrical isomers* (*cis* and *trans*), the peptide bond can have two configurations:



In the *cis* configuration, bulky R groups, attached to the α-carbons, tend to interfere with each other sterically. Consequently, in most proteins, the *trans* configuration is preferred.

We write sequences of amino acids in a polypeptide chain according to a convention such that the remaining free α-NH_2 group is on the left and the free α-COOH group is on the right. We refer to the end of the polypeptide chain containing the free amino group as the **amino-** or **N-terminus;** the end containing the free carboxyl

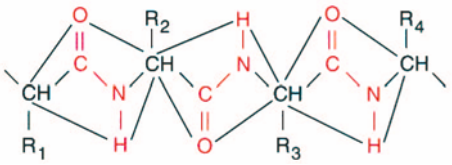


Figure 2.8. The polypeptide chain. The carbon backbone has a zigzag configuration with amino acid side chains protruding from it on both sides in alternating fashion.

group is termed the **carboxyl-** or **C-terminus.** We call the amino acids in the chain (*AA*) **amino acid residues** and number them from left to right:

$$\overset{+}{H_3N}-AA_1-AA_2-AA_3-AA_4-AA_5\ldots\ldots\ldots AA_n-COO^-$$

N-terminal residue — Internal residues — C-terminal residue

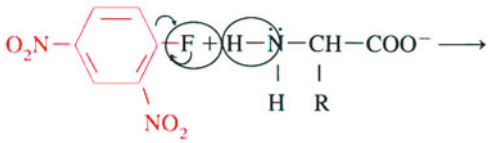
In naming peptides, we proceed, amino acid by amino acid, from the N- to the C-terminus. Names of all residues, except the C-terminal one, end in "yl." Thus, the tetrapeptide glycyl-phenylalanyl-leucyl-tyrosine has glycine as N-terminus, phenylalanine and leucine as internal residues, and tyrosine as C-terminus.

2.4. END-GROUP ANALYSIS

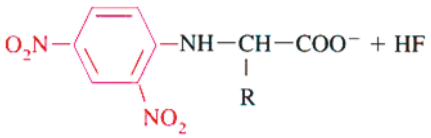
Amino acids undergo various chemical reactions. Of particular interest for the biochemical laboratory are those used in **end-group analysis.** End-group analysis, as the name implies, refers to an analytical determination involving a group located at the end of a molecule, especially a polymer. End-group analysis is not limited to peptides and proteins but can be applied to other biopolymers as well. Fred Sanger, a British biochemist, originally developed the method as part of his effort to elucidate the structure of insulin. Insulin became the first protein to be completely sequenced (1953), and Sanger received the Nobel Prize for his work in 1958.

2.4.1. The Sanger Reaction

Sanger's end-group analysis is based on the reaction of 1-fluoro-2,4-dinitrobenzene (**FDNB, Sanger reagent**) with the α-NH_2 groups of amino acids, peptides, and proteins in alkaline solution. Reaction of the reagent with a free amino acid produces a *dinitrophenyl amino acid,* or **DNP-amino acid (DNP-AA):**



1-Fluoro-2,4-dinitrobenzene
(Sanger reagent)



Dinitrophenyl amino acid
(DNP-amino acid)

In end-group analysis, we first derivatize the protein or peptide by reaction with FDNB, yielding a DNP-protein (Figure 2.9). Acid hydrolysis of the DNP-protein cleaves all peptide bonds but does not break the acid-stable linkage between FDNB and the N-terminal amino acid. We obtain a mixture of free amino acids and one DNP-AA, namely, that corresponding to the N-terminus. We can extract the DNP-AA from the mixture because of differences in solubilities and can identify it by chromatography (see Appendix C). Typically, researchers use paper chromatography because DNP-amino acids have intense yellow colors; unknowns can be identified by visually inspecting the chromatogram and comparing the unknown spot with those of standard DNP-amino acids.

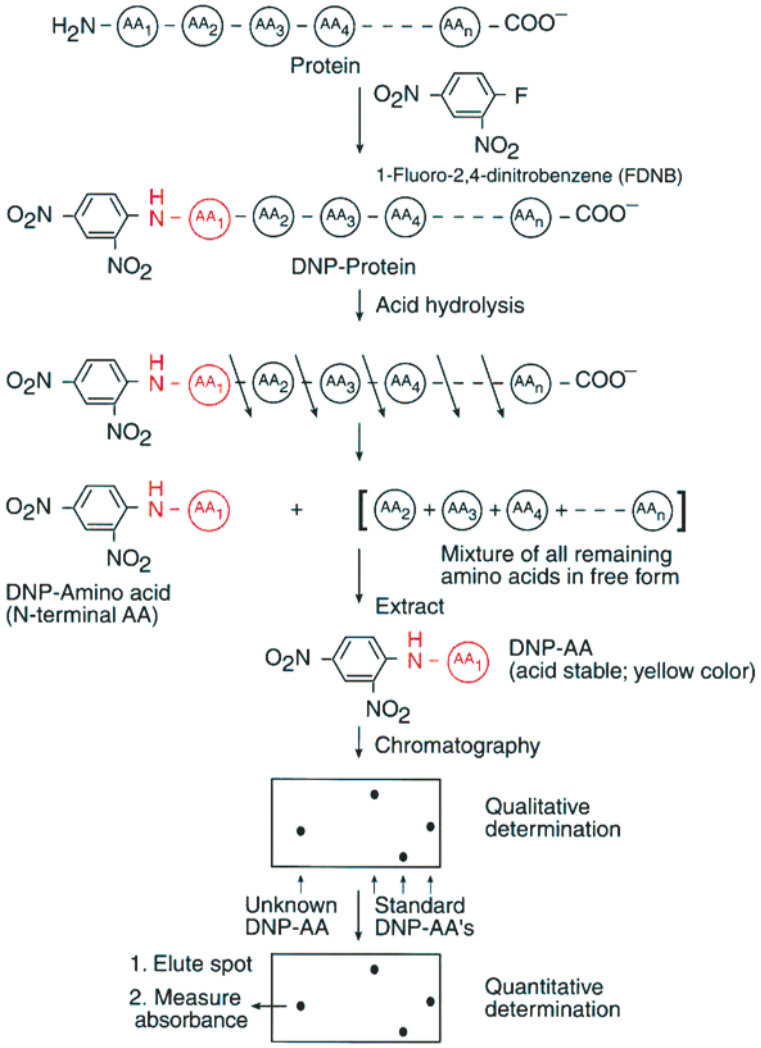


Figure 2.9. End-group analysis of a peptide or protein by means of the Sanger reaction.

Quantitative end-group analysis involves eluting the colored spot from the paper and determining the amount of DNP-AA from the absorbance of the solution (see Appendix C). From these data, the **minimum molecular weight** of the original protein can be calculated, based on the argument that there must be *at least*

one molecule (mole) DNP-AA/one molecule (mole) protein

To illustrate this approach, assume that you obtained X moles of DNP-AA starting with Y grams of protein. If the molecular weight of the protein is M, its mass is M gram/mole, and you have used Y/M moles for your experiment. This number of moles must equal the number of moles of DNP-AA produced so that $X = Y/M$. Knowing X and Y, you can calculate M, the minimum molecular weight. The true molecular weight could be larger but cannot be smaller. Assume, for example, that the protein consists of four individual polypeptide chains, each of which carries the same N-terminal amino acid. In that case, one mole of protein would yield four moles of DNP-AA. Consequently, $X = 4(Y/M)$, and the true molecular weight is the minimum molecular weight, multiplied by four. A popular method for determining true molecular weights of individual polypeptide chains is *sodium dodecyl sulfate-polyacrylamide gel electrophoresis* (*SDS-PAGE*) (see Appendix C). You can calculate minimum molecular weights whenever you can argue that there must be at least one particular functional group, atom, ion, or the like, per molecule. Based on the above, the Sanger end-group analysis provides answers to three questions:

1. What is the N-terminal amino acid of the protein?
2. Are all polypeptide chains in the sample identical with respect to the N-terminal amino acid? (If they are not, the experiment would yield a mixture of DNP-amino acids, resulting in more than one spot on the paper chromatogram.)
3. What is the minimum molecular weight of the protein? (Actually, the calculated value is the minimum molecular weight of the average chain length because polypeptide chains in the sample could have identical N-termini but different chain lengths.)

2.4.2. The Dansyl Chloride Reaction

Because of some disadvantages of the Sanger reaction, such as toxicity of the reagent, it has now been largely replaced by the **dansyl chloride reaction** (Figure 2.10). The

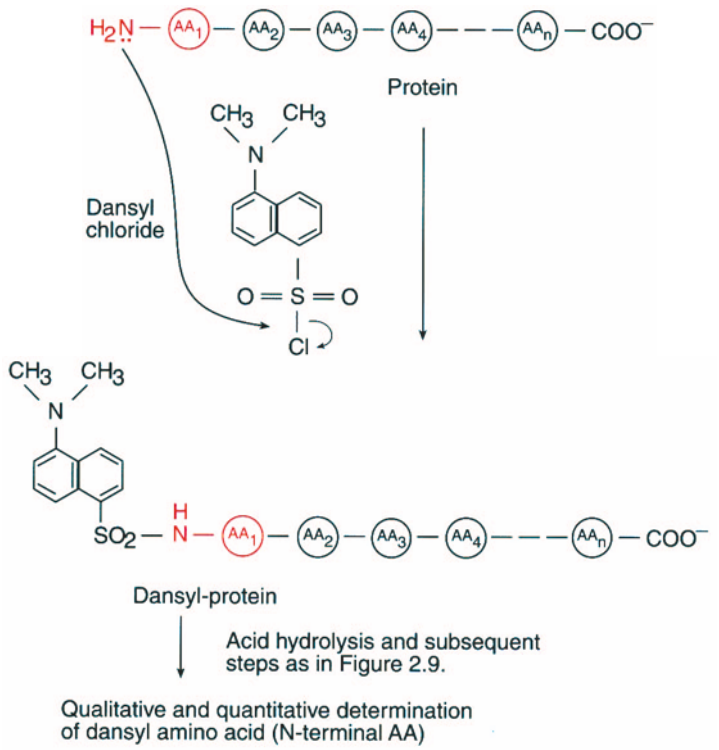


Figure 2.10. The use of dansyl chloride for end-group analysis.

method has great sensitivity because dansyl derivatives of amino acids are fluorescent and can be detected even at very low concentrations. As little as one nanomole of amino acid (one nanomole = 10^{-9} mole) can be determined. In terms of the basic procedure and the information derived, the dansyl chloride reaction is similar to the Sanger reaction.

2.4.3. The Edman Degradation

The Edman degradation constitutes a third type of end-group analysis (Figure 2.11). In this procedure, α-amino groups of amino acids, peptides, and proteins react with *phenylisothiocyanate* (**Edman reagent**). The protein is first treated with the reagent under alkaline conditions, forming a *phenylthiocarbamyl derivative* (*PTC-protein*). Subsequent treatment with strong anhydrous acid cleaves the N-terminal amino acid from the polypeptide chain and cyclizes it to a *thiazolinone* derivative. The great tendency for forming this cyclic compound is responsible for breaking the peptide bond between the N-terminal amino acid and the next amino acid in the polypeptide chain. The remaining peptide bonds, by contrast, remain intact; conditions are insufficient to bring about their hydrolysis. Thus, in addition to the cyclic compound, we obtain a second reaction product that consists of the original polypeptide chain shortened by one amino acid (the N-terminal AA).

We can separate the thiazolinone derivative and the shortened protein because of differences in solubility. Treating the thiazolinone derivative with aqueous acid leads to a rearrangement and formation of a *phenylthio-*

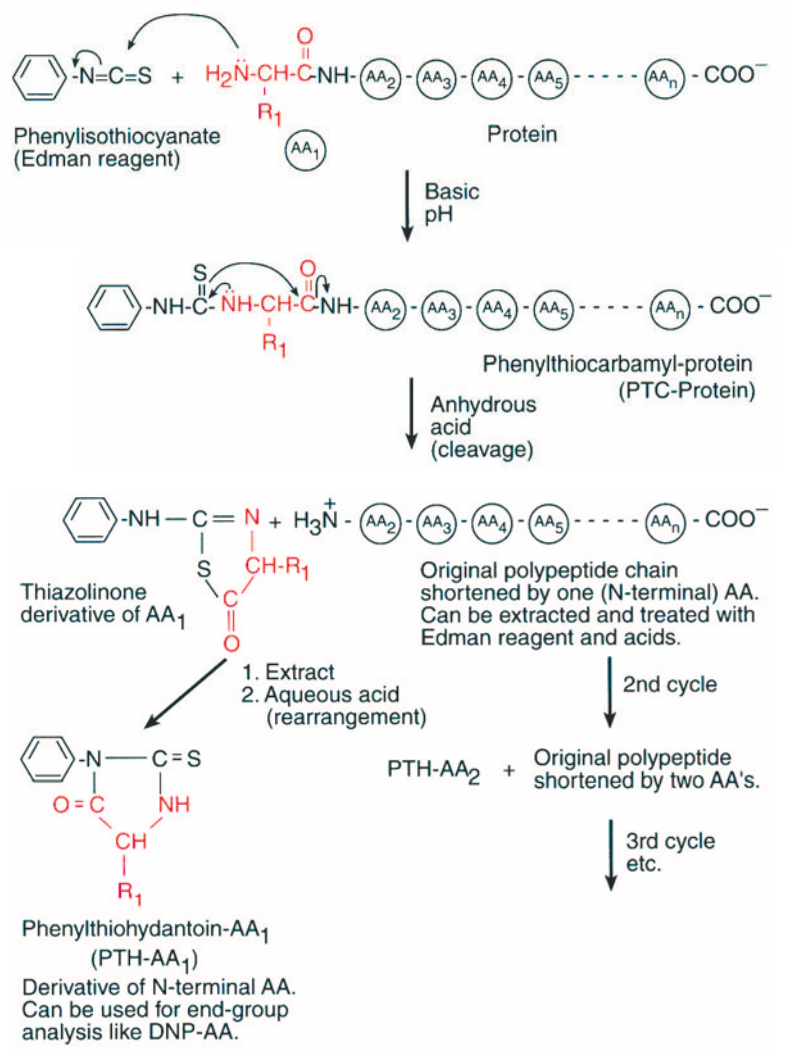


Figure 2.11. The Edman degradation of peptides and proteins.

hydantoin amino acid (*PTH-AA*). You can use a PTH-AA, like a DNP-AA or a dansyl-AA, for end-group analysis.

However, the importance of the Edman degradation is due to its potential to provide even more information than that obtained from an end-group analysis. This is accomplished by repeating the above steps. You can subject the original protein, shortened by one amino acid, to a second cycle of reactions starting with the Edman reagent. The second cycle yields the PTH derivative of the second amino acid in the chain and the original protein, now shortened by two amino acids. You can repeat the process, shortening the polypeptide chain by one amino acid at a time—hence the term *degradation.* Identifying the released amino acids [commonly by *high-performance liquid chromatography* (*HPLC*); see Appendix C] provides the **amino acid sequence** of the polypeptide chain. The entire set of reactions has been automated by means of an apparatus called a *protein sequenator.* Practical limitations of the method result from accumulation of impurities and incomplete reactions, so that a maximum of 40–60 amino acid residues can be sequenced reliably.

2.5. NATURALLY OCCURRING PEPTIDES

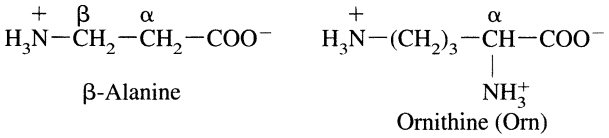
Naturally occurring peptides vary in size from dipeptides, the smallest possible, to large ones containing 20–30 amino acids (Table 2.5). Many peptides contain only the standard amino acids that serve as building blocks of proteins. Some peptides contain modified amino acids, like pyroglutamic acid and glycinamide. Still other peptides

Table 2.5. Some Naturally Occurring Peptides

Name	Structure	Source and function
Carnosine	β-Ala-L-His	Found in vertebrate and human muscle; function unknown
Glutathione	γ-Glu-Cys-Gly	Widely distributed; scavenger of harmful oxidizing agents (antioxidant)
Thyrotropin-releasing factor[a,b]	pyroGlu-His-Pro(NH_2)	Secreted by the hypothalamus; causes the pituitary gland to secrete thyrotropic hormone
Leu-enkephalin	Tyr-Gly-Gly-Phe-Leu	Found in the brain; relieves pain (analgesic)
Met-enkephalin	Tyr-Gly-Gly-Phe-Met	Found in the brain; relieves pain (analgesic)
Vasopressin[b]	NH_3^+—Cys-Tyr-Phe | | S Gln | | S | | Cys — Asn | Pro-Arg-Gly(NH_2)	Secreted by the pituitary gland; increases blood pressure and stimulates reabsorption of water by the kidney (antidiuretic)
Oxytocin[b]	NH_3^+—Cys-Tyr-Ile | | S Gln | | S | | Cys — Asn | Pro-Arg-Gly(NH_2)	Secreted by the pituitary gland; stimulates uterine contraction and ejection of milk from mammary gland
Gramicidin S	L-Val—L-Orn—L-Leu—D-Phe—L-Pro | | L-Pro—D-Phe—L-Leu—D-Orn—L-Val	Produced by *Bacillus brevis;* an antibiotic
Gastrin[a] (human)	pyroGlu—Gly—Pro—Trp—Leu—Glu—Glu—Glu—Glu—Ala—Tyr— Gly—Trp—Met—Asp—Phe (NH_2) | SO_3	Secreted by the stomach; causes the stomach to secrete acid

[a] Pyroglutamic acid (pyroGlu) is a ring structure formed by linking the γ-COOH and the α-NH_2 groups via an amide bond.
[b] (NH_2) refers to formation of an amide at the C-terminal carboxyl group of the peptide.

contain amino acids not found in proteins, such as β-alanine, ornithine (Orn), and D-amino acids:



Peptides vary greatly in their function. For some, like *carnosine,* no precise biological function has yet been determined. Many biologically active peptides serve as *hormones: thyrotropin-releasing factor, vasopressin, oxytocin,* and *gastrin.* Vasopressin and oxytocin possess a cyclic structure because a disulfide bond links two cysteine residues in the molecule. Not all hormones are peptides; some proteins and steroids also function in that capacity (see Section 8.2).

Gramicidin S is an **antibiotic**—a compound produced by a microorganism or a plant, or a close chemical derivative of such a compound, that is toxic to microorganisms from a number of other species. *Gramicidin S* has an unusual structure; it is a cyclic peptide that contains or-

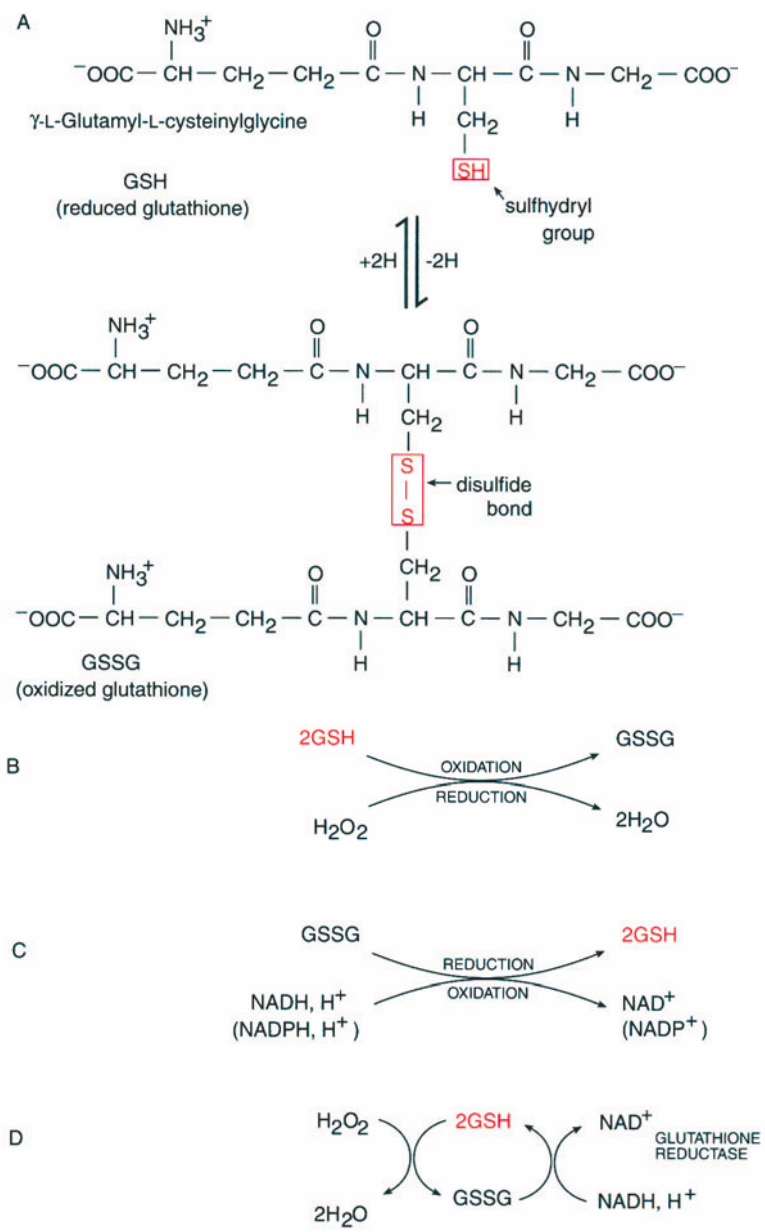


Figure 2.12. Major reactions of glutathione. (A) Oxidation and reduction of glutathione; (B) deactivation of hydrogen peroxide by glutathione; (C) regeneration of reduced glutathione; (D) coupled reactions involving glutathione.

nithine, an amino acid not found in proteins, and several D-amino acids. Scientists believe that the antibiotic activity is due, in part, to these D-amino acids, thereby illustrating the interdependence of structure and function. The proper function of biomolecules depends on specific structural aspects, and slight changes in structure frequently lead to profound changes in function.

The two pentapeptides *leu-enkephalin* and *met-enkephalin* are **neurotransmitter peptides,** naturally occurring peptides that influence nerve transmission in some parts of the brain. There are three groups of neurotransmitter peptides: *endorphins, enkephalins,* and *dynorphins.* All are known as *opioid peptides* because they bind to specific receptors in the brain to which opiates bind. As a result of their binding, opioid peptides mimic some of the pharmacological properties of *opiates* (narcotics that resemble opium in their action). Unlike opiates, however, opioid peptides degrade rapidly after being released and do not accumulate in large enough amounts to induce the kind of tolerance experienced by morphine addicts.

The remaining compound in Table 2.5, *glutathione,* is a tripeptide that readily undergoes oxidation. In the process, two molecules of *reduced glutathione,* or *GSH,* combine by forming an *interchain disulfide bond* between the sulfhydryl groups of the two cysteine residues. The product is *oxidized glutathione,* or *GSSG* (Figure 2.12A). This reaction accounts for the protective action of glutathione in biological systems.

Glutathione acts as a scavenger of harmful oxidizing agents, produced by certain metabolic pathways. These agents include *hydrogen peroxide,* H_2O_2, the *hydroxyl radical,* $OH\cdot$, and the *superoxide anion radical,* $\cdot O_2^-$. If not deactivated, these oxidizing agents can do serious damage to proteins, nucleic acids, and lipids. Reduced glutathione deactivates the oxidizing agents by reducing them, either in the absence or in the presence of an enzyme. Simultaneously, the reduced glutathione is oxidized (Figure 2.12B). Recall that no oxidation takes place without reduction, and vice versa.

Oxidized glutathione is subsequently reduced by means of *glutathione reductase.* This reaction, too, requires that some other substance be oxidized at the same time. The coenzymes nicotinamide adenine dinucleotide, NADH, and nicotinamide adenine dinucleotide phosphate, NADPH (see Section 11.1), serve in that capacity (Figure 2.12C). We frequently describe linked oxidation–reduction reactions by a sequence of curved arrows (Figure 2.12D).

The ease with which glutathione and other SH-group-containing compounds (e.g., *2-mercaptoethanol* and *dithiothreitol*) can be reversibly oxidized to disulfide structures has two experimental applications. First, we frequently add glutathione or other sulfhydryl compounds to reaction mixtures to prevent oxidation of essential SH groups of enzymes and other proteins. In these instances, the added SH compound acts as an **antioxidant;** it, rather than the protein or the enzyme, undergoes oxidation. The second application involves the intentional cleavage of disulfide bonds in proteins by treatment with glutathione or other SH-containing compounds. In these cases, we reduce disulfide bonds in a protein to sulfhydryl groups; reduction of each disulfide bond is coupled to oxidation of an added SH compound to its corresponding disulfide (Fig. 2.13).

As a final note to this chapter, let us look at a synthetic peptide of recent interest, namely, *aspartame.* This dipeptide is an artificial sweetener, approved by the U.S. Food and Drug Administration.

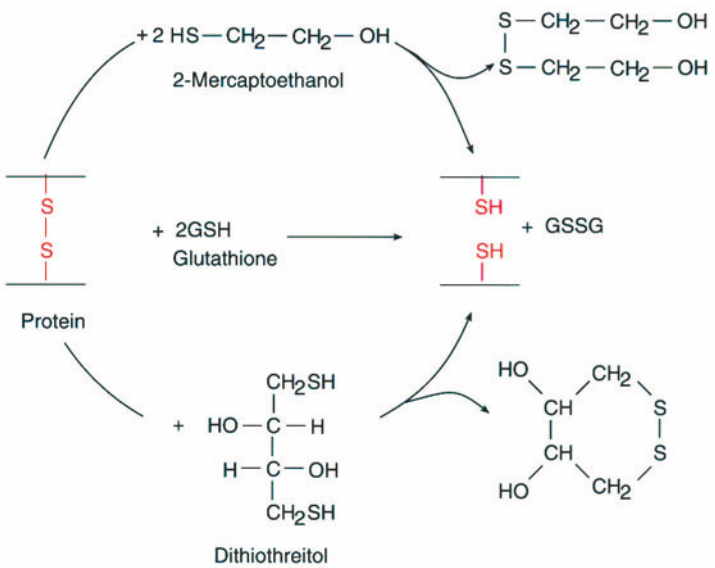
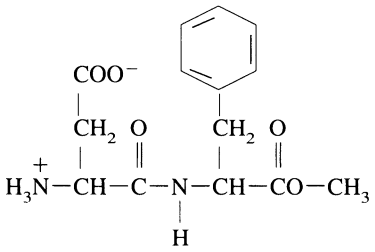


Figure 2.13. Cleavage of disulfide bonds in proteins by treatment with glutathione or other sulfhydryl compounds.



Aspartame (L-aspartyl-L-phenylalanine methyl ester)

Aspartame is marketed under the trademark NutraSweet when used as an ingredient in a product and under the trademark Equal when sold as a sugar substitute. It has a sweetness rating about 200 times that of table sugar (sucrose). Interestingly, substitution of a D-amino acid for either of the two L-amino acids in aspartame produces a derivative that is bitter rather than sweet.

SUMMARY

Twenty amino acids commonly occur as building blocks of proteins. A number of other amino acids serve as intermediates in metabolism. We divide amino acids into two groups, polar and nonpolar amino acids, based on the nature of the side chain. All amino acids have at least two functional groups, an α-COOH and an α-NH_2 group ($-NH_2^+-$ for proline). In addition, many amino acids have a third functional group in their side chain. Some side-chain functional groups ionize and can be titrated with either acid or base. Because of their ionizable groups, the net charge of amino acids changes as a function of pH. At the isoelectric point (pI), an amino acid has a net charge of zero. At pH values below the pI, it has a net positive charge, and at pH values above the pI, it has a net negative charge. Amino acids differ in their isoelectric points and in the dependence of their net charge on pH. These differences permit us to separate them by ion-exchange chromatography.

Amino acids are linked covalently via peptide bonds to form peptides. Oligopeptides contain from 2 to 10 amino acids; polypeptides contain more than 10 amino acids. Large, naturally occurring polypeptides are proteins. We call the string of linked amino acids a polypeptide chain.

End-group analysis involves a reaction of the free α-NH_2 group of peptides or proteins with one of three commonly used reagents: 1-fluoro-2,4-dinitrobenzene, dansyl chloride, or phenylisothiocyanate. End-group analysis provides information about the type and amount of N-terminal amino acid and the minimum molecular weight of the peptide or protein. The reaction with phenylisothiocyanate, when used in repeating cycles, allows us to determine the sequence of amino acids in the polypeptide chain. Naturally occurring peptides vary widely in their size, amino acid sequence, and function. Some peptides serve as hormones, antibiotics, or neurotransmitters.

SELECTED READINGS

Changeux, J. P., Chemical signaling in the brain, *Sci. Am.* 269:58–62 (1993).

Davies, J. S. (ed.), *Amino Acids and Peptides,* CRC Press, Boca Raton, Florida (1993).

Gray, W. R., End group analysis using dansyl chloride, *Methods Enzymol.* 25:121–138 (1972).

Jones, J. H., *Amino Acid and Peptide Synthesis,* Oxford University Press, Oxford (1992).

Konig, W., *Peptide and Protein Hormones: Structure, Regulation, Activity—A Reference Manual,* VCH Publishers, Weinheim (1992).

Meister, A., and Anderson, M. E., Glutathione, *Annu. Rev. Biochem* 52:711–760 (1983).

Nicolas, P., and Mor, A., Peptides as weapons against microorganisms in the chemical defense system of vertebrates, *Annu. Rev. Microbiol.* 49:277–304 (1995).

Sheldon, K., Liu, D., Ferguson, J., and Gariépy, J., Loligomers: Design of *de novo* peptide-based intracellular vehicles, *Proc. Natl. Acad. Sci. USA* 92:2056–2060 (1995).

Stegink, L. D, and Filer, L. J., Jr., *Aspartame—Physiology and Biochemistry,* Marcel Dekker, New York (1984).

Stenesh, J., *Core Topics in Biochemistry,* Cogno Press, Kalamazoo, Michigan (1993).

Takamoto, K., Kamo, M., Kubota, K., Satake, K., and Tsugita, A., Carboxy-terminal degradation of peptides using perfluoroacyl anhydrides—a C-terminal sequencing method, *Eur. J. Biochem.* 228: 362–372 (1995).

A. Define each of the following terms:

Polypeptide chain	Zwitterion
Peptide bond	End-group analysis
Minimum molecular weight	Isoelectric point
Antibiotic	Neurotransmitter peptides

B. Differentiate between the two terms in each of the following pairs (AA = amino acid):

C-terminus/N-terminus	AA side chain/AA sequence
Essential AA/nonessential AA	Polar AA/nonpolar AA
Antibiotic/antioxidant	Sanger reagent/Edman reagent

C. (1) Why are the pK'_a values of functional groups in a peptide or a protein likely to be different from those in the free amino acids?
(2) Explain the principle of the Sanger end-group analysis of proteins. What specific questions can be answered by this method and how are the answers obtained?
(3) Name the side-chain functional groups of the amino acids, and indicate which amino acid carries which group. Write out the mode of ionization for each group.
(4) What conventions are embodied in writing out the amino acid sequence of a polypeptide chain? How is the configuration of this chain affected by the tetrahedral nature of the α-carbon and the planarity of the peptide bond?
(5) Give some examples of naturally occurring peptides, and list some of their properties.

2.1. How would cysteine and tyrosine be grouped if the classification of polar amino acids in Figure 2.1 were based on: (a) pH 4; (b) pH 12?

2.2. Write out the complete structure for carnosine in its fully protonated state (pH 1).

2.3.* The ionization of a neutral amino acid can be written as

$$H_2A^+ \rightleftarrows HA^{\pm} + H^+ \ (K'_{a_1})$$
$$HA^{\pm} \rightleftarrows A^- + H^+ \ (K'_{a_2})$$

Show that, for such amino acids, $pI = \frac{1}{2}(pK'_{a_1} + pK'_{a_2})$. (Hint: Write out the equilibrium constants for the above two equations.)

2.4. The isoelectric point of peptides can be calculated by means of the equation $pI = \frac{1}{2}(pK'_{a_m} + pK'_{a_{m+1}})$, in which m represents the maximum number of positive charges the peptide can have at low pH. On this basis, calculate the isoelectric point of the tetrapeptide Tyr-Ser-Glu-Lys discussed at the end of Section 2.2.5.

2.5. Is there a pH at which an ordinary protein, composed of the standard 20 amino acids, has an *absolute,* rather than *net,* zero charge? In other words, is there a pH at which the protein has all of its functional groups in uncharged form? Explain.

2.6. A tripeptide (Lys-Gly-Asp) is hydrolyzed with sodium hydroxide at pH 12. The resulting hydrolysate is neutralized to pH 7.0, applied to an anion-exchange column, and eluted with buffers of decreasing pH. What is the order of elution of the amino acids?

2.7. Justify the statement "all peptide bonds are amide bonds, but not all amide bonds are peptide bonds."

2.8.* A certain protein is known to consist of two polypeptide chains linked by a number of interchain disulfide bonds (like the two disulfide bonds in insulin, Figure 2.7). A 1.00-g sample reacts fully with 25.0 mg of reduced glutathione (GSH; MW = 307). (a) What is the minimum molecular weight of the protein? (b) How many disulfide bonds occur per molecule if the true molecular weight of the protein is 98,240? (c) How many milligrams of 2-mercaptoethanol (MW = 78.0) would be required for complete reaction with the original 1.00-g sample?

2.9. Using one-letter symbols, write out all of the differ-

ent tripeptides that can be made from the following three amino acids: alanine (A), glycine (G), and proline (P).

2.10. Write out the structures of (a) aspartic acid and (b) lysine as a function of increasing pH, beginning with the fully protonated forms. Indicate which form predominates at pH 6 and which predominates at pH 12.

2.11. A peptide is subjected to end-group analysis using the Sanger reagent. Chromatography yields no spot of DNP-AA. Suggest two possible explanations for this result.

2.12. A peptide derived from a naturally occurring protein contains glycine and several acidic and basic amino acids. Would you expect glycine to have the L- or the D-configuration?

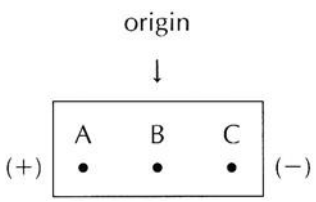
2.13. What is the net charge of the tetrapeptide Asp-Cys-Arg-Lys at: (a) pH 5; (b) pH 14? What is the approximate isoelectric point of the tetrapeptide?

2.14.* Calculate the ratio of the zwitterionic form of glycine to that of the uncharged form, that is, $[^{+}H_3N—CH_2—COO^{-}]/[H_2N—CH_2—COOH]$, at the isoelectric point of glycine. (Hint: Consider the dissociation of cationic glycine to form either the uncharged or the zwitterionic species, and assume that the pK'_a values in Table 2.1 apply to both reactions.)

2.15.* Twenty milliliters of 0.100*M* isoelectric arginine is titrated with 0.10*M* NaOH. What is the pH after: (a) addition of 5.00 ml of NaOH; (b) 75.0% of the guanido group has been titrated?

2.16.* One hundred milliliters of an amino acid solution (0.365 g/100 ml) is titrated with 0.100*M* NaOH. The amino acid is in the fully protonated form. Three distinct inflection points are observed, each requiring precisely 25.0 ml of NaOH for titration. What is the molecular weight of the amino acid?

2.17. The three dipeptides Asp-Phe, Gly-Cys, and Tyr-Lys are separated by electrophoresis at pH 6.0, resulting in the pattern shown. Identify each spot.



2.18. For what pH ranges could alanine serve as a buffer? At what pH values would alanine buffers have maximum capacity for resisting the addition of *either* acid or base?

2.19. At which of the following pH values will proline, aspartic acid, and lysine be separable by electrophoresis? (a) pH 2; (b) pH 7; (c) pH 12.

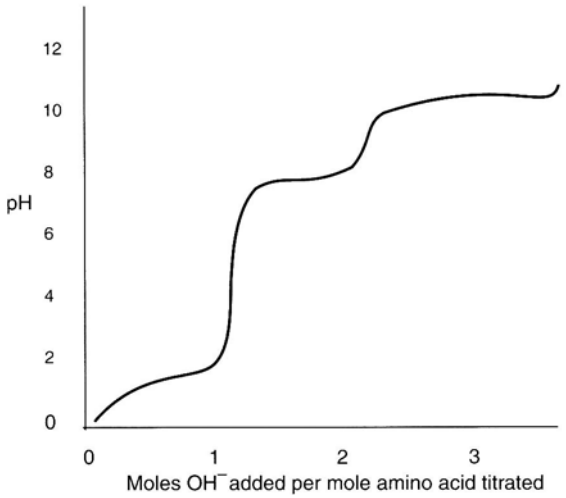
2.20. Classify the following peptides as aliphatic, aromatic, acidic, basic, polar, or nonpolar:

(a) Asn-Ala-Gln	(b) Met-Trp-Trp	(c) Phe-Arg-Glu
(d) Pro-Ser-Tyr	(e) Lys-His-Leu	(f) Ile-Val-Gly

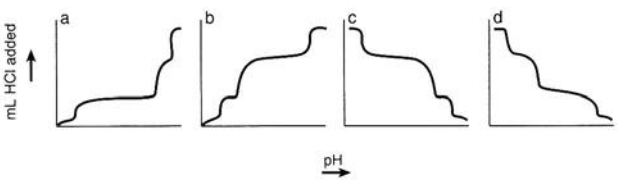
2.21.* Ten grams of a protein is subjected to end-group analysis using the Sanger reagent (FDNB). A single DNP-AA spot is obtained, identified as DNP-Phe, and eluted to yield 5.00 ml of a yellow solution. The solution has an absorbance of 0.600 at 350 nm when measured with a light path of 1.00 cm. The extinction coefficient (see Appendix C) of DNP-Phe at 350 nm is 15.0 $cm^{-1}M^{-1}$. What is the minimum molecular weight of the protein?

2.22. A dansyl amino acid spot on a chromatogram has an R_f of 0.20. If the solvent front is 30.0 cm from the origin, how far is the dansyl amino acid spot from the origin? (See Appendix C.)

2.23. An amino acid titration curve is sketched below. (a) Indicate the pK'_a values and identify the amino acid from them. (b) Indicate the isoelectric point and the buffering regions. (c) Indicate the regions where the amino acid has a net charge of +1, −1, or −2.



2.24. Students are asked to plot the titration curve for the amino acid lysine, using a set of axes reversed from that commonly used. Among the plots submitted are the four sketched below. Is there a correct plot among these four? If so, which one?



Proteins

3

In the first half of the 19th century, the Dutch chemist Gerardus Mulder was investigating the properties of substances extractable from both animal and plant tissues. He found these to contain carbon, hydrogen, nitrogen, and oxygen and believed them to be "without doubt the most important of the known substances in living matter, and without them life would be impossible on our planet." In 1838, at the suggestion of the Swedish chemist Jöns Jakob Berzelius, Mulder named these substances "proteins" (from the Greek, meaning "first" or "foremost").

Mulder's surmise that proteins are of primary importance in biological systems proved to be correct. We now know that proteins are the most abundant of cellular components, constituting more than half of the dry weight of most organisms. It is the variety of functions performed by proteins, however, that makes them so essential for living systems. Proteins run the gamut of biological function, stretching from relatively inert structural components to compounds of high biological activity and specificity such as enzymes, hormones, antibodies, and the like.

A typical cell contains several million protein molecules, representing thousands of different kinds of macromolecules (see Table 1.1). In each protein, a unique amino acid sequence determines the properties of the molecule. Despite their tremendous diversity, all proteins are derived from 20 major amino acids. How can such few building blocks be assembled into so many different proteins? Mathematically, it is quite feasible.

The number of sequences that can be formed from n different objects is *n* factorial (*n*!), where $n! = n(n-1)(n-2)\ldots 1$. Thus, a dipeptide composed of two distinct amino acids (A, B) can have 2!, or $2 \times 1 = 2$, sequences (AB and BA). For a tripeptide of three different amino acids (A, B, C), a total of 3!, or $3 \times 2 \times 1 = 6$, unique sequential arrangements are possible (ABC, ACB, BAC, BCA, CAB, CBA). For a polypeptide of 20 different amino acids, with each occurring only once, the number of possible sequences increases to 20!, or approximately 2×10^{18}. The number of possible amino acid sequences for still larger polypeptides and proteins becomes truly astronomical. For an average protein, composed of 300 amino acids, the number of possible sequences is so large that if only *one*

molecule of each possible sequence were to exist, the total mass of these molecules would far exceed the mass of the Earth. Out of this myriad of possible sequences, several thousand have evolved into the proteins characteristic of a given organism.

3.1. THE NATURE OF PROTEINS

3.1.1. Definition

In terms of their structural organization, we can define **proteins** as *polymers of amino acids, joined via peptide bonds.* Let us examine the different parts of this definition. You have read about the properties of amino acid building blocks and their linkage via peptide bonds in Chapter 2. Amino acids that occur in proteins generally have the L-configuration, and all but glycine are optically active. Peptide bonds form between the α-carboxyl group of one amino acid and the α-amino group of a second amino acid. The peptide bond is planar, and its *trans* configuration is usually preferred for protein structure.

3.1.1A. Occurrence of Peptide Bonds. Early researchers found that the peptide bond constitutes the major, and in most cases the only, covalent bond linking amino acids together in a protein. Titrations of proteins with acid and base revealed that few α-NH_2 and α-COOH groups, but large numbers of side-chain functional groups, were accessible to titration. Unavailability of α-NH_2 and α-COOH groups for titration suggested that these groups served to link amino acids together in the polypeptide chain. This concept was supported by experiments in which the protein was hydrolyzed and aliquots were titrated as a function of time. As hydrolysis proceeded, the number of titratable α-NH_2 and α-COOH groups increased. Moreover, the two types of groups were released in equal numbers. Investigators concluded that, in the intact protein, amino acids were linked by a bond involving the α-COOH group of one amino acid and the α-NH_2 group of a second amino acid.

Additional evidence for the preponderance of peptide bonds in proteins came from using enzymes. When researchers used enzymes known to catalyze hydrolysis of peptide bonds in small peptides, the same enzymes also caused extensive hydrolysis of proteins. Lastly, occurrence of peptide bonds was indicated by the findings that specific colorimetric reagents formed identical colored complexes when reacted with either proteins or model compounds containing CO—NH groups.

These early studies were followed by direct determinations of amino acid linkages in proteins via X-ray diffraction. William Astbury carried out pioneering studies in the early 1930s in England. Astbury found that fibers of hair and wool yielded characteristic diffraction patterns, suggesting a periodicity of structure. Subsequently, Linus Pauling and Robert Corey, working in the United States, analyzed amino acids and peptides. From studies carried out in the 1940s and 1950s, they ultimately deduced the precise structure of the peptide bond.

3.1.1B. High Molecular Weights. Biochemists showed early on that proteins were high-molecular-weight compounds, or **macromolecules** (Table 3.1). One indication came from *minimum molecular weight* calculations (Section 2.4) for *hemoglobin,* the oxygen-carrying protein of the blood. Isolated hemoglobin contained 0.335% (w/w) iron (at. wt. = 55.85), and thus its minimum molecular weight (MW) could be calculated from the equation

$$\frac{\text{MW}(0.335)}{100} = 55.85$$

yielding MW = 16,700. We now know that hemoglobin consists of four polypeptide chains, each surrounding an atom of iron, so that there are actually *four atoms of iron per molecule.* On this basis, the molecular weight of he-

Table 3.1. Molecular Data for Selected Proteins

Protein	Molecular weight (MW)[a]	Number of polypeptide chains
Insulin (bovine)	5,700	2
Myoglobin (horse)	16,900	1
Hemoglobin (human)	64,500	4
Hexokinase (mammalian)	100,000	1
Phosphofructokinase (muscle)	340,000	4
Satellite tobacco necrosis virus (virus coat)	1,300,000	60
Pyruvate dehydrogenase complex (bovine)	8,400,000	192
Tobacco mosaic virus (virus coat)	40,000,000	2130

[a] *Molecular mass* may be expressed in terms of *molecular weight (MW)* or in units of *daltons (D).* The molecular weight is a dimensionless quantity, equal to the ratio of particle mass to that of a ^{12}C atom. A dalton is identical to an *atomic mass unit (amu).* In biochemistry, we use molecular weight and molecular mass synonymously with *particle weight* and *particle mass.* Even for huge aggregates like the tobacco mosaic virus, which contains thousands of molecules, we use the terms molecular weight and molecular mass.

moglobin is 16,700 × 4 = 66,800, which is close to the accepted value (Table 3.1).

3.1.2. Classifications and Functions

Originally, biochemists classified proteins on the basis of their solubility in various solvents and solutions. While you may still see an occasional reference to such solubility properties and while some of the early group names such as albumin and globulin survive, the general nomenclature has been discarded. Currently, we use three classification schemes based on the function, composition, and shape of proteins.

3.1.2A. Specific Functions. When we classify proteins on the basis of their function, we refer to catalytic proteins, storage proteins, transport proteins, and the like. Often, in addition to the categories listed in Table 3.2, we group proteins according to their location or sets of reactions in which they participate. Such groups include blood proteins (proteins in blood that participate in clotting, transport, and other reactions), membrane receptor proteins (proteins that form specific surface sites on membranes), digestive proteins (enzymes catalyzing the degradation of macromolecules), electron transfer proteins (proteins participating in oxidation–reduction reactions), vision proteins (proteins responsible for the biochemistry of vision), and so on.

3.1.2B. Chemical Composition. On the basis of their composition, we place proteins in two broad classes, **simple proteins** and **conjugated proteins.** Simple proteins consist of only a protein component (i.e., amino acids). Conjugated proteins contain some component in addition to amino acids. The additional component may be a metal ion (such as the iron in hemoglobin), an inorganic group (such as the phosphate in phosphoproteins), a low-molecular-weight organic compound (such as a sugar or a lipid), or a high-molecular-weight organic compound (such as a polysaccharide or a nucleic acid). In all cases, the nonprotein component may be tightly bound to the protein or only associated with it loosely. We call a tightly bound nonprotein component a **prosthetic group** and base the designation on the results of *dialysis.* In this technique, one places the protein solution in a bag made of a semipermeable membrane. The bag is immersed in water or an aqueous solution and is agitated. Ions and low-molecular-weight compounds pass through the pores of

Table 3.2. Some Specialized Functions of Proteins

Type and examples	Occurrence or function
Catalytic proteins (enzymes)	
Trypsin	Hydrolysis of peptide bonds
DNA polymerase	Synthesis of DNA
Regulatory proteins (hormones)	
Insulin	Stimulates glucose metabolism
Growth hormone	Stimulates bone growth
Protective proteins	
Antibodies	Combine with foreign proteins
Interferon	Impairs virus replication
Storage proteins	
Casein	Major protein in milk
Ferritin	Iron storage in liver
Transport proteins	
Hemoglobin	Transports oxygen in blood
Myoglobin	Transports/stores oxygen in muscle
Structural proteins	
Collagen	Fibrous connective tissue
Ribosomal proteins	Associated with RNA in ribosomes
Contractile proteins	
Myosin	Thick filaments of muscle
Actin	Thin filaments of muscle
Genetic function proteins	
Histones	Associate with DNA of chromosomes
Repressor	Blocks expression of genes
Toxic proteins	
Ricin	Toxic protein of castor beans
Cholera toxin	Bacterial toxin causing cholera

the membrane, but macromolecules remain trapped inside the bag. If dialysis does not remove the nonprotein component, then we call it a prosthetic group.

Frequently, we describe proteins in terms of their nonprotein component. *Lipoproteins, glycoproteins, nucleoproteins,* and *metalloproteins* refer to proteins that contain lipids, carbohydrates, nucleic acids, and metals, respectively.

3.1.2C. Molecular Shape. Based on their overall shape, we distinguish two large categories of proteins, **globular** and **fibrous proteins.** Globular proteins have roughly spherical or ellipsoidal shapes. They are relatively compact and may consist of a single polypeptide chain, as in *myoglobin,* or of several polypeptide chains, as in *hemoglobin.* The protein may be a simple or a conjugated one and is usually water-soluble. Enzymes are typically globular proteins.

Fibrous proteins have elongated, cylindrical shapes. They, too, may consist of either one or several polypeptide chains. The protein may be a simple or a conjugated one and is usually water-insoluble. Structural proteins of connective tissue, contractile tissue, hair, and skin are typically fibrous proteins.

3.1.2D. General Functions. In addition to their specialized functions, proteins have four general functions: they serve as a source of energy and as a source of nitrogen, they are effective buffers, and they contribute to osmotic pressure.

Degradative metabolism (*catabolism*) of proteins produces roughly 17 kJ/g (4 kcal/g), a yield comparable to that of carbohydrates. However, normally carbohydrates and lipids serve as primary fuels, and proteins are degraded only as a last resort. Proteins provide amino acids for metabolism but do not serve directly as nutrients. In higher organisms, a *sparing mechanism* prevents tissue proteins from being degraded until all of the stored carbohydrate and lipid has been used up for production of energy. This mechanism protects vital organs from degradation.

Most nitrogen that humans and other animals obtain through the diet comes in the form of proteins. Some bacteria and plants can assimilate nitrogen in other ways, but animals get most of their dietary nitrogen from proteins.

Because of the large number of weakly basic and acidic functional groups in the amino acid side chains, proteins have good buffer capacity. Proteins serve as buffers of biological fluids such as blood and cytosol. The bulk of blood buffering is due to the protein hemoglobin, located inside the red blood cells, and to albumin and other proteins present in the plasma, rather than to inorganic buffers like bicarbonate and phosphate.

Osmotic pressure is one of the *colligative properties* of solutions—properties that depend on the *number* of particles and not on their size or shape. Osmotic pressure functions in maintaining cell and tissue structure and in regulating fluid balance. Water-soluble proteins contribute to the osmotic pressure of both intracellular and extracellular fluids.

3.1.3. Isolation, Purification, and Characterization

3.1.3A. Isolation and Purification of Proteins. Before we can determine the structure and properties of a specific protein, we have to isolate and purify it. This amounts to a difficult undertaking that is part art and part science. To put protein isolation and purification in perspective, recall that a typical bacterial cell (Table 1.1) may contain about 3 million protein molecules, representing some 3000 different types. Additionally, the cell contains about 50 million nonprotein molecules. You can now appreciate the magnitude of the task of isolating a few hundred or a few thousand molecules of a specific protein from this mixture. The work requires a great deal of perseverance, ingenuity, biochemical expertise, and luck. The overall approach can be broken down into four basic stages: extraction from source, removal of impurities, protein fractionation, and storage. At every stage, we must take care to use conditions and methods that are as mild as possible so as not to alter the protein. You will read more about that in Section 3.6.

3.1.3A.a. Extraction from Source. After selecting a known or suspected source, you must extract the protein of interest. Sometimes you can do this directly, as in isolating an extracellular protein from blood plasma, but in most cases you have to begin by breaking cells. This you can accomplish by freezing and thawing tissue repeatedly, grinding it with an abrasive, homogenizing it in a blender, or exposing it to high-frequency sound waves (*sonication*). You can also break cells by exposing them to large and rapid changes in osmotic pressure (*osmotic shock*), or by subjecting them first to high pressure in a *French pressure cell* and then exploding them by suddenly releasing the pressure.

When cell breakage is involved, we generally remove tissue fragments, cells, and cell debris by low-speed *centrifugation* (see Appendix C) and obtain a **cell-free extract.** We usually keep the temperature low to prevent degradation of proteins by proteolytic enzymes present in the extract. Assuming that the desired protein is located in the cell-free extract, the next steps entail a workup of this solution.

3.1.3A.b. Removal of Impurities. To purify the cell-free extract, we generally remove nonprotein compo-

nents first. High-molecular-weight contaminants, such as polysaccharides and polynucleotides, may be removed by centrifugation, precipitation, degradation, or adsorption onto a suitable adsorbent. Low-molecular-weight contaminants, such as simple sugars and amino acids, are commonly removed by dialysis.

3.1.3A.c. Protein Fractionation. If you carried out the previous stages successfully, you are now left with a mixture containing from a few to several thousand different types of protein molecules. To isolate one specific type of protein molecule from this mixture, you will use any known or suspected properties of the protein, such as acid–base character, content of specific amino acids, or approximate molecular weight, to aid isolation.

You can remove proteins that are either smaller or larger than the protein of interest by *differential centrifugation,* using various speeds and centrifugal forces. Other techniques include electrophoresis, chromatography, and *fractional precipitation* by adding salt (Section 3.6). Usually, you will apply a number of different experimental approaches before obtaining a *pure* protein—one in which all of the macromolecules are of one kind (e.g., insulin).

3.1.3A.d. Storage. We generally obtain pure proteins in the form of aqueous solutions. In most cases, such solutions cannot be kept for any length of time even in the refrigerator, because of the growth of microorganisms. Instead, the aqueous solution may be frozen and kept in a freezer at −20°C, or even at the temperature of liquid nitrogen (−196°C). At other times, the protein may be precipitated by the addition of acetone, and the precipitate collected, dried, and stored in the form of an *acetone powder.* In most cases, however, the method of choice is *freeze-drying* or *lyophilization.* This involves freezing the protein solution in a round-bottom flask and then evacuating the flask so that the water *sublimes.* Lyophilization effectively removes water from a preparation and is least likely to damage the protein.

3.1.3B. Characterization of Proteins. Once you have isolated and purified a protein, you have to characterize it to understand its function. This process requires many different types of measurements: two fundamental ones are determining the amount and the activity of the protein. Preferred methods for quantitative estimations of proteins use absorbance measurements (see Appendix C). Two common methods employ the *Biuret* and *Lowry* reactions, in which colored complexes form between copper ions and peptide bonds in alkaline solution.

You can carry out various *assays* (determinations) of protein activity by measuring absorbance, mass, radioactivity, or some other property. Frequently, proteins such as enzymes and hormones possess such high activity at low levels that describing their solutions in terms of common measures of concentration is not very useful. In these cases, you will find it helpful to express concentration in terms of defined *units* based on the activity of the molecule (see Section 4.3).

In addition to quantitative determinations and assays of activity, you must determine other chemical, physical, and biological parameters. A chemical description of the protein includes determining its amino acid sequence, acid–base properties, and the like. A physical description consists of determining the protein's size, shape, and related properties. A biological description may entail pinpointing the exact intracellular location of the protein and the various control mechanisms regulating its function.

3.2. PRIMARY STRUCTURE

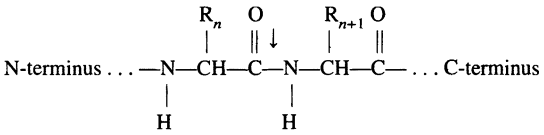
We find it convenient to discuss protein structure in terms of four levels, known as *primary, secondary, tertiary,* and *quaternary* structure. These levels roughly follow the order of increasing molecular complexity. Primary structure corresponds to the covalent structure of the polypeptide chain. Secondary and tertiary structures describe periodic and irregular folding of this chain, respectively. Lastly, quaternary structure consists of the interactions of multiple polypeptide chains.

We define **primary structure** as *the type, number, and sequence of the amino acids in the polypeptide chain.* Only one bond—the covalent peptide bond—functions to maintain the primary structure of proteins; covalent disulfide bonds are a component of the tertiary structure. The primary structure includes the configurations of atoms about the α-carbon, but spatial arrangements of other atoms form part of the secondary and tertiary structures.

3.2.1. Amino Acid Composition and Peptide Maps

3.2.1A. Amino Acid Composition. In order to study any protein in detail, we must determine its primary structure. We can obtain two aspects of the primary structure—types of amino acids present and their relative amounts—from ion-exchange chromatography of a protein hydrolysate. The data provide the **amino acid composition** of the protein, usually expressed in terms of *mole percent (mol %),* that is, the number of moles of an amino acid per 100 moles of total amino acids. If you know the molecular weight of the protein, you can calculate the actual number of amino acids per protein molecule. If you do not know the molecular weight of the pro-

Table 3.3. Specific Hydrolysis of Polypeptide Chains by Endopeptidases



Enzyme	AA$_n$	AA$_{n+1}$
Trypsin	Lys, Arg	*Not* Pro
Chymotrypsin	Phe, Trp, Tyr	*Not* Pro
Pepsin	Phe, Leu, others	*Not* Pro
Thermolysin	*Not* Pro	Phe, Trp, Tyr
Elastase	Ala, Gly, Ser	*Not* Pro

tein, you can calculate its minimum molecular weight by assuming that the amino acid present in smallest amount occurs only once per protein molecule (see Section 2.4).

3.2.1B. Peptide Maps. Determining the amino acid sequence represents the most difficult aspect of elucidating a protein's primary structure. Before discussing this process, let us examine a method that does not yield the amino acid sequence but provides information about it. At times, it is desirable to *compare* amino acid sequences of several proteins without *determining* them. We can do this by constructing **peptide maps** or **fingerprints.** The technique requires partial hydrolysis of each protein into a number of fragments, followed by two-dimensional separation of the fragments using chromatography, electrophoresis, or both.

We can cleave the polypeptide chain by using one or more *proteolytic enzymes (proteases)* or other chemical reactions. Proteolytic enzymes catalyze hydrolysis of peptide bonds; some work on all peptide bonds, whereas others work only on certain bonds, formed from particular amino acids linked in a specific fashion. Table 3.3 illustrates the hydrolysis step of creating a peptide map. The enzymes listed are *endopeptidases,* which catalyze hydrolysis in the *interior of the polypeptide chain.* Other peptidases, called *exopeptidases* (see below), catalyze hydrolysis beginning at an *end of the polypeptide chain.* Polypeptide chains can also be fragmented by means of cyanogen bromide (Br—C≡N), which cleaves the chain at the carboxyl end of a methionyl residue (Figure 3.1).

To prepare peptide maps, we treat each protein with the *same* enzyme(s) and/or with cyanogen bromide. We then separate the mixture of peptides in one dimension,

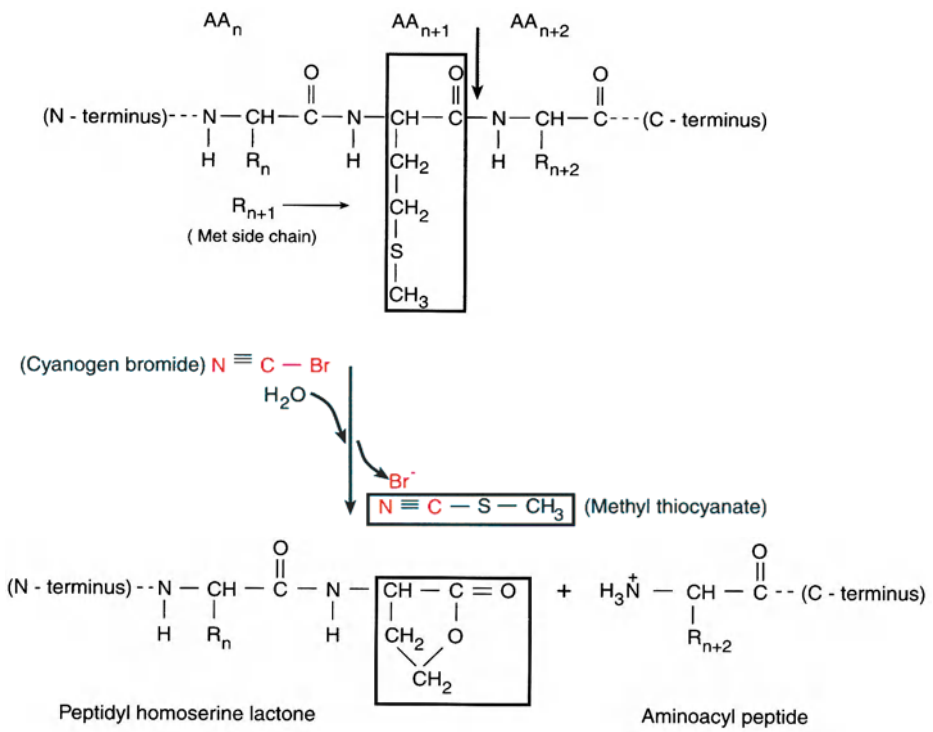


Figure 3.1. Cleavage of polypeptide chains with cyanogen bromide. A lactone is an intramolecular ester, and homoserine describes a compound having one more CH_2 group than serine.

turn the supporting medium (filter paper, gel slab, etc.) by 90°, and separate the mixture again (Figure 3.2) in order to spread out the fragments. After treatment with a color-producing reagent, the fragments form a pattern of spots called a peptide map or fingerprint.

If you compare two unknown proteins that are in fact identical, you will obtain identical peptide maps. The *number* of spots, their *positions,* and their *intensities* will all be identical. If the two proteins differ greatly in their amino acid sequences, they will yield very different peptide maps. Lastly, if the two proteins are similar in their amino acid sequences, the maps produced will also be similar. Thus, peptide maps provide a good indication of the degree to which any two amino acid sequences are similar or dissimilar.

3.2.1C. Sickle-Cell Anemia. The peptide map method was originally developed by Vernon Ingram, a British biochemist, in his studies on hemoglobin. Hemoglobin consists of two α-chains and two β-chains. Ingram isolated the β-chains from both normal individuals and patients afflicted with *sickle-cell anemia.* He obtained two peptide maps (Figure 3.2) that were almost identical, differing in only two spots. By eluting the variant spots and determining the amino acid sequence of the corresponding peptides, Ingram was able to show that the different patterns resulted from replacement of one amino acid in the β-chain. In normal hemoglobin or *hemoglobin A,* glutamic acid occupies position 6; in sickle-cell hemoglobin or *hemoglobin S,* valine occupies this position.

Hemoglobin A	Val-His-Leu-Thr-Pro-**Glu**-Glu-Lys
Hemoglobin S	Val-His-Leu-Thr-Pro-**Val**-Glu-Lys
β-Chain	1 2 3 4 5 **6** 7 8

Replacement of the polar and acidic glutamic acid by the nonpolar valine leads to a remarkable aggregation of deoxyhemoglobin molecules. Aggregation results from noncovalent interactions, mediated in large part by the mutant valine residues (Figure 3.3). The multimolecular aggregates have less affinity for binding oxygen than normal deoxyhemoglobin. In addition, formation of these long linear aggregates inside red blood cells distorts the cells, converting ordinary biconcave cells to ones having a crescent shape. The distorted, elongated red blood cells rupture easily and get stuck within blood capillaries. Both effects further retard delivery of oxygen to the tissues. All these factors combine to produce anemia. Sickle-cell anemia is an example of a *molecular disease,* a disease that can be traced to a change in a single type of molecule.

3.2.2. Determining Amino Acid Sequence

We now turn to the actual determination of amino acid sequences. The general approach for sequencing a protein was originally devised by Sanger in his work on insulin.

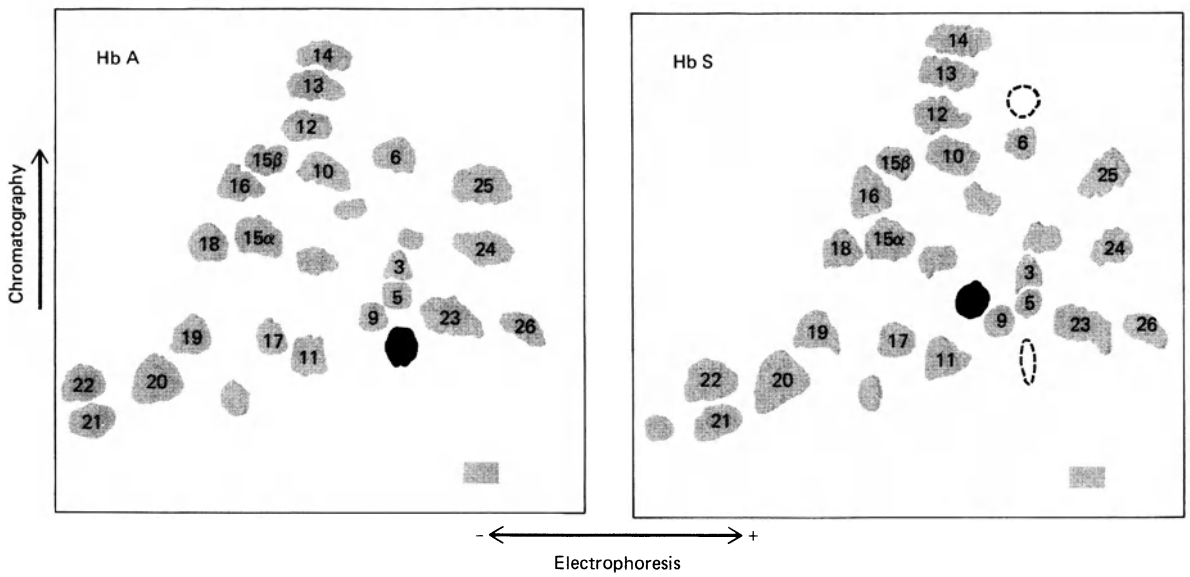


Figure 3.2. Peptide maps of the β-chain from normal (HbA) and sickle-cell (HbS) hemoglobin after digestion with trypsin. The two nonidentical spots (dark spots) represent N-terminal octapeptides that contain glutamic acid and valine, respectively at position no. 6. [Reprinted from C. Baglioni, *Biochim. Biophys. Acta* 48:392–396 (1961) with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV, Amsterdam, The Netherlands.]

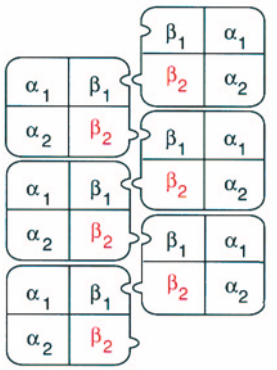


Figure 3.3. Double-stranded polymers formed by aggregation of deoxygenated sickle-cell hemoglobin. Valine 6 from a β_2 subunit in one strand fits into a hydrophobic pocket formed by phenylalanine 85 and leucine 88 in the β_1 subunit of a second strand. [After B. C. Wishner *et al., J. Mol. Biol.* 98:179–194 (1975). With permission.]

That important research (late 1950s, early 1960s), which established definitively that proteins have unique covalent structures, was carried out by a number of scientists over a period of some 10 years, using about 100 g of insulin. Since then, the relevant techniques have been improved tremendously so that nowadays a skilled technician can sequence a protein of similar size in a few days, using just a few micrograms. The sequencing strategy can be divided into three parts, which we will discuss in turn: (1) preparing the protein for sequencing; (2) sequencing the polypeptide chains; and (3) deriving the complete structure.

3.2.2A. Preparing the Protein for Sequencing. Many proteins consist of smaller fragments held together noncovalently. We call such proteins **oligomers,** and the smaller fragments **subunits** or **monomers.** An oligomeric protein can be dissociated into its subunits without cleaving covalent bonds. A subunit may consist of an individual polypeptide chain or of two or more polypeptide chains linked covalently. In hemoglobin, the four chains (two α and two β) constitute four subunits. Insulin, on the other hand, can form a dimer composed of two subunits, with each subunit consisting of two covalently linked polypeptide chains (Figure 2.7).

The material that you plan to sequence must consist of a single type of polypeptide chain. Consequently, you must first separate an oligomeric protein into its subunits. If necessary, purify the subunits to remove nonprotein components. Next, you must separate the subunit into its component parts if it contains more than one polypeptide chain. You do this by breaking **interchain disulfide bonds,** formed between two cysteine residues located in *different polypeptide chains.* You must also break **intrachain disulfide bonds,** formed between two cysteine residues located in the *same polypeptide chain.* Both types of disulfide bonds can be broken either by reduction to sulfhydryl groups (Figure 2.13) or by oxidation to cysteic acid groups ($-SO_3H$). Breaking intra- or interchain disulfide bonds, by either reduction or oxidation (Figure 3.4), produces a compound(s) with new ionizable groups ($-SH$ and $-SO_3H$) and different chromatographic and electrophoretic properties.

Molecular weight, on the other hand, is a property that may or may not be affected by breaking disulfide bonds. Cleaving an intrachain disulfide bond will not change the molecular weight of the protein. The increase in mass due to the change of $-S-S-$ to 2 SH or 2 SO_3H is trivial for a molecule as large as a protein. Cleaving an interchain disulfide bond, however, will produce two molecules, each having a smaller molecular weight than the original protein. Hence, if the molecular weight of an unknown protein decreases upon breakage of disulfide bonds, the presence of *interchain* bonds is indicated. This does not rule out the possibility that the protein may contain *intrachain* disulfide bonds as well.

A final step in preparing the protein for sequencing requires a determination of the amino acid compositions and the molecular weights of the separated polypeptide chains. We have already discussed how to obtain the amino acid composition by ion-exchange chromatography. Molecular weights of peptides and proteins can be determined by various methods, including gel-filtration

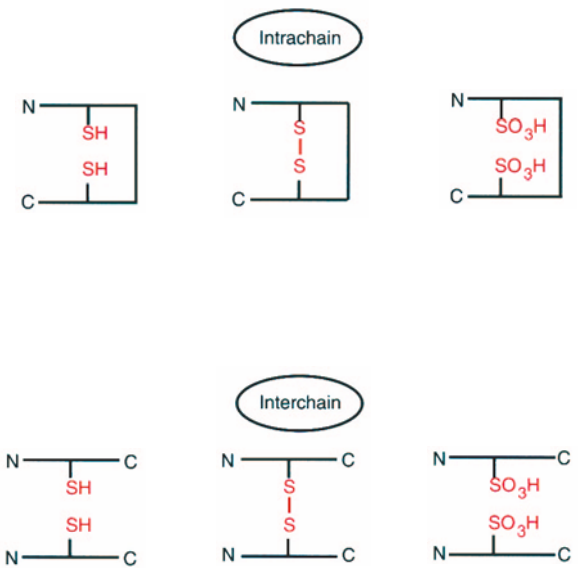


Figure 3.4. Breakage of intra- and interchain disulfide bonds. Interchain bonds can occur in parallel or antiparallel strands, and functional groups may be ionized: pK'_a (SH) = 8.33; pK'_a (SO_3H) = 7.20.

chromatography, electrophoresis, ultracentrifugation, osmotic pressure, and light scattering.

3.2.2B. Sequencing the Polypeptide Chains

3.2.2B.a. N-Terminal Amino Acid. You can identify the N-terminal amino acid of a polypeptide chain using any of the three end-group analysis reactions discussed in Section 2.4. A different approach employs *aminopeptidase,* an enzyme that catalyzes the hydrolysis of all types of peptide bonds and does so by moving from the N- toward the C-terminus (Figure 3.5). The enzyme is an exopeptidase, as distinct from the endopeptidases listed in Table 3.3.

Using this approach, we incubate aliquots of the protein with aminopeptidase for brief periods, precipitate residual protein, and identify amino acids released into the supernatant by chromatography. Suppose that three such experiments yield the following results:

Incubation time (min)	Amino acids detected in the supernatant
0.25	Tyr
1	Tyr, Ala
2	Tyr, Ala, Leu

We can conclude that tyrosine is the first amino acid to be released by enzymatic digestion, followed by alanine and leucine, in this order. Thus, the sequence at the N-terminus must be

$$H_2N\text{-Tyr-Ala-Leu}\ldots$$

You may wonder why no one uses this method for complete sequencing of a protein. To do so would require that the specific hydrolysis proceed with a 100% yield. Unfortunately, that is difficult to achieve experimentally. As time goes on, the amino acid released becomes increasingly contaminated with amino acids released from previously unreacted chains. Interpreting the data becomes impossible. In practice, the N-terminal amino acid and a few subsequent ones in the chain can usually be identified reliably.

3.2.2B.b. C-Terminal Amino Acid. To identify the C-terminus, you can use an enzymatic procedure analogous to that described for the N-terminus. The enzyme *carboxypeptidase* is also an exopeptidase but has opposite specificity to that of aminopeptidase; it moves along the polypeptide chain from the C- to the N-terminus (Figure 3.5). Here, too, we can usually identify the C-terminal amino acid plus a few preceding residues in the chain.

Two other methods exist for identifying the C-terminal amino acid (Figure 3.6). In one, we treat the polypeptide chain with *lithium borohydride ($LiBH_4$).* This reagent *modifies only the C-terminal amino acid residue,* reducing its carboxyl group to an alcohol group. Hydrolysis of the treated peptide yields free amino acids and the amino alcohol of the C-terminal amino acid. The amino alcohol is identified by chromatography. In the second method, we treat the polypeptide chain with *hydrazine* (NH_2-NH_2). This reagent cleaves all of the peptide bonds (*hydrazinolysis*) and *converts all of the amino acid residues, except the C-terminal one, to aminoacyl hydrazides.* The unchanged C-terminal amino acid is identified by chromatography.

3.2.2B.c. Internal Amino Acid Residues. Whereas identifying the N- and C-terminal amino acids is relatively simple, determining the amino acid sequence of the remaining internal residues is much more complex. The essence of the method lies in breaking the original protein down to ever smaller fragments until we obtain peptides of such small size that we can determine their amino acid

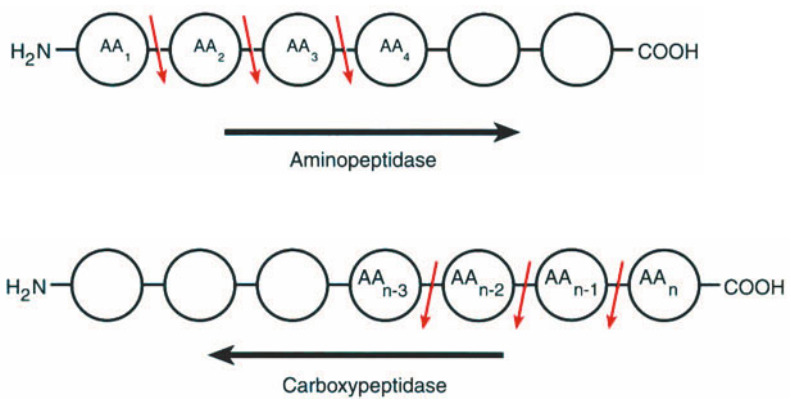


Figure 3.5. Exopeptidase mode of action. Exopeptidases catalyze hydrolysis of all peptide bonds, proceeding from the N- to the C-terminus (aminopeptidase) or vice versa (carboxypeptidase).

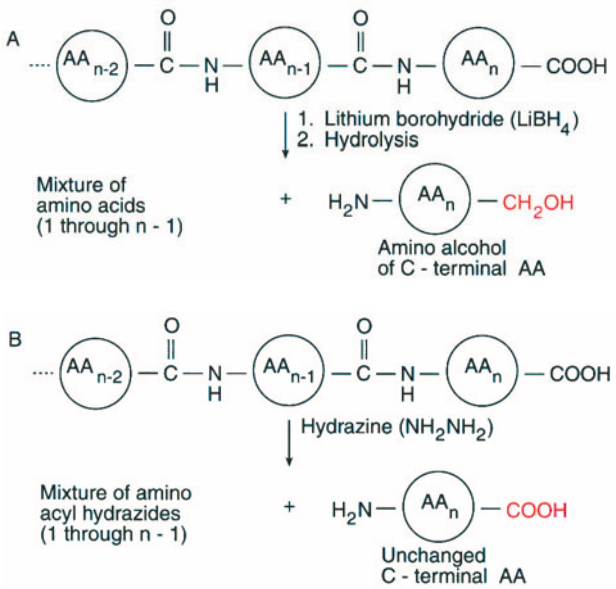


Figure 3.6. Chemical methods for identifying the C-terminus. (A) Reduction with lithium borohydride, followed by hydrolysis; (B) cleavage with hydrazine.

sequence unambiguously. The strategy consists of three stages:

1. *Partial hydrolysis of the protein.* The protein is partially hydrolyzed by treatment with one or more endopeptidases, producing a number of smaller peptide fragments of varying lengths.
2. *Separation of the peptides.* The peptides obtained in Stage 1 are separated by chromatography, typically high-performance liquid chromatography (HPLC), so that they can be studied individually.
3. *Sequencing of the peptides.* Each peptide fragment is fully sequenced. This requires a determination, *for each fragment,* of:

   - Amino acid composition
   - Molecular weight (size)
   - N-terminal amino acid
   - C-terminal amino acid
   - Sequence of internal amino acids

Researchers routinely carry out stage 3 by means of the Edman degradation and an amino acid sequenator (Section 2.4). Recall that this allows sequencing of a peptide containing about 40–60 amino acids. Thus, if you happened to isolate a decapeptide (10 amino acids), determined the N- and C-terminal amino acids, and then sequenced the internal residues by the Edman degradation, the peptide would be fully sequenced.

But suppose that one large fragment contained 85 amino acids. What then? In that case you would subject this fragment again to partial hydrolysis, using different enzymes, to break it down to smaller fragments that, at this point, you hope could be sequenced. If not, you would break these secondary fragments down further. To obtain different size fragments, you must often hydrolyze the original protein or some of its fragments with a number of different proteases, producing *overlapping fragments* (Figure 3.7). For this reason, the procedure is known as the *overlap method.*

3.2.2C. Deriving the Complete Structure. Once you have fully sequenced all of the fragments, you can deduce the amino acid sequence of the original peptide by working backward, lining up the fragments. All of the isolated fragments must fit into the deduced sequence; there can be no exception. Figure 3.7 illustrates this approach with a simplified example in which we deduce the sequence of a 24-amino acid peptide, which has leucine as N-terminus and glycine as C-terminus.

Lastly, you must establish the *number* of inter- and intrachain disulfide bonds and their precise *location* in the molecule. To do this, you subject the native protein, with all of its disulfide bonds intact, to partial hydrolysis with endopeptidases. You identify and isolate fragments containing a disulfide bond. After cleaving the disulfide bonds, you sequence each fragment and locate it within the protein.

A completely different approach to protein sequencing has been made possible by rapid developments in the sequencing of nucleic acids. At present, biochemists can sequence nucleotides in a polynucleotide faster and with greater ease than they can sequence amino acids in a polypeptide. Accordingly, it is becoming increasingly common to deduce the amino acid sequence of a polypeptide from the nucleotide sequence in the DNA gene that codes for the polypeptide. However, direct amino acid sequencing retains its importance because DNA sequencing does not provide information about the number or types of disulfide bonds, does not indicate whether messenger RNA (mRNA) processing occurs, and does not identify amino acid residues modified after the protein was synthesized.

3.2.3. Primary Structure as a Molecular Determinant

Elucidation of the primary structure of a large number of proteins has led to acceptance of the following principle:

> The three-dimensional structure and the functional aspects of a protein are determined by the primary structure of the polypeptide chain. The structure and prop-

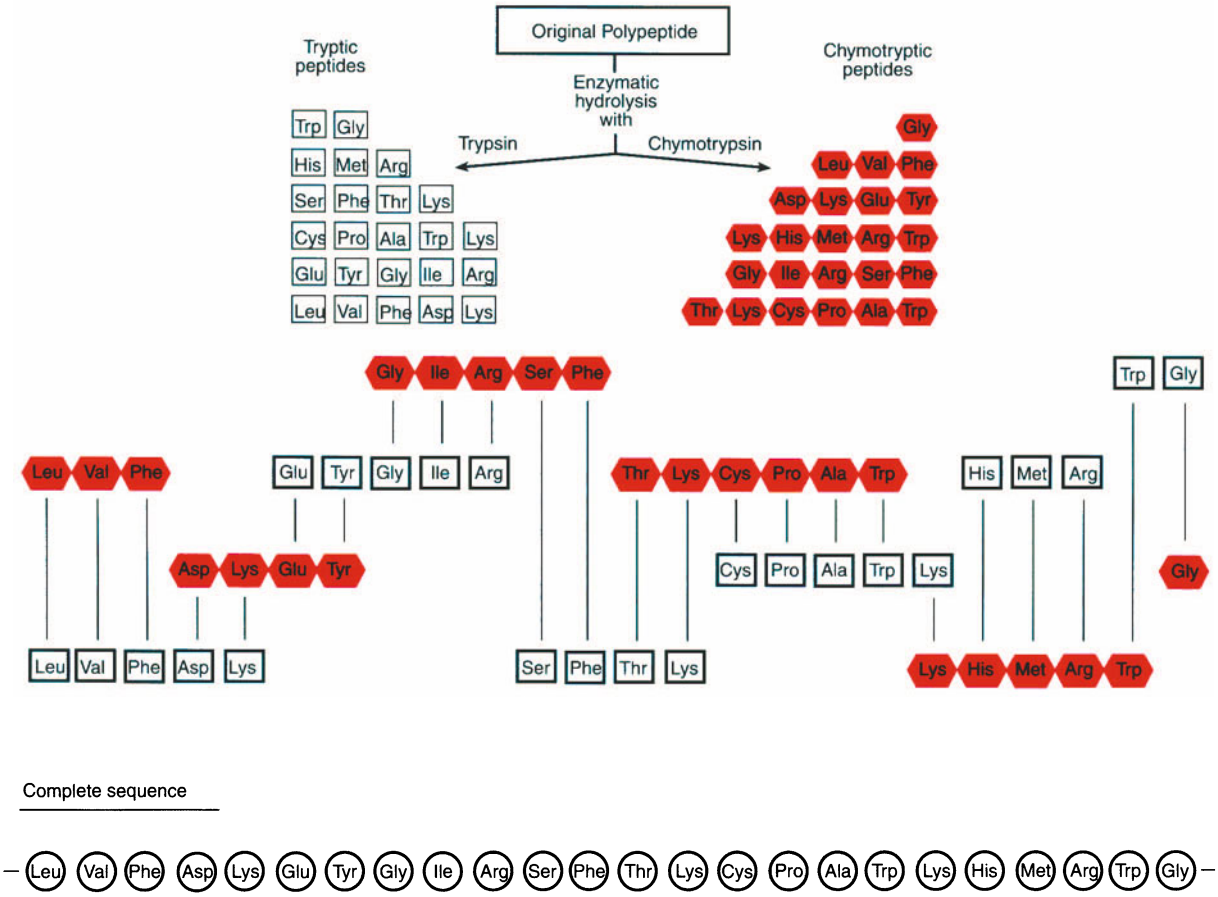


Figure 3.7. Principle of the overlap method for sequencing amino acids. In this example, aliquots of the original peptide have been fragmented by digestion with the endopeptidases trypsin and chymotrypsin.

erties of the protein are effectively "built into" it by virtue of the structure, properties, and sequence of its component amino acids.

This statement illustrates an important general biochemical principle that we encountered already in our discussion of soap micelles. The supramolecular complex of the micelle formed spontaneously as a result of the structure and properties of the component individual fatty acid molecules. Many other large multimolecular aggregates (viruses, ribosomes, multienzyme systems, and the like) assemble from their component parts in a similar manner.

For proteins, the primary structure of the polypeptide chain serves as a major determinant—and for simple proteins (nonconjugated ones) it serves as the only determinant—of the structure and properties of the molecule. Once the amino acids join together in proper sequence, the polypeptide chain folds spontaneously to take on the molecule's characteristic three-dimensional shape.

Laboratory synthesis of polypeptide chains of defined sequences exemplifies this principle. Once produced, a synthetic chain can fold to yield a functional protein. *In vivo,* however, the process is more complex. Protein folding requires the action of specific proteins, called **chaperones,** that help in the proper assembly of protein structure. Several types of chaperones exist; some are enzymes. Chaperones function either by assisting in the correct folding of a polypeptide chain or by preventing improper interactions between parts of one chain or between two different chains (more in Section 19.7).

Studies of the primary structure of proteins have led to several other important generalizations about protein structure and function:

1. No single partial amino acid sequence exists that is common to all proteins. No single peptide segment exists that is essential for the structure of all proteins.
2. Every possible combination of two successive amino acids can occur in nature. There exist no re-

strictions regarding the occurrence of any amino acid prior to, or following, any other amino acid.

3. Proteins that have different functions generally have different amino acid sequences. The amino acid sequence determines the functional properties of the molecule.
4. Proteins that have similar functions have similar partial amino acid sequences. The degree of similarity varies from slight to extensive. Sequence comparisons of such proteins help pinpoint sections that are critical for the function of the molecule.
5. Proteins that have the same function and are derived from different members of the same species have identical amino acid sequences except for the occurrence of mutations. Normal (HbA) and sickle-cell hemoglobin (HbS), isolated from different individuals, illustrate this principle.
6. Proteins that have the same function and are derived from different species have varying degrees of similarity of amino acid sequences. We use the term **sequence homology** to describe the occurrence, in different proteins, of segments having identical amino acid sequences. The greater the sequence homology is, the more closely related the species from which the proteins are derived.

3.2.4. Primary Structure as an Index of Evolution

The point just made suggests that sequence comparisons of the same protein, but isolated from different sources, may be used to evaluate evolutionary relationships of species. We find that **evolutionary trees,** constructed on the basis of biochemical studies involving *a single protein,* are in excellent agreement with those obtained on the basis of a large body of biological evidence. The finding of close relationships among organisms at the molecular level provides strong support for the concept of chemical evolution discussed in Section 1.1.

One protein that has been studied extensively with respect to evolutionary relationships is *cytochrome c*. Researchers have determined the amino acid sequence for cytochrome *c* isolated from 38 different organisms, ranging from yeast to humans, and spanning 1.2 billion years of biological evolution. Cytochrome *c* from all of these organisms consists of a single polypeptide chain that varies in length from 103 to 112 amino acids. Of these, 38 amino acids show *total homology;* they are identical or *invariant* in all species. Additionally, 23 amino acids show *functional homology;* they are replaced by other, functionally similar amino acids. In these instances, one polar amino acid replaces another, one acidic amino acid replaces another, and so on. Amino acid replacements include Ser/Thr, Glu/Asp, Arg/Lys, Val/Leu, and similar substitutions.

Table 3.4 gives the number of different or *variant* amino acids for cytochrome *c* from 25 species. Figure 3.8 shows a plot of such data for three different proteins. These proteins differ not only in their structure and function but also in the degree to which they are essential for overall metabolism and survival of the organism.

Functionally speaking, cytochrome *c* represents the most highly specialized protein of the three in Figure 3.8. Cytochrome *c* has a critical role in the complex array of the electron transport system. Without cytochrome *c*, the electron transport system cannot operate and aerobic metabolism cannot function. Hemoglobin comes next in importance as a deliverer of oxygen to the tissues of eukaryotes. Each hemoglobin chain must interact with three others to form a functional oligomer. Least important are the *fibrinopeptides*—two peptides removed from fibrinogen during its conversion to fibrin as a blood clot forms. On the basis of these comments and the graphs of Figure 3.8, we can draw the following conclusions:

1. Evolutionary changes in a given protein can be described by a linear relationship between the percentage of amino acid replacements and time of divergence of the species (time at which the species separated). From the slope of the line we compute an evolutionary rate, called the **unit evolutionary period (UEP)** and defined as *the time required for the amino acid sequence of a protein to change by 1% after two species have diverged.*

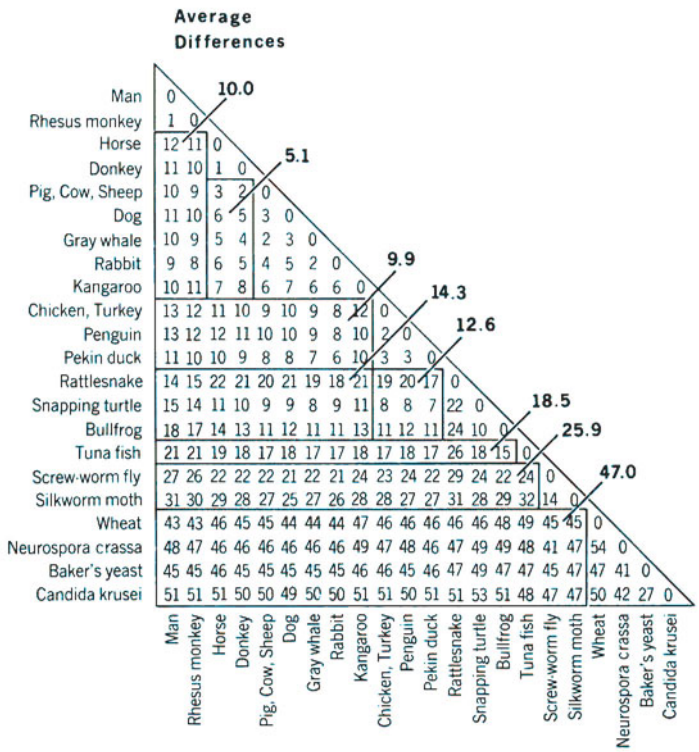
2. The evolutionary rate varies greatly from protein to protein but seems reasonably constant for a given protein, regardless of the organisms from which the protein is derived.

3. The more highly specialized and critical the protein, the longer it takes for a change in amino acid sequence to result in an altered but still functional protein. For essential proteins, it takes a long time to produce an "acceptable" mutation, resulting in a large unit evolutionary period. Of the proteins shown in Fig. 3.8, fibrinopeptides have the smallest and cytochrome c has the largest UEP, with hemoglobin falling in between.

3.3. SECONDARY STRUCTURE

Determination of protein structure involves a great deal more than merely establishing the correct primary structure. Once the amino acids have been covalently linked in their proper sequence, the polypeptide chain must fold

Table 3.4. Amino Acid Difference Matrix for 25 Species of Cytochrome *c*[a,b]



[a]Table copyrighted @ by Irving Geis. Reproduced with permission.
[b]Each table entry indicates the average number of amino acid differences when comparing multiple individuals of the species noted to the left and below that entry. For example, the average difference between the cytochromes *c* of rattlesnakes and rabbits is 18 amino acids.

and twist to generate the final complete three-dimensional structure of the molecule. This requires the elements of secondary and tertiary structure.

We define **secondary structure** as *the regular or periodic folding of the polypeptide chain along an axis.* Secondary structure describes the local spatial arrangement of segments of the polypeptide chain without regard to the conformation of side chains or to the relation of one segment to other segments. The major bond responsible for maintaining the secondary structure is the noncovalent hydrogen bond formed between two peptide bonds. Hydrogen bonds form between the NH group of one peptide bond, serving as hydrogen bond donor, and the CO group of a second peptide bond, serving as hydrogen bond acceptor. Hydrogen bonds can be *intrachain,* formed between two peptide bonds located in different segments of the same polypeptide chain, or they can be *interchain,* formed between two peptide bonds located in segments of different polypeptide chains.

Formation of either intra- or interchain hydrogen bonds generates two basic structural models that were first described by Linus Pauling and Robert Corey in 1951. Development of these models was based on interatomic distances derived from X-ray diffraction patterns and on the following three assumptions:

1. The amide group (CONH) of the peptide bond plus the two linked α-carbons form a planar structure that has a *trans* configuration.
2. Bond angles and interatomic distances in a protein correspond to those found in small organic compounds.
3. Protein structure requires forming the *maximum number* possible of hydrogen bonds. Every CO group must H-bond to an NH group, and vice versa. Pauling and Corey made this assumption in order to describe a structure of greatest stability.

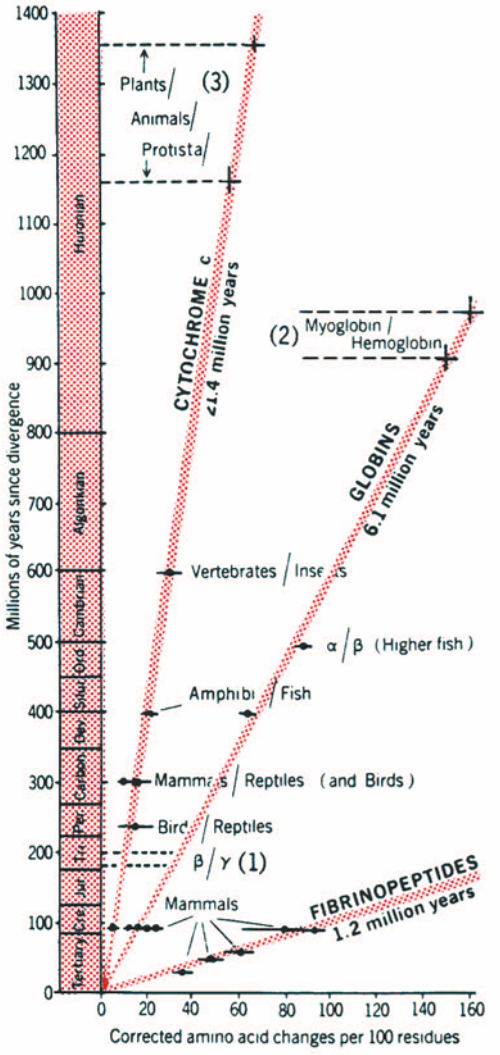


Figure 3.8. Chemical evolution of proteins. A plot of the average difference in amino acid sequence as a function of time (since the species diverged) for three unrelated proteins. The unit evolutionary period, calculated from the slope, is listed under each protein. Numbers in parentheses denote critical branch points in evolution. Standard names of geological periods are abbreviated on the left. (Figure copyrighted © by Irving Geis. Reproduced with permission.)

3.3.1. Intrachain Hydrogen Bonds—the Alpha Helix

In order to form intrachain H-bonds according to the Pauling–Corey model, the polypeptide chain must fold or twist. The resulting configuration constitutes a spiral or coil, like a spring, and is called a **helix.** Theoretically, an infinite number of helices can exist, depending on the diameter and *pitch* of the helix. Pitch is the distance along the helix axis from one point to an identical point: from one crest to another or from one trough to another; the distance for one complete turn of the helix. Helices can be either *right-handed* or *left-handed,* much like the thread of a right-handed or left-handed screw. We call the specific helix formed in a protein by intrachain H-bonds an **alpha helix** (α-helix) (Figure 3.9). It is a right-handed helix that has a pitch of 0.54 nm and 3.6 amino acid residues per complete turn. Every CO group is H-bonded to the NH group of the third amino acid residue behind it in the chain: $AA_4 \ldots AA_1$, $AA_5 \ldots AA_2$, and so on.

A helix can exist in the form of mirror images. A right-handed coil constitutes a mirror image of a left-handed coil. In the case of proteins, a right-handed helix of L-amino acids is a mirror image of a left-handed helix of D-amino acids. Thus, regardless of the optical activity of the component amino acids, a helical structure by itself makes a contribution to the total optical activity of the molecule. The optical rotation of a protein consists of the sum of the rotations of its amino acids and those of its α-helical segments. Accordingly, we can get an indication of the extent of helical structure by determining the difference between the optical rotation of the native protein and that of the amino acid mixture obtained by hydrolyzing the protein. That a helix has optical activity can be demonstrated directly. Of all the amino acids, only glycine lacks a chiral carbon and has no optical rotation. However, polyglycine, a synthetic polypeptide of glycine, forms an α-helix and possesses optical activity. The optical rotation of polyglycine results entirely from the asymmetry of the α-helix.

3.3.2. Constraints on α-Helix Formation

The sequence of amino acids in a polypeptide chain plays a major role in determining whether or not that chain can fold to form an α-helix. There exist four constraints on formation of the α-helix: helix-breaking amino acids, amino acids with bulky side chains, ionic interactions, and planarity and *trans* configuration of the peptide bond.

3.3.2A. Helix-Breaking Amino Acids. The occurrence of proline or hydroxyproline in the polypeptide chain causes a disruption in the α-helical structure. Proline and hydroxyproline contain an *imino* rather than an *amino* group. When the imino group forms a peptide bond, its nitrogen loses its hydrogen atoms and can no longer serve as a hydrogen bond donor. Consequently, the helix is disrupted at this point; proline and hydroxyproline are **helix-breaking amino acids.**

3.3.2B. Amino Acids with Bulky Side Chains. Amino acids that have bulky side chains may disrupt the α-helix. The effect depends on the positions of the amino acids in the polypeptide chain and on the size and shape of their R-groups. If a number of such amino acids (asparagine and leucine, for example) occur in close

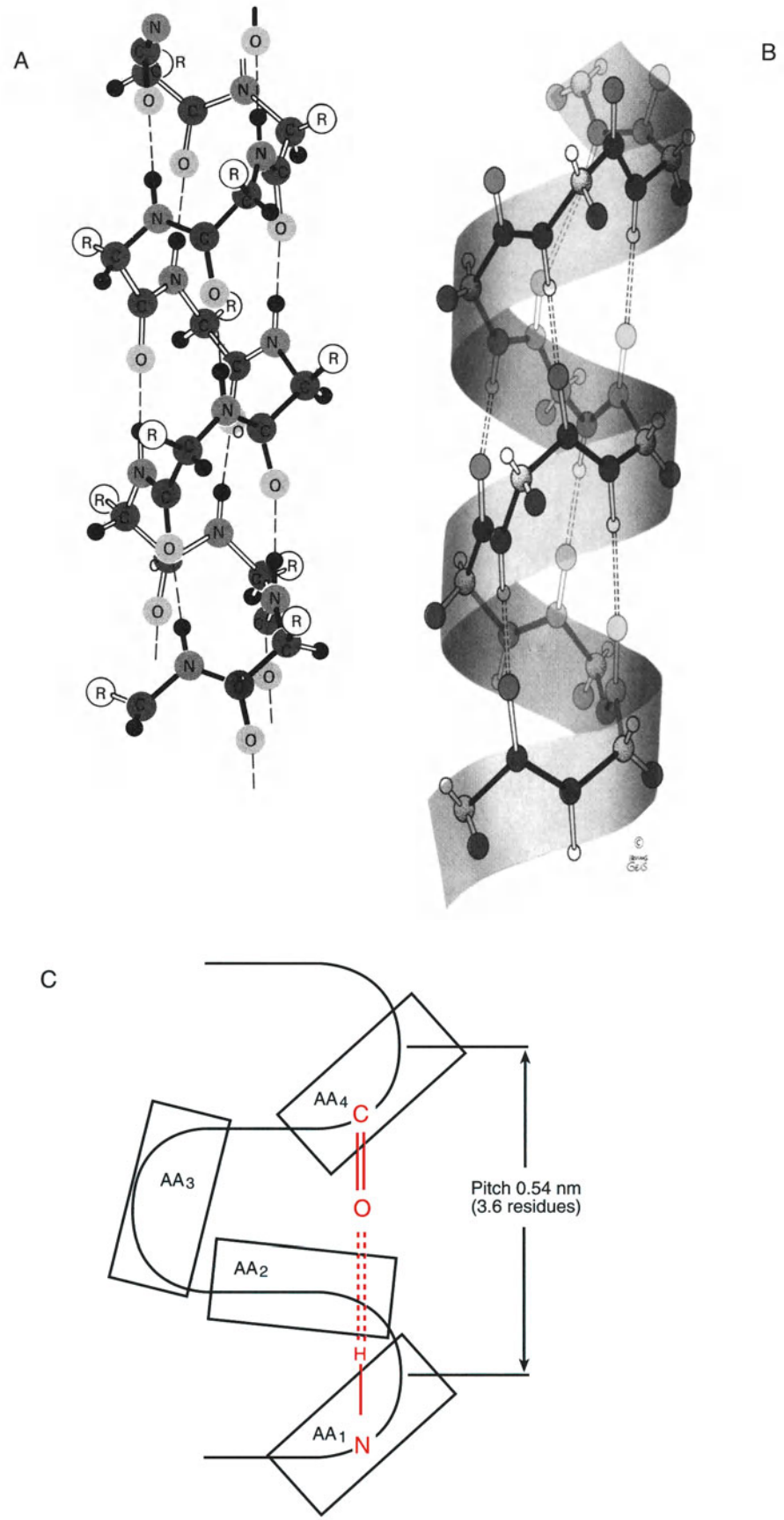


Figure 3.9. The right-handed α-helix. (A) Early illustration. [Reprinted, with permission, from L. Pauling, *The Nature of the Chemical Bond,* 3rd ed., Cornell University Press, Ithaca, New York (1960)]. (B) Recent illustration. (Figure copyrighted © by Irving Geis. Reproduced with permission.) (C) The CO group of each amino acid is H-bonded to the NH group of the third amino acid residue behind it in the chain.

proximity, steric interaction of the R-groups may effectively prevent α-helix formation.

3.3.2C. Ionic Interactions. Ionic interactions among amino acid residues in close proximity may also prevent helix formation. The synthetic polypeptide *polylysine* does not form a helix at pH 7 but forms one at pH 12. At pH 7, the ϵ-NH_2 groups of the lysine residues carry positive charges (pK'_a = 10.5), and the electrostatic repulsion generated by the large number of like charges along the polypeptide chain prevents it from folding into a helix. At pH 12, on the other hand, the ϵ-NH_2 groups carry no charge, and the polypeptide chain folds to form an α-helix. Likewise, the synthetic polypeptide *polyglutamic acid* does not form an α-helix at pH 7 but forms one at pH 2. The γ-COOH groups of the glutamic acid residues carry negative charges at pH 7 (pK'_a = 4.3) but carry no charge at pH 2.

3.3.2D. Planarity and *Trans* Configuration of the Peptide Bond. Planarity of the peptide bond limits rotations of the backbone of the polypeptide chain to those about the C(α)—N and the C(α)—C(1) bonds. We designate the angles of rotation about these two bonds as ϕ (phi) and ψ (psi), respectively (Figure 3.10). We can describe conformations of the polypeptide chain by specifying the ϕ and ψ values for each amino acid residue.

Knowing the van der Waals distances between atoms, we can calculate whether a given combination of ϕ and ψ values results in a possible conformation of two amino acid residues, or whether the structure is impossible due to steric hindrance between nearby atoms. Such calculations have been made for polyalanine. The data can be plotted to yield a diagram, called a *Ramachandran plot* after its inventor (Figure 3.11). You can see that out of the total number of possible conformations, only a small fraction are sterically feasible. These conformations are indicated by the three small shaded regions of the diagram and correspond to left- and right-handed α-helices and to β-pleated sheets (see below). Note that the left-handed α-helix appears to be feasible based on the Ramachandran plot for polyalanine. However, for other amino acids, significant steric interaction exists between

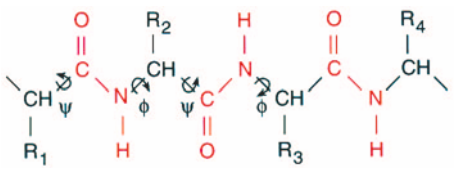


Figure 3.10. Restricted rotation about single bonds in a polypeptide chain. The C—N bond cannot rotate, but the α-carbon to carbonyl carbon (ψ) and the α-carbon to nitrogen (ϕ) bonds can rotate.

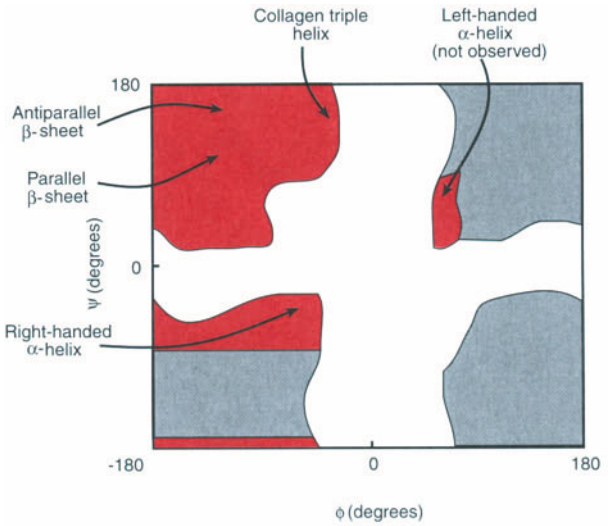


Figure 3.11. A Ramachandran plot showing the allowed values of ψ and ϕ for L-alanine (red). These regions, except for a left-handed α-helix, are allowed for all amino acid residues. Additional regions (shaded) are allowed only for glycine because of its small side chain. [Adapted, with permission, from G. N. Ramachandran and V. Sasissekharan, *Adv. Protein Chem.* 23:283–437 (1968).]

amino acid R-groups so that generally the right-handed α-helix is preferred.

3.3.3. Interchain Hydrogen Bonds—the Beta-Pleated Sheet

We call the Pauling–Corey model of secondary structure generated by forming interchain hydrogen bonds a **beta-pleated sheet,** or simply a **β-sheet.** As in the case of the α-helix, this model requires forming the maximum number possible of H-bonds: every CO group in one chain is H-bonded to an NH group in another chain, and vice versa (Figure 3.12). Much as the helical model yields two structures (a right- and a left-handed helix), so pleated sheets can occur in two forms, a *parallel* **β-sheet** and an *antiparallel* **β-sheet,** depending on the directions in which the polypeptide chains run. However, whereas only the right-handed α-helix occurs in proteins, both parallel and antiparallel β-sheets serve as structural components. Both types of sheet structure represent allowable conformations on the basis of Ramachandran plots (Figure 3.11).

3.3.4. Reverse Turns

The confines of globular proteins dictate that helices and β-pleated sheets cannot extend indefinitely. Turns must be an integral part of the folded polypeptide chain. Some of these turns are very sharp, abruptly reversing the direction in which the chain runs by almost 180°. Such turns, called

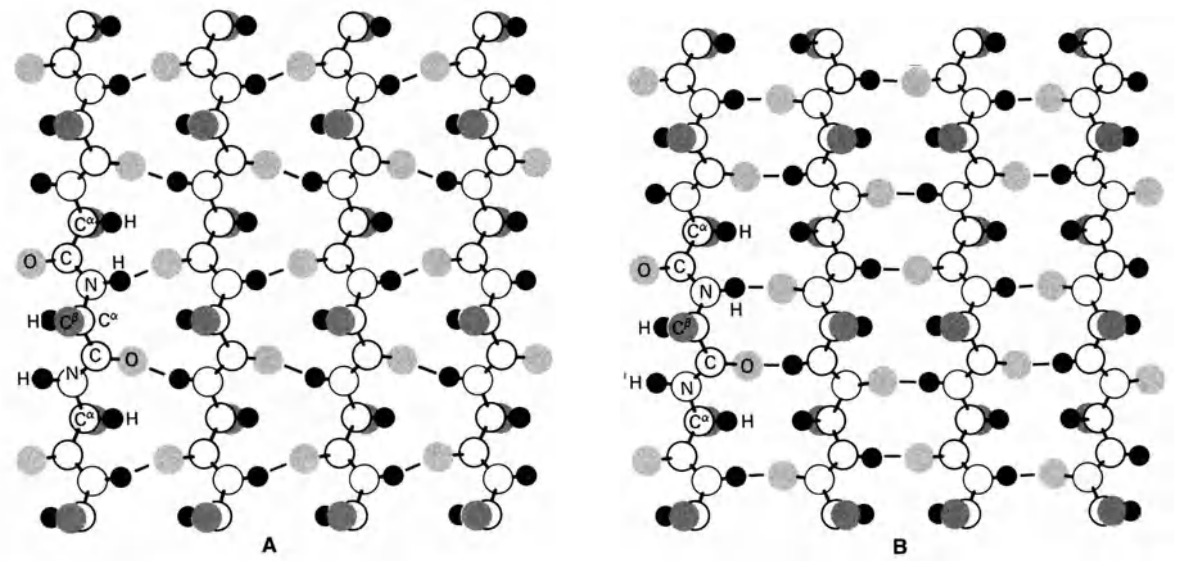


Figure 3.12. Structures of parallel (A) and antiparallel (B) β-pleated sheets. [Reprinted, with permission, from L. Pauling, *The Nature of the Chemical Bond,* 3rd ed., Cornell University Press, Ithaca, New York (1960).]

reverse turns, are largely responsible for giving globular proteins their compact spherical structure. We also call these turns **β-bends** because they frequently connect successive strands of antiparallel β-pleated sheets.

A reverse turn usually consists of four successive amino acid residues. It often includes glycine, proline, and polar amino acids, and frequently a hydrogen bond stabilizes the turn (Figure 3.13). Glycine constitutes part of the turn because its small size (R group = H) presents little steric interference. Proline occurs in the turn because its cyclic structure naturally contributes to a change in direction of the polypeptide chain. Lastly, polar amino acids participate because in globular proteins polar amino acids tend to be located preferentially at the surface of the molecule, where reverse turns generally occur.

3.3.5. Supersecondary Structures

X-ray diffraction has revealed that, in addition to the α-helix and the β-pleated sheet, proteins contain structural components called **supersecondary structures.** These structural components do not occur with the regularity and periodicity of a helix or a pleated sheet. Accordingly, they straddle the dividing line between *secondary* and *tertiary* structure. Figure 3.14 shows a number of the more common supersecondary structures.

3.4. TERTIARY STRUCTURE

Whereas the secondary structure involves periodic folding of the polypeptide chain, the tertiary structure consists of folding that is not regular. Forming the **tertiary structure** requires that *a polypeptide chain that already contains folded α-helical and β-sheet segments be folded once more, but in an irregular manner.* As such, the tertiary structure represents the arrangement in space of all the atoms of a protein or of a subunit. It constitutes the specific three-dimensional structure of the molecule. Tertiary structure results from the interaction of amino acid side chains, located either close or far apart along the chain.

We consider disulfide bonds and nonprotein components as part of the tertiary structure. Even though the

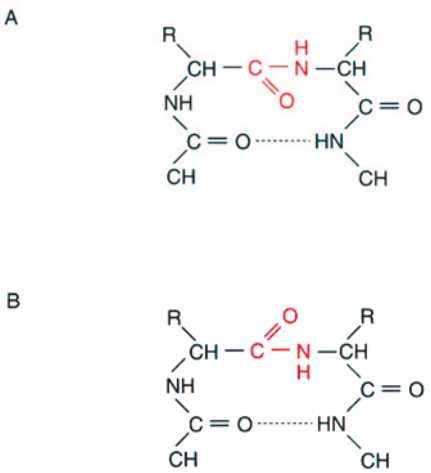


Figure 3.13. Two types of reverse turns, each consisting of four amino acid residues, with a H-bond between residues 1 and 4. The two types differ by a 180° flip of the peptide bond linking residues 2 and 3. [Adapted, with permission, from P. Y. Chou and G. D. Fasman, *J. Mol. Biol.* 115:135–175 (1977).]

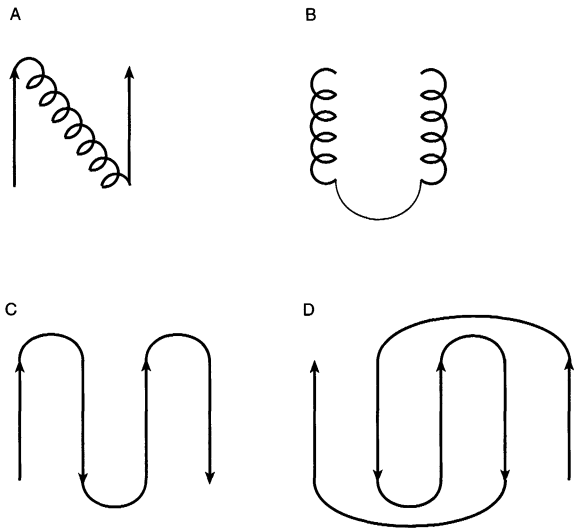


Figure 3.14. Schematic drawings of supersecondary structures in proteins. An α-helix is indicated by a coil, and a β-sheet strand by an arrow. (A) βαβ unit; (B) αα unit; (C) β meander; (D) Greek key motif. Other supersecondary structures, called beta barrels, contain rolled-up sections of β-sheets.

polypeptide chain folds in an irregular manner, the folding is *unique* for any given protein, fully reproducible from molecule to molecule, and critical for the protein's function. Any slight alteration in the tertiary (and/or the secondary) structure may cause the molecule to lose its biological activity.

3.4.1. X-Ray Diffraction

We can investigate the tertiary structure of proteins by X-ray diffraction, a technique that provides information about the positions of atoms and the distances between them. The method was first successfully applied to proteins by John Kendrew, studying myoglobin (1957), and by Max Perutz, working on hemoglobin (1959).

In *X-ray diffraction* we expose protein crystals to beams of X rays, a high-energy radiation of short wavelengths. The protein crystal is mounted in a capillary and positioned with a precise orientation relative to the impinging X-ray beam. As the X rays pass through the crystal lattice, they strike the atoms in that lattice and are scattered by them. The intensity of an atom's scattering depends on its electron cloud. Heavier atoms contain more electrons in their electron clouds than lighter atoms and produce more scattering. The scattered X rays are allowed to strike a photographic film, giving rise to a diffraction pattern (Figure 3.15A) that consists of a multitude

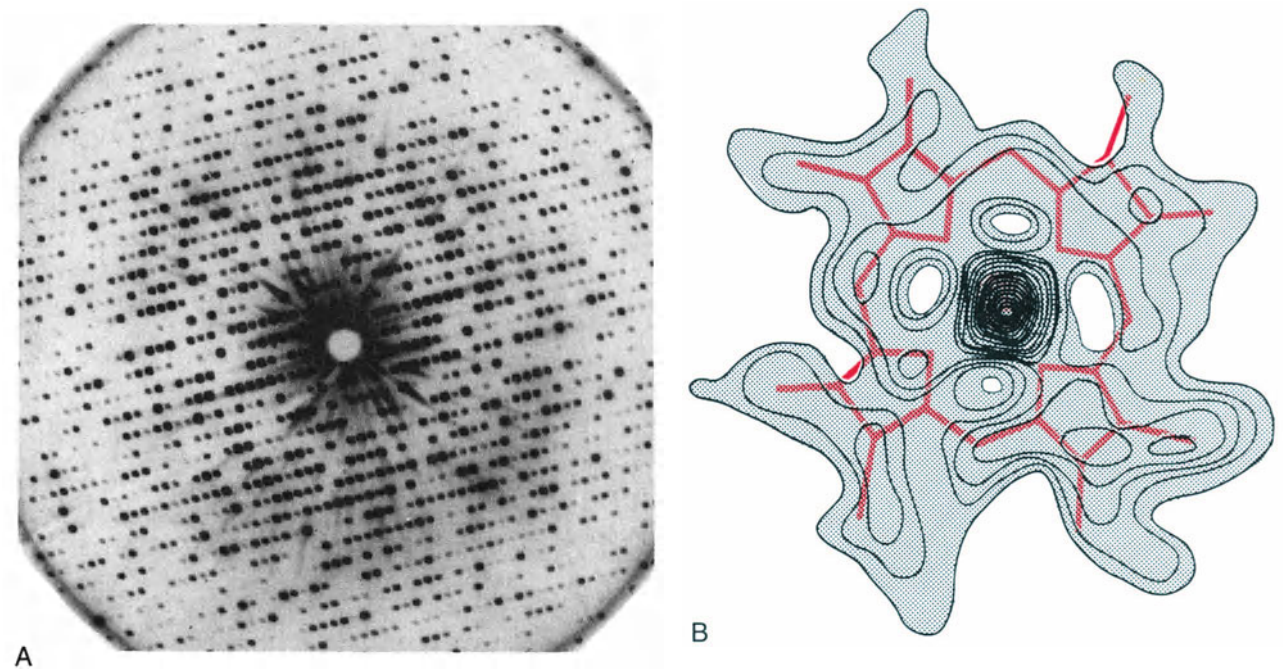


Figure 3.15. X-ray diffraction analysis. (A) Photograph of the diffraction pattern from a myoglobin crystal. [Reprinted, with permission, from J. C. Kendrew, *Science* 139:1259–1266 (1963). Copyright © 1963 American Association for the Advancement of Science.] (B) Section of the 2.0-Å resolution electron density map of myoglobin, showing location of the heme group (red) and the central iron atom (red circle). (Figure copyrighted © by Irving Geis. Reproduced with permission.)

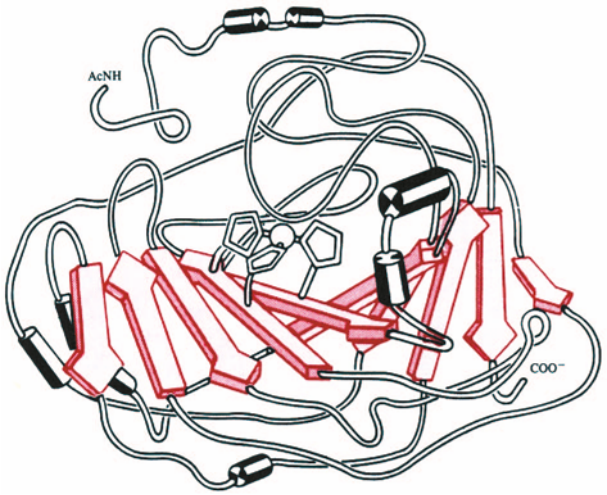


Figure 3.16. Tertiary structure of a globular protein. In this schematic representation of carbonic anhydrase, α-helices are indicated by cylinders, and strands of β-sheets are shown as arrows, pointing to the C-terminus of the polypeptide chain. The enzyme contains a Zn^{2+} ion (red) coordinated to three histidine side chains. [Reprinted, with permission, from K. K. Kannan *et al., Cold Spring Harbor Symp. Quant. Biol.* 36:221–231 (1971).]

of spots varying in both position and intensity. A change in the angle at which the X-ray beam strikes the crystal produces a different diffraction pattern.

Perutz discovered a way to extract information from X-ray diffraction patterns. His approach requires introducing a heavy metal atom into the protein without changing either the protein's conformation or the size and symmetry of the crystallographic unit cell. We term this an *isomorphous replacement.* The heavy metal ion produces intense scattering of the X-ray beam and yields an exceptionally dense spot in the diffraction pattern. With this spot as a reference point, and by means of thousands of calculations performed with a computer, one can analyze a diffraction pattern and convert it to an *electron density map* (Figure 3.15B), analogous to a geological contour map. Mapping electron densities in this fashion is tantamount to mapping the positions of individual atoms, since a region of high electron density pinpoints an atom's location.

3.4.2. Globular and Fibrous Proteins

The tertiary structure of globular proteins consists of varying portions of α-helices, β-sheets, and irregularly folded polypeptide chain segments (Figure 3.16). We refer to the percentage of amino acid residues that occur in the form of α-helical segments as the *helical content* of the molecule. As you can see from Table 3.5, the helical content of proteins varies greatly and does not necessarily correlate with the amount of β-pleated sheet and the number of reverse turns.

Fibrous proteins can form from polypeptide chains in either the α-helical or the β-sheet configuration. Some fibrous proteins, like α-keratin (see Section 3.8), consist entirely of bundles of α-helices. Other proteins, like β-keratin and silk fibroin, consist entirely of β-pleated sheets. Additionally, protein fibers, while not strictly fibrous proteins, can form by linear aggregation of globular proteins.

3.4.3. Forces Stabilizing the Tertiary Structure

Both covalent and noncovalent bonds are responsible for maintaining the tertiary structure of proteins (Figure 3.17). Covalent bonds consist of *intra-* and *interchain* disulfide bonds. Noncovalent bonds include van der Waals interactions, hydrogen bonds, ionic interactions, and hydrophobic interactions.

Table 3.5. Structural Features of Some Globular Proteins[a]

	α-Helix		β-Pleated Sheet			
Protein	Percent	Number of helices	Percent	Number of sheets	Number of strands	Number of Reverse turns
Hemoglobin	79	8	0	0	0	6
Insulin	52	3	6	1	1	0
Lysozyme	40	6	12	2	6	6
Cytochrome *c*	39	5	0	0	0	6
Ribonuclease	26	3	35	1	6	2
Chymotrypsin	14	3	45	2	12	17
Concanavalin A	2	1	57	3	18	0

[a]Adapted with permission from A. Liljas and M. G. Rossman, *Annu. Rev. Biochem.* 43:475–507 (1980), @ 1980 by Annual Reviews, Inc.

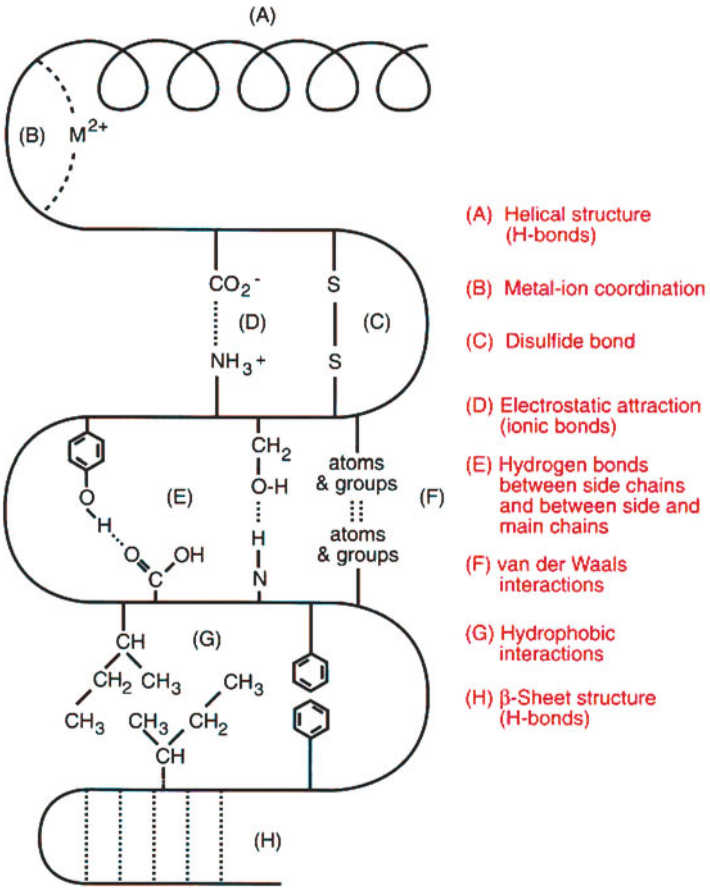


Figure 3.17. Forces that stabilize the tertiary structure of proteins.

Van der Waals interactions between uncharged molecular groups can make a significant contribution to the stability of proteins. Globular proteins generally have a tightly packed interior, as indicated by X-ray diffraction. Accordingly, interatomic contacts abound, and a variety of interactions can take place between adjacent atoms and molecular groups. H-bonds can form between peptide bonds, both in a regular manner as in an α-helix or a β-sheet, and in an irregular manner between any two peptide bonds located suitably close to each other in the folded polypeptide chain. Additionally, H-bonds can form between proton donor and acceptor groups in amino acid side chains. Ionic interactions occur between oppositely charged functional groups of the amino acids, and hydrophobic interactions involve nonpolar amino acid side chains.

Although a protein does not constitute a micelle, a globular protein is constructed somewhat along similar principles. X-ray diffraction has shown that, in globular proteins, polar amino acids tend to be spread over the surface of the molecule while nonpolar amino acids tend to be located inside the protein. This resembles the structure of a soap micelle (Figure 1.10), where polar carboxyl groups lie on the outside while nonpolar hydrocarbon chains cluster inside. Likewise, hydrophobic interactions between nonpolar amino acid side chains stabilize globular proteins much as interactions between hydrocarbon chains stabilize soap micelles. Lastly, polypeptide chains fold spontaneously to produce three-dimensional structures and soap molecules aggregate spontaneously to form micelles.

As an interesting footnote, we might mention that the fundamental structural models of proteins, the α-helix and the β-pleated sheet, were proposed solely on the basis of H-bond considerations. We now realize that, in addition to H-bonds, hydrophobic interactions are quite important in maintaining the structural integrity of proteins. Indeed, at times they may be the major force responsible for stabilizing the three-dimensional structure of the molecule.

3.4.4. Domains

As X-ray diffraction data and other information about protein structures accumulate, we see that a protein frequently has recognizable subsections of its overall three-dimensional structure, called **domains.**

A domain is an independently folded, relatively globular region of a polypeptide chain. A domain occurs as a spatially isolated structure that can be physically separated from other parts of the molecule by a suitable cut, or cuts, in the polypeptide chain. Proteins containing over 200 amino acid residues usually fold into two or more domains. Domains may interact slightly or extensively with each other; they may be associated with specific functions; and they vary in size. Most domains contain about 100–200 amino acid residues. Typically, one or, at times, two polypeptide segments connect one domain to another. Figure 3.18 shows an enzyme having two distinct domains. The variable and constant regions of immunoglobulins also constitute separate domains (see Figure 3.31).

3.5. QUATERNARY STRUCTURE

We can fully describe a monomeric protein by its primary, secondary, and tertiary structures. However, oligomeric proteins composed of subunits require still another structural level for their description. We call that level the **quaternary structure.** It consists of the *arrangement and interaction of subunits* without regard to the internal structures of the subunits. The interaction between the four subunits of hemoglobin (two α-chains and two β-chains) that results in formation of the functional molecule illustrates quaternary structure.

Forces responsible for maintaining the quaternary structure are generally noncovalent. Because subunits are typically globular proteins, nonpolar amino acids tend to cluster inside each subunit while polar amino acids tend to spread on its surface. These structural characteristics minimize hydrophobic interactions among subunits and maximize van der Waals interactions, H-bonds, and ionic interactions. Exceptions do, however, occur. The quaternary structure of hemoglobin, for example, results from numerous H-bonds as well as from extensive hydrophobic interactions among the subunits.

3.6. PROPERTIES OF PROTEINS

3.6.1. Net Charge

Acid–base properties of proteins reflect their molecular structure. Many of the functional groups in the amino acid side chains undergo ionization. Recall that you can use pK'_a values of free amino acids only as a rough guide to pK'_a values of the same groups in a protein. Actual pK'_a values depend on the location of each group within the polypeptide chain and on the electronic environment surrounding it. Every protein, like every amino acid and peptide, has its specific isoelectric point (pI), the pH at which the molecule has a net zero charge. At pH values more basic than the pI, the protein has a net negative charge; at pH values more acidic than the pI, the protein has a net positive charge.

Researchers make use of this fact in procedures requiring the precipitation of proteins by means of heavy metal ions. For example, *casein,* the major protein in milk, and *egg albumin,* the major protein in egg white, have pI values of 4.7 and 4.6, respectively. To precipitate these two proteins with heavy metal ions like Hg^{2+}, Pb^{2+}, or Ag^{+}, it is best to carry out the reaction at a pH above 5.0. Above pH 5.0, either protein has a net negative charge and interacts well with the metal cations.

A clinical application of this principle involves the treatment of poisoning caused by the ingestion of such heavy metal ions. In these cases, the patient receives a dose of egg white or milk—effectively a dose of egg albumin or casein—or both. Although some digestion of casein and egg albumin occurs in the stomach, a large fraction of each protein passes from the stomach into the intestine in undigested form. Intestinal fluid has a pH of about 7. At that pH, both proteins are negatively charged and combine readily with any heavy metal cations that have gotten into the intestine. The protein/metal-cation complex precipitates out and can be excreted, thereby *detoxifying* the intestine. This treatment does not remove metal cations left in the stomach, in which the pH is about 2; to eliminate them, the stomach must be pumped out.

3.6.2. Protein Solubility

We saw earlier that in order to become soluble, solute particles must interact more with solvent molecules, generally water, than with each other. The pH dependence of protein solubility reflects the same principle (Figure 3.19).

At pH values below or above the pI, the protein has a net positive or a net negative charge, respectively, and the charged protein particles repel each other. Extensive hydration by water dipoles further ensures that protein molecules stay apart. At those pH values, the protein is soluble. At the isoelectric point, however, the protein has a net zero charge, and protein molecules tend to react more with each other than with water. At this pH, solute–solute interactions exceed solute–solvent interactions, and protein molecules aggregate and precipitate out.

Protein solubility also varies with the ionic strength

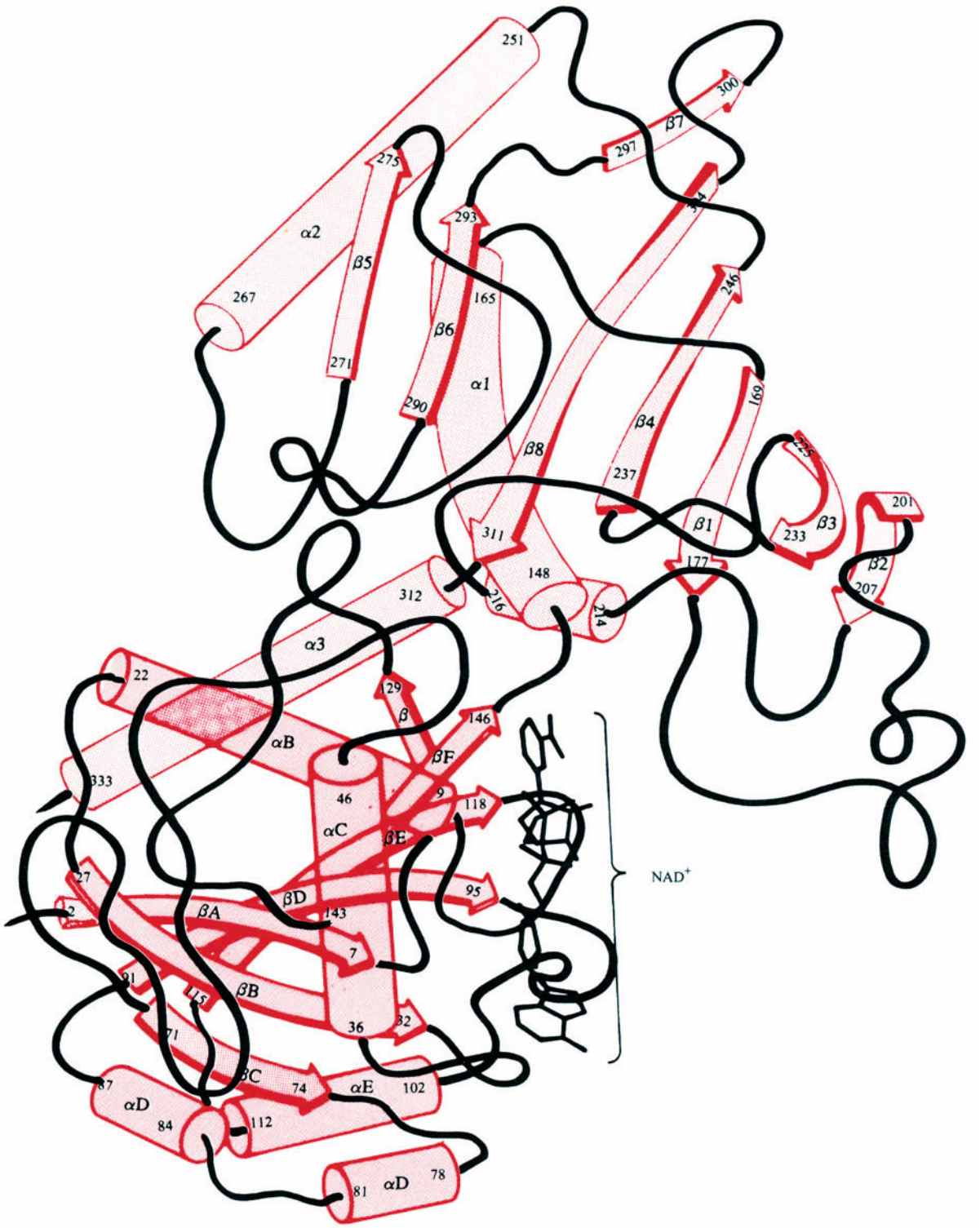


Figure 3.18. Subunit structure of glyceraldehyde 3-phosphate dehydrogenase from *Bacillus stearothermophilus*. Two distinct domains can be seen. The first (residues 1–146) binds the coenzyme NAD^+, and the second (residues 147–333) binds the substrate, glyceraldehyde 3-phosphate (not shown). [Reproduced, with permission, from G. Biesecker, *Nature (London)* 266:328–333 (1977). Copyright © 1977 Macmillan Magazines Limited.]

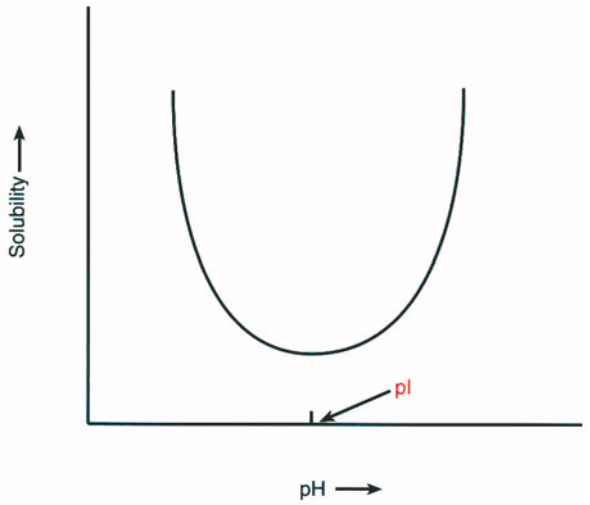


Figure 3.19. Variation of protein solubility with pH.

(see Appendix A) of the solution. As the ionic strength increases, protein solubility at first increases. We term this "salting in." However, beyond a certain point, the solubility begins to decrease, a phenomenon called "salting out" (Figure 3.20). At low ionic strengths, added salt ions surround and shield the protein's ionizable groups. This prevents interactions among ionizable groups so that protein–protein interactions decrease and protein solubility increases. At high ionic strengths, so much water has become bound by the added salt ions that not enough remains to properly hydrate the protein. This increases interactions among ionizable groups so that protein–protein interactions increase and protein solubility decreases.

Because of differences in structure and amino acid sequence, proteins differ in their salting-in and salting-out behavior. Moreover, solubility curves vary depending on the type and valence of the salt ions added. These properties form the basis for *fractional precipitation* by adding salt, a useful technique in isolating and purifying proteins. In this technique, one slowly adds a salt, typically ammonium sulfate, to a protein solution. As the salt concentration increases, different proteins precipitate out and can be removed or collected.

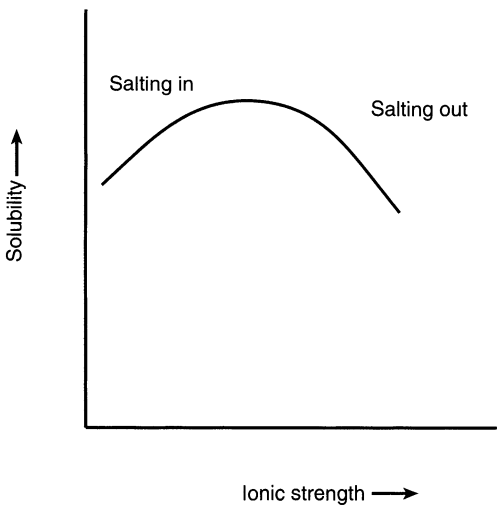


Figure 3.20. Variation of protein solubility with ionic strength.

3.6.3. Denaturation

It is apparent from our discussion that the complete, three-dimensional structure of a protein represents a delicate balance of covalent and noncovalent interactions. We must take care when isolating and studying proteins that this balance is not disturbed and that we do not alter the structure of the molecule in any way. Even a slight structural alteration may result in a significant change in the protein's properties and a decrease or loss of its biological activity. We refer to such alterations in, or loss of, the *native* structure of a protein as **denaturation.** Hans Neurath defined the term as "any non-proteolytic modification of the unique structure of a protein giving rise to definite changes in chemical, physical, or biological properties."

Denaturation includes cleavage of disulfide bonds, disruption of noncovalent interactions, and any other change in the protein except the breaking of peptide bonds (termed *degradation*). Denaturation may be reversible or irreversible. A reversal of denaturation is called *renaturation.* Denaturation applies not only to proteins. Breaking noncovalent interactions in nucleic acids also constitutes denaturation.

Protein denaturation consists of a partial or complete unfolding of the polypeptide chain, an opening up of the molecule's native structure (Figure 3.21). Regularly folded segments of α-helix and β-sheet lose their specific configurations, and the polypeptide chain forms a **random coil**—a flexible chain folded in an unordered manner. A dish of spaghetti resembles a collection of denatured polypeptide chains. These structural alterations lead to observable changes in the molecule's properties. To take a well-known example, consider boiling an egg. The heat leads to alterations in the configuration of egg albumin. The native polypeptide chains undergo denaturation and subsequently aggregate to form the coagulated egg white. Thus, denaturation amounts to the unseen cause that produces the observable effects.

Denatured proteins are generally more susceptible to enzymatic attack. Random-coil structures provide easier access to enzymes catalyzing peptide bond hydrolysis. Denaturation during cooking produces proteins that are more

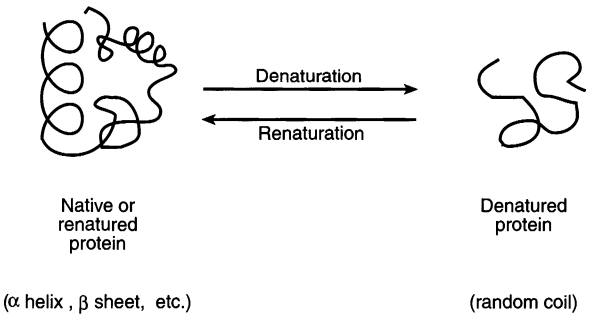
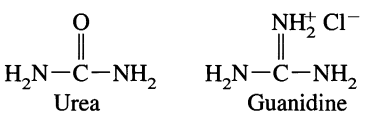


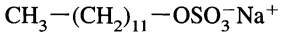
Figure 3.21. Interconversion of native and denatured proteins.

readily digested by proteolytic enzymes. Denaturation usually leads to exposure of some nonpolar groups, previously buried inside the protein, so that protein solubility decreases. Optical activity also decreases during denaturation as helical segments unfold to form random coils. Denatured proteins generally have less or no biological activity. Unraveling a polypeptide chain may alter the configuration of an active site and prevent an enzyme from binding a substrate, an antibody from binding an antigen, and so on.

Many factors can bring about denaturation. Any agent that disrupts either a covalent or a noncovalent bond critical for the structure of the molecule may cause partial or complete unfolding of the protein. Temperature is a common agent of denaturation. An increase in temperature disrupts hydrogen bonds and, at higher temperatures, hydrophobic bonds as well. Various compounds that reduce or oxidize disulfide bonds have already been described (Figures 2.12, 2.13, and 3.4).

Increasing the ionic strength or changing the pH may break ionic interactions among amino acid side chains. Added ions interact with oppositely charged groups on the protein and thereby disrupt intraprotein ionic interactions. Changes in pH alter the charges of specific functional groups and thus affect electrostatic interactions involving these groups. Compounds capable of forming hydrogen bonds or participating in hydrophobic interactions may cause denaturation by competing with corresponding groups of the protein. For example, *urea* and *guanidine hydrochloride* break intraprotein H-bonds, and *soaps* and *detergents* disrupt intraprotein hydrophobic interactions.





Sodium dodecyl sulfate (SDS), a detergent

Lastly, physical agents such as ultraviolet light, pressure, and sonication can lead to protein denaturation.

3.7. PROTEIN BINDING—HEMOGLOBIN AND MYOGLOBIN

The capacity to bind atoms, ions, and small or large molecules constitutes an essential aspect of protein function. The bound entity, called **ligand,** is usually, but not always, linked to the protein noncovalently. The **binding site** on the protein might be a simple functional group, such as a carboxyl group binding a proton. Alternatively, the binding site may be a section of the tertiary structure—a region on the surface of the protein defined by one or more segments of the polypeptide chain. The site can consist of one or several amino acid residues that may be close together or far apart along the chain, depending on its folding. There may be one or several binding sites per molecule. Multiple binding sites may bind the same ligand or different ligands. You can see that binding is a varied phenomenon, and we shall encounter many examples of it in our study of biochemistry. A powerful application of specific binding interactions, *affinity chromatography,* is discussed in Appendix C. In this section, we will illustrate some important aspects of protein binding by examining the binding of oxygen to hemoglobin and myoglobin.

3.7.1. Oxygen Saturation Curves

Hemoglobin is present in high concentration in red blood cells; it binds oxygen in the lungs and releases it to the tissues and cells around the blood capillaries. Hemoglobin functions primarily as an oxygen transport protein. Myoglobin is located in muscle cells; it binds and releases oxygen in the cytoplasm of these cells in response to changes in the oxygen concentration. Myoglobin functions primarily as an oxygen storage protein.

A plot of the extent of oxygen binding by these two proteins as a function of increasing oxygen concentration yields two curves, called *oxygen saturation curves,* that differ in their shape (Figure 3.22). You can see from this figure that myoglobin has great affinity for oxygen. Only a small increase in oxygen concentration (partial pressure) is needed to fully saturate myoglobin with oxygen. The curve rises steeply; it is a *hyperbolic curve.* Hemoglobin, on the other hand, initially binds oxygen with difficulty since the curve rises slowly at low oxygen concentrations. As the oxygen concentration increases, the binding affinity also increases, as detected by an increase in the slope of the curve. The curve levels off when hemoglobin be-

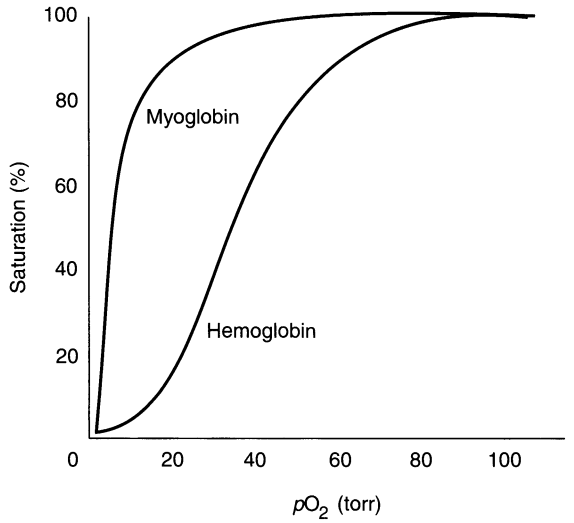


Figure 3.22. Oxygen saturation curves of myoglobin and hemoglobin. Oxygen concentration is expressed as partial pressure (pO_2) in torr (1 torr = 1 mm Hg; 1 atm = 760 torr).

comes fully saturated with oxygen. We call this type of curve a *sigmoidal curve* or an *S-shaped* curve.

The different properties of hemoglobin and myoglobin allow an organism to shift oxygen between these two proteins according to need. For example, starting with the deoxygenated proteins and increasing the partial pressure of oxygen (as in moving a person from high altitude to sea level), the myoglobin stores are replenished before the circulating hemoglobin is fully oxygenated. Conversely, starting with the fully oxygenated proteins and decreasing the partial pressure of oxygen (as in moving a person from sea level to high altitude), most of the hemoglobin is deoxygenated before the oxygen stores of myoglobin are depleted as a last resort. Similar changes in oxygen concentration occur in animals as the blood flows from the tissues (lower pO_2) to the lungs (higher pO_2) or vice versa.

The shift of oxygen between hemoglobin and myoglobin is analogous to shifting money between a savings account (myoglobin) and ready cash (hemoglobin). As money becomes available, prudent individuals build up savings accounts before spending excess funds as cash. Conversely, as spending becomes necessary, most individuals use up cash reserves before dipping into savings accounts.

3.7.2. Cooperativity—Multiple Binding Sites

Apart from the usefulness of having two proteins with unique oxygen binding properties, we must ask what accounts for this different behavior. The answer is that the properties of hemoglobin and myoglobin are a function of their respective structures. Myoglobin occurs as a *monomeric protein* in which the polypeptide chain, or *globin,* surrounds a *heme.* Heme consists of an iron atom coordinated in a planar tetrapyrrole ring system, called *protoporphyrin IX* (Figure 3.23). Hemoglobin, on the other hand, occurs as a *tetrameric protein* in which each subunit consists of a polypeptide chain (globin) surrounding a heme. Hemoglobin contains two types of polypeptide chains (two α-chains and two β-chains), and both resemble the single polypeptide chain of myoglobin. Additionally, the heme structure is identical in both myoglobin and hemoglobin.

Thus, there exists only one major difference between the two proteins: hemoglobin is a tetramer, with four binding sites for oxygen, whereas myoglobin is a monomer, with only one binding site for oxygen. Each binding site in the two proteins binds a *molecule* of oxygen, O_2. Myoglobin has great affinity for oxygen, and binding proceeds rapidly. Moreover, binding of a single O_2 fills the protein's binding site so that no additional oxygen binding can occur. Such binding results in a hyperbolic oxygen saturation curve.

Hemoglobin has a lower affinity for oxygen, and binding proceeds slowly at first. However, once some oxygen has become bound, hemoglobin's affinity for additional oxygen increases. The change in affinity is only possible because hemoglobin has more than one binding site per molecule. The initial binding of oxygen to *some* sites enhances oxygen binding to *subsequent* sites. We refer to this phenomenon as **cooperative interactions,** or

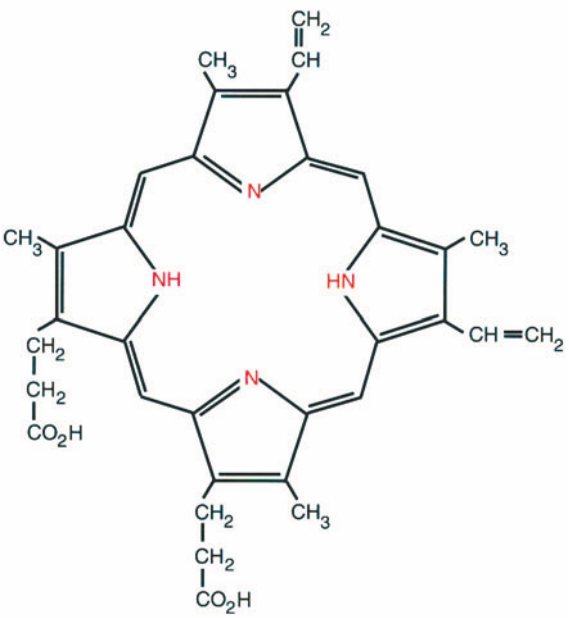


Figure 3.23. Protoporphyrin IX.

cooperativity. This type of binding results in a sigmoidal oxygen saturation curve.

> Cooperativity exists whenever the binding of one ligand affects the binding of subsequent ligands (identical or different ones) to other sites (identical or different ones).

Multiple binding sites may be located on a single macromolecule or distributed over a macromolecular aggregate. Cooperative interactions characterize oligomeric proteins that consist of two or more subunits, with each subunit having one or more binding sites. Cooperative interactions in such proteins are revealed by a sigmoidal binding curve.

Cooperative interactions may be *positive* or *negative.* With positive cooperativity, the binding of a ligand enhances the binding of a subsequent ligand. With negative cooperativity, the binding of a ligand diminishes the binding of a subsequent ligand. Note that if ligand binding is cooperative, then ligand dissociation must likewise be cooperative. Thus, removal of the first O_2 from *oxyhemoglobin* proceeds slowly, but subsequent molecules of O_2 dissociate more readily.

3.7.3. Changes in Hemoglobin upon Oxygenation

We understand the nature of the cooperative interactions in hemoglobin quite well (Figure 3.24). Upon oxygena-

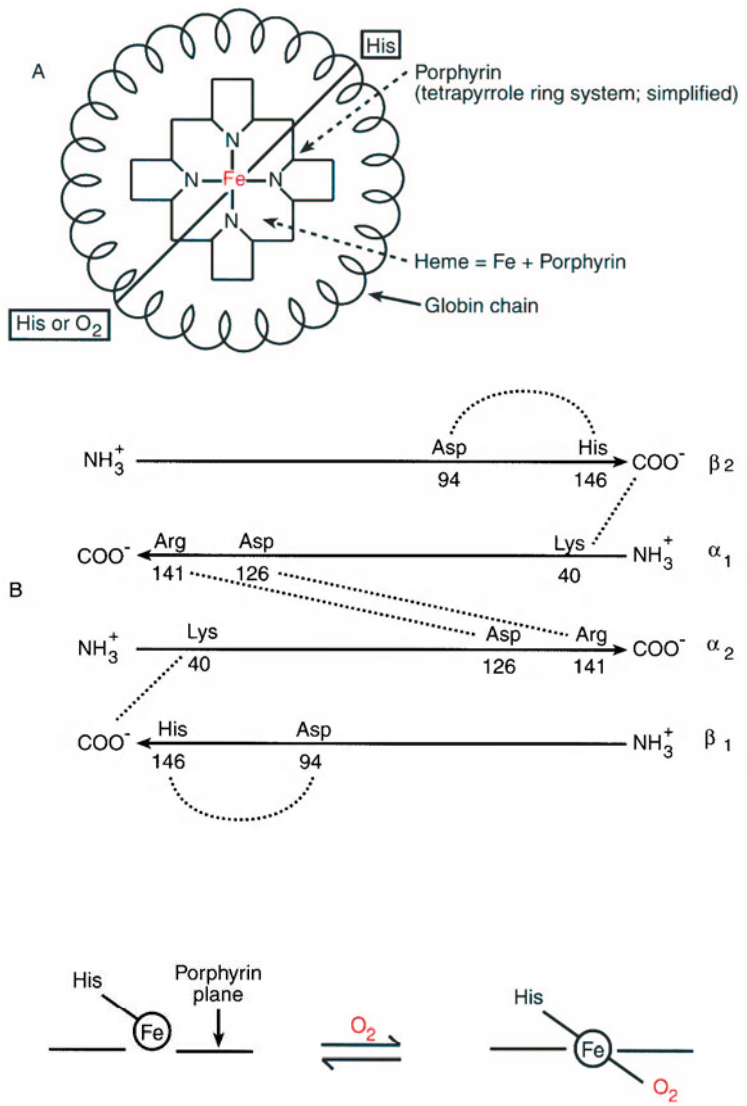


Figure 3.24. Cooperative interactions in hemoglobin. (A) Schematic drawing of a subunit. (B) Electrostatic interactions within and between subunits. (C) Movement of iron into the porphyrin plane upon oxygenation.

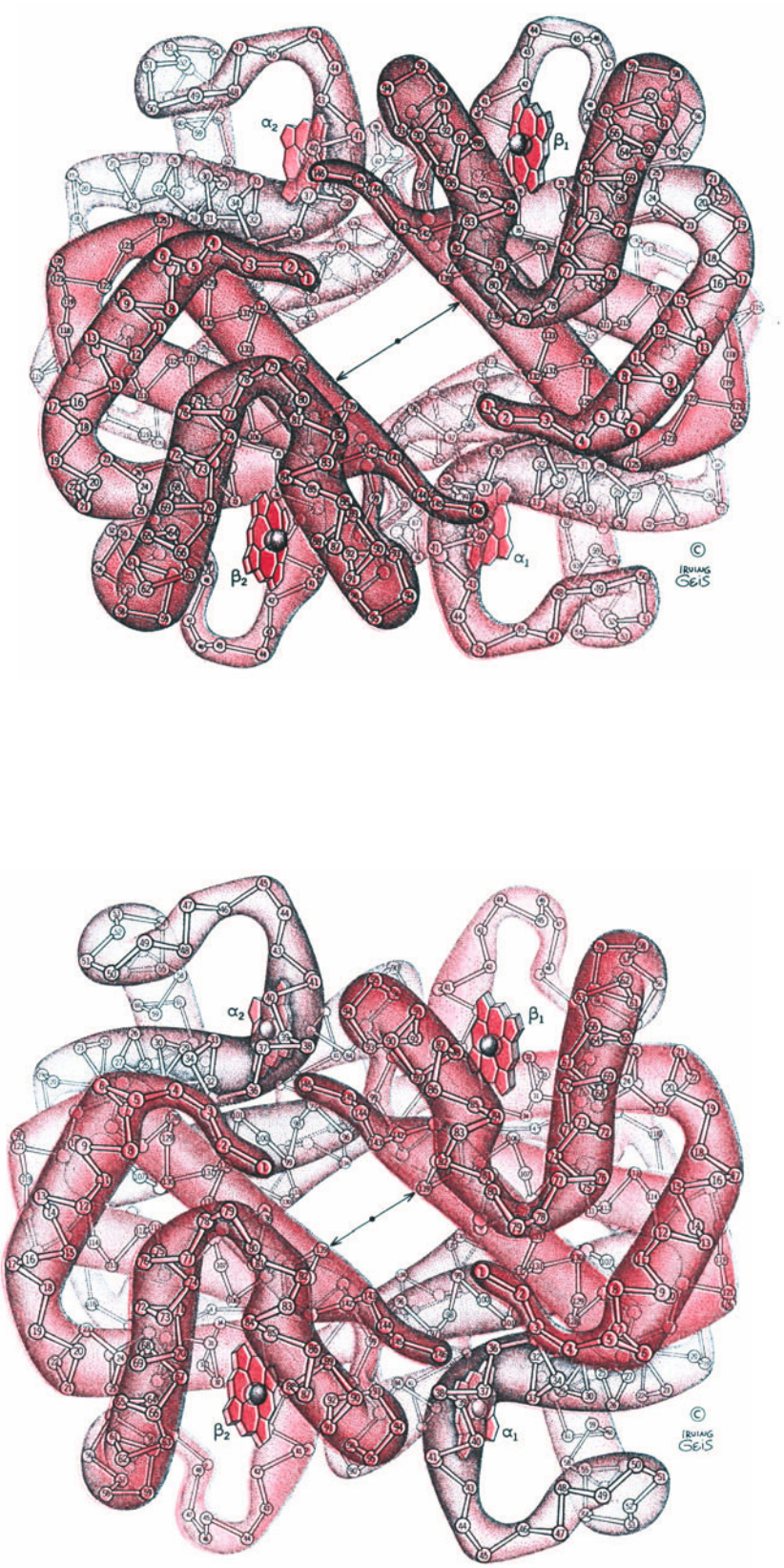


Figure 3.25. The structure of oxyhemoglobin and deoxyhemoglobin. Subunits are closer together in oxyhemoglobin than in deoxyhemoglobin. (Figure copyrighted © by Irving Geis. Reproduced with permission.)

tion, the effective radius of the iron atom decreases; introducing the electronegative oxygen molecule compresses the electron cloud surrounding the iron. As a result, the iron atom, which was originally not centered in the plane of the protoporphyrin, can now move into that plane. Displacement of the iron amounts to 0.6 Å.

Iron has six coordination positions, four of which are taken up by linkage to the nitrogens of the protoporphyrin ring. A fifth position links the iron to a histidine residue in the polypeptide chain of the surrounding globin. The sixth position is the one at which an oxygen molecule (O_2) can bind. The fifth and sixth coordination positions are on opposite sides of the protoporphyrin plane. Since the iron is coordinated to the globin polypeptide chain, movement of the iron into the protoporphyrin plane distorts the globin chain.

This conformational change breaks several interchain ionic bonds. The structural change in one subunit, resulting from oxygenation, becomes translated into structural changes at the interfaces of other subunits. Properties of these subunits, including their affinity for oxygen, are thereby altered. In this way, binding of oxygen to one site affects the binding affinity of other sites. Such effects constitute the essence of cooperative interactions.

The conformational changes brought about in the subunits by oxygenation result in *oxyhemoglobin* being more compact than *deoxyhemoglobin.* In oxyhemoglobin, the subunits are closer together and the *central cavity* is smaller than in deoxyhemoglobin (Figure 3.25). The conformational changes also cause alterations in acid–base properties. For example, the distance between an aspartic acid residue (#94) and a histidine residue (#146) on the same β-chain increases. Upon oxygenation, the proton of the imidazole group of histidine is less attracted to the negatively charged β-carboxyl group of aspartic acid. In other words, the imidazole proton dissociates more readily; its pK_a' has been lowered. The histidine has become a stronger acid.

Several other such changes alter the acid–base properties in a similar manner. As a result, oxyhemoglobin is a *stronger acid* than deoxyhemoglobin, which may be indicated schematically by writing:

$$HHb + O_2 \rightleftarrows HBO_2^- + H^+ \qquad (3.1)$$

3.7.4. The Bohr Effect

Equation (3.1) describes one of the central aspects of the chemistry of respiration. It also explains a phenomenon known as the **Bohr effect.** The Bohr effect refers to the *decrease in the oxygen affinity of hemoglobin produced by either a decrease in the pH or an increase in the par-*

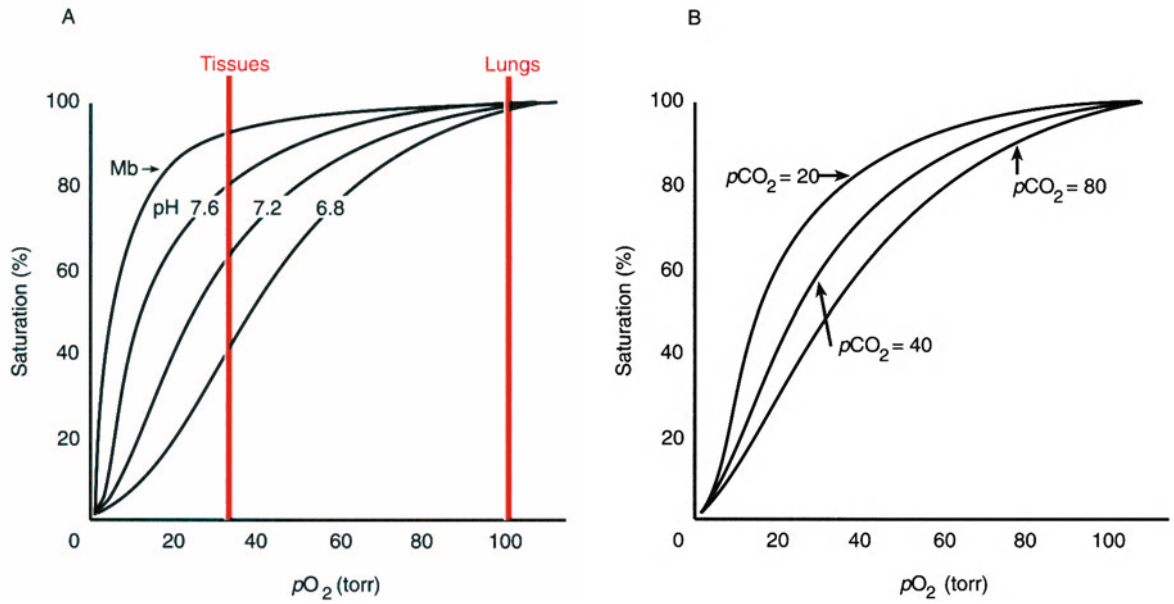


Figure 3.26. The Bohr effect. Variation of oxygen binding by hemoglobin with pH (A) and partial pressure of CO_2 (B). Binding of oxygen by myoglobin varies little with pH so that the curve is approximately correct at all three pH values.

tial pressure of carbon dioxide (Figure 3.26). Put somewhat differently, a decrease in pH or an increase in [CO_2] favors the dissociation of oxygen from oxyhemoglobin.

You can see from Eq.(3.1) that a decrease in pH, or increase in [H^+], drives the reaction from right to left according to Le Chatelier's principle, causing oxygen to dissociate from oxyhemoglobin. The effect of CO_2 results from two reactions in which CO_2 participates. In one reaction, CO_2 is hydrated to carbonic acid, which then ionizes to bicarbonate and a proton:

$$CO_2 + H_2O \rightleftarrows H_2CO_3 \rightleftarrows HCO_3^- + H^+$$

In the second reaction, CO_2 reacts with amino groups of hemoglobin and other proteins to form *carbamino compounds* that also ionize:

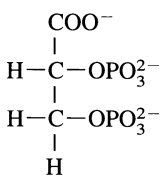
$$CO_2 + R\text{-}NH_2 \rightleftarrows R\text{-}NH\text{-}COOH \rightleftarrows R\text{-}NH\text{-}COO^- + H^+$$

An increase in [CO_2] drives both of these reactions from left to right, thereby increasing the H^+ concentration. Hence, an increase in [CO_2] has the same effect as an increase in [H^+] and favors oxygen dissociation from oxyhemoglobin.

Conditions of low pH and high [CO_2] are precisely those that exist at the level of actively metabolizing tissue. As we shall see, most metabolic intermediates are weak acids. Metabolizing tissue, therefore, tends to have a somewhat lower pH than inactive tissue. Likewise, the *citric acid cycle,* which produces most of the respiratory CO_2, is fully operative during aerobic metabolism so that the concentration of CO_2 is higher in metabolizing tissue than in inactive tissue. You can see that the conditions in actively metabolizing tissue produce a perfect match between the tissue's increased need for oxygen and the enhanced release of oxygen from oxyhemoglobin.

3.7.5. 2,3-Bisphosphoglycerate—a Regulator of Oxygen Binding

In addition to the Bohr effect, there exists another important regulatory factor in the chemistry of respiration—*2,3-bisphosphoglycerate (BPG):*



BPG binds electrostatically to the central cavity of intracellular hemoglobin. The binding involves specific interactions between the five negative charges present on BPG at intracellular pH and eight positively charged sites on hemoglobin (Figure 3.27). In deoxyhemoglobin, the central cavity is large enough so that BPG fits into it well. In oxyhemoglobin, the cavity is too small to accommodate BPG. Consequently, BPG binds to deoxyhemoglobin but not to oxyhemoglobin.

Upon binding to deoxyhemoglobin, BPG effectively "locks" the protein into this form, retards its oxygenation to oxyhemoglobin, and thereby lowers deoxyhemoglobin's affinity for oxygen. While binding of oxygen increases the proportion of oxyhemoglobin,

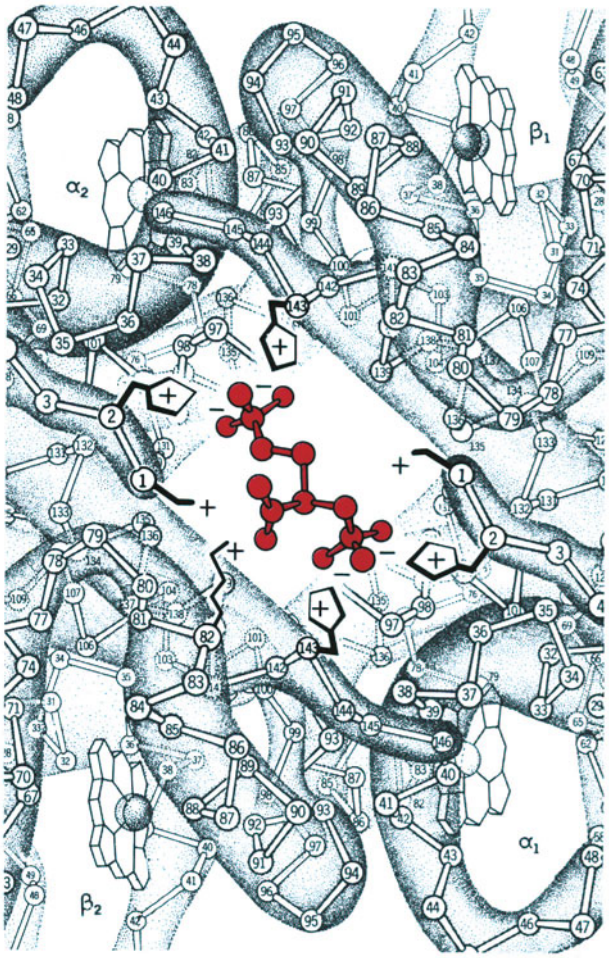


Figure 3.27. Binding of 2,3-bisphosphoglycerate (BPG) to the central cavity of deoxyhemoglobin. Binding involves electrostatic interactions between BPG's five negative charges and eight positive charges of two β-subunits of deoxyhemoglobin: α-NH_2 of Val 1; imidazole of His 2; ε-NH_2 of Lys 82; and imidazole of His 143. (Figure copyrighted © by Irving Geis. Reproduced with permission.)

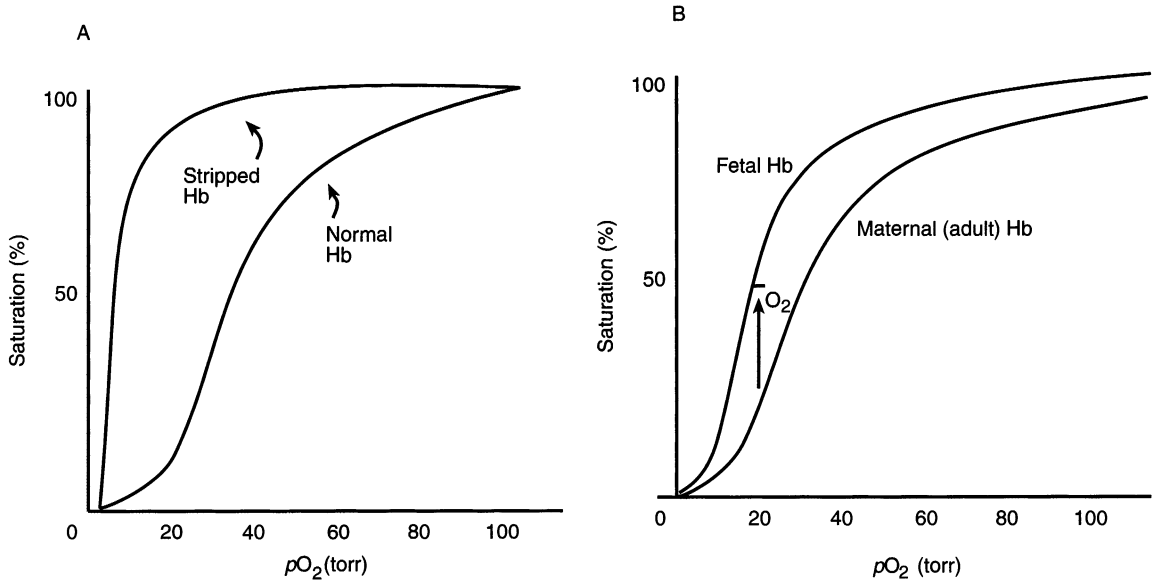


Figure 3.28. The effect of 2,3-bisphosphoglycerate (BPG) on oxygen saturation curves. (A) Hemoglobin, stripped of BPG, binds oxygen almost as well as myoglobin. (B) Fetal hemoglobin has greater affinity for oxygen because it binds BPG less strongly than adult hemoglobin.

binding of BPG increases the proportion of deoxyhemoglobin.

BPG concentration in red blood cells is high and nearly identical to the concentration of hemoglobin. As a result, normal hemoglobin has significant amounts of BPG bound to it. "Stripping" such hemoglobin of its BPG leads to an almost complete loss of the characteristic sigmoidal oxygen saturation curve, and the binding of oxygen approaches that shown by myoglobin (Figure 3.28A).

BPG has an important physiological role because it binds less strongly to fetal hemoglobin than to adult hemoglobin. Consequently, fetal hemoglobin has a greater affinity for oxygen than adult hemoglobin (Figure 3.28B). This difference allows oxygen released by maternal hemoglobin in the placenta to become bound to fetal hemoglobin and transported to fetal tissues.

The decreased binding of BPG to fetal hemoglobin results from the structure of its central cavity. Fetal hemoglobin, like adult hemoglobin, is a tetramer. It contains two α-chains, but instead of two β-chains it contains two γ-chains, in which a histidine has been replaced by serine. Because this residue forms part of the lining of the central cavity, fetal hemoglobin has only six positive charges in its central cavity, compared to eight for maternal hemoglobin. Because of the fewer charges, fetal hemoglobin has a lower binding affinity for BPG.

3.8. STRUCTURES OF SOME SPECIFIC PROTEINS

We can illustrate some of the principles discussed in this chapter by describing the structures of several proteins, including keratins, collagen, myosin, and immunoglobulins.

3.8.1. Keratins

Keratins are fibrous proteins. α-Keratins occur in hair and wool of mammals; their basic structural unit is a coiled coil of two α-helices. Two of these coiled coils form a **protofibril.** Protofibrils are arranged in a "9 + 2" pattern, or **microfibril,** in which nine protofibrils form a ring that surrounds a central core of two protofibrils. A large number of microfibrils make up a single hair or strand of wool.

The springiness of hair and wool fibers results from the coiled coil's tendency to untwist when stretched and to reform when the force is relaxed. Stretching the fibers breaks hydrogen bonds, and relaxing the force lets these bonds reform so that the fibers regain their original conformation.

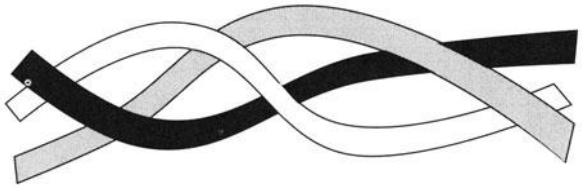


Figure 3.29. The triple helix of collagen—a right-handed supercoil, composed of three left-handed helices. Each individual helix has 3.3 residues per turn and a pitch of 1.0 nm. Crowding of the helices requires glycine at every third position of each polypeptide chain. Hydrogen bonding is not shown.

When human hair is given a "permanent," disulfide bonds linking the helices are first broken and then allowed to reform in different positions after the hair has been set.

β-Keratins have β-sheet structures and occur in feathers, skin, and scales of most birds and reptiles.

3.8.2. Collagen

Collagen, the major component of connective tissue in all multicellular animals, represents the most abundant protein of vertebrates. Collagen is a right-handed triple helix, composed of three left-handed helices, linked by H-bonds (Figure 3.29). Individual triple helices are cross-linked covalently, resulting in a structure that has great tensile strength. The polypeptide chains of collagen contain large amounts of proline, hydroxyproline, and glycine.

3.8.3. Myosin

Myosin, the structural unit of the *thick filaments* of striated muscle, constitutes about 60–70% of total muscle protein (Figure 3.30). Myosin consists of six polypeptide chains, two heavy ones (MW = 230,000 each) and two pairs of light ones (MW = 20,000 each). The functional complex has a molecular weight of 540,000, and its structure is part fibrous and part globular. The N-terminal half of each heavy chain folds into a globular head, while the C-terminal half forms a long α-helical tail. Two light chains are complexed to each heavy-chain head. A thick filament consists of several hundred myosin molecules with their rodlike tails packed end-to-end in a regular staggered array.

3.8.4. Immunoglobulins

Immunoglobulins, or **antibodies,** form a related but enormously diverse group of proteins that have a common structural unit (Figure 3.31) composed of two heavy chains (MW ≈ 50,000) and two light chains (MW ≈ 25,000). Immunoglobulins contain five different types of heavy chains and two types of light chains. The chains are linked via disulfide bonds and noncovalent bonds to form an asymmetric, Y-shaped, globular molecule. Each heavy chain has an oligosaccharide linked to it so that immunoglobulins constitute a group of *glycoproteins.* The

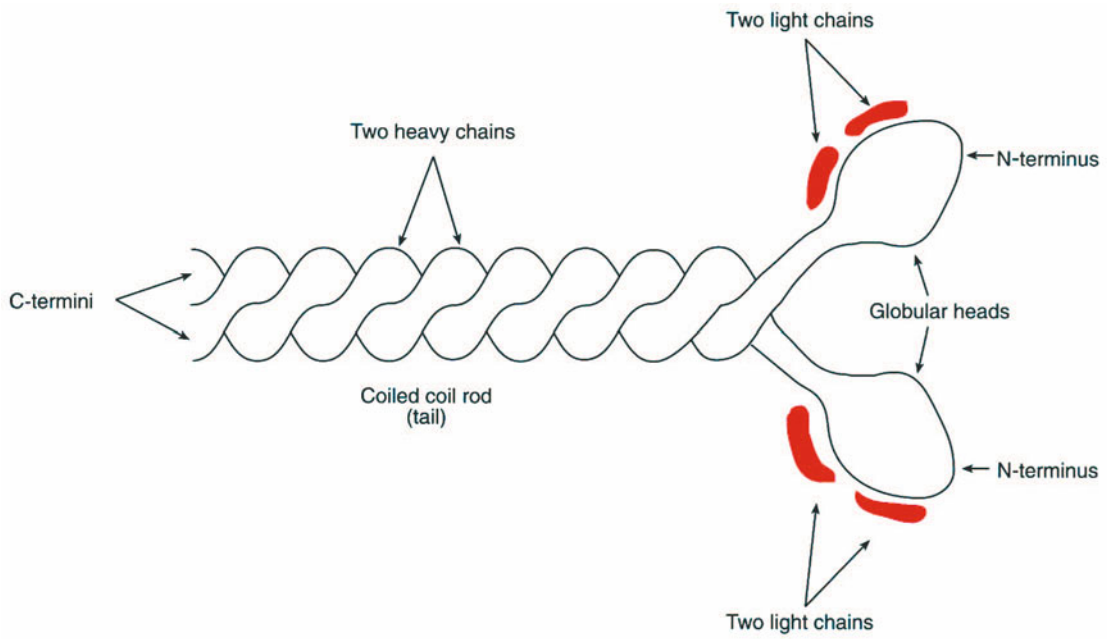


Figure 3.30. Schematic structure of myosin.

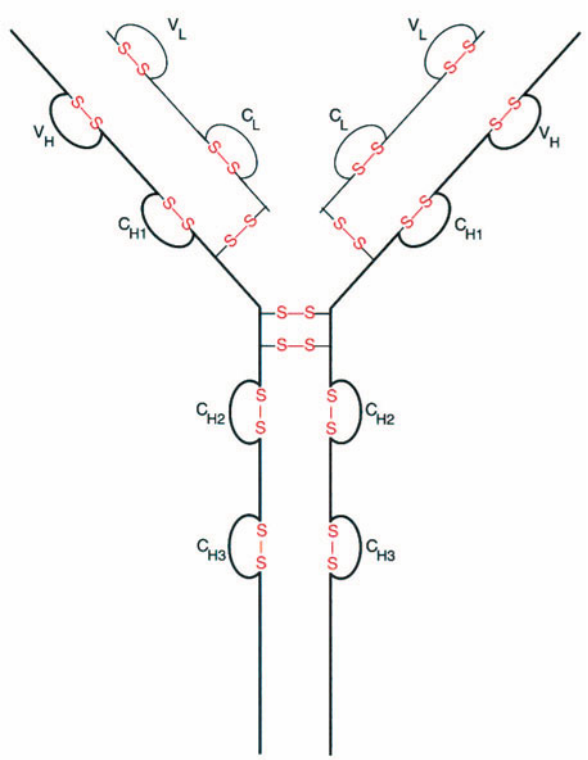


Figure 3.31. Immunoglobulins. Basic antibody structure. C and V designate constant and variable regions; L and H designate light and heavy chains. IgG is a monomer of the basic structure; IgA occurs as a monomer, dimer, or trimer, and IgM is a pentamer. Monomers are linked by a joining chain (J-chain) via disulfide bonds.

five classes of immunoglobulins are designated *IgG*, *IgA*, *IgM*, *IgD*, and *IgE*. They vary in their structures and function in different protective roles.

Both the light and the heavy chains have *variable regions*, in which many amino acid substitutions occur, and *constant regions*, in which the amino acid sequence is conserved. The variable region of a light chain, together with that of a heavy chain, forms an *antigen binding site*. For a given class of antibodies, the amino acid sequences are constant in the constant regions. Variations in amino acid sequences in the variable regions are responsible for the great specificity of antibodies within each class.

SUMMARY

Proteins have many specialized and several general functions. They are a source of energy, provide dietary nitrogen, serve as buffers, and contribute to osmotic pressure. We recognize four levels of protein structure: primary, secondary, tertiary, and quaternary. The primary structure refers to the type, number, and sequence of amino acids in the polypeptide chain. We determine it by partially hydrolyzing the protein, separating the peptide fragments, and then sequencing each fragment.

Regular folding of the polypeptide chain along one axis produces the secondary structure. According to the Pauling–Corey models, secondary structure results from hydrogen bonding between peptide bonds. H-bonds form between the NH group of one peptide bond and the CO group of another, and all peptide bonds participate in H-bond formation. Formation of intrachain hydrogen bonds produces an α-helix; formation of interchain hydrogen bonds produces a β-pleated sheet. We can estimate the helical content of proteins by measuring optical rotation.

Tertiary structure constitutes the three-dimensional structure of the molecule, the irregular but reproducible folding of an already folded polypep-

tide chain. The tertiary structure is stabilized by disulfide bonds, H-bonds, and van der Waals, hydrophobic, and ionic interactions. We can determine tertiary structure from X-ray diffraction data. The arrangement of subunits in an oligomeric protein and the interactions among different subunits constitute the quaternary structure. Subunits are generally linked noncovalently.

Proteins, like amino acids and peptides, have isoelectric points (pI), at which the molecules carry a net zero charge. At a pH above its pI, a protein has a net negative charge; at a pH below its pI, it has a net positive charge. Protein solubility varies as a function of pH and ionic strength.

Binding of oxygen to myoglobin and hemoglobin yields two different types of oxygen saturation curves: a hyperbolic one for myoglobin and a sigmoidal (S-shaped) one for hemoglobin. These curves reflect the different structures and properties of the two proteins. Because of its subunit structure, hemoglobin shows cooperative effects in oxygen binding. Cooperative interactions occur when the binding of some ligands to some sites affects the binding of subsequent ligands to other sites. Cooperativity can be positive (enhance binding) or negative (diminish binding). Binding of oxygen to hemoglobin leads to conformational changes and to changes in the acid–base properties of the molecule. Changes in pH, [CO_2], and concentration of 2,3-bisphosphoglycerate affect the binding.

SELECTED READINGS

Borman, S., Scientists refine understanding of protein folding and design, *Chem. Eng. News,* May 27, 1996.

Bradley, W. A., Gianturco, S. H., and Segrest, J. P. (eds.), Plasma lipoproteins, Part C, *Methods in Enzymology,* Vol. 263, Academic Press, San Diego (1996).

Creighton, T. E., *Proteins—Structures and Molecular Properties,* 2nd ed., W. H. Freeman, New York (1993).

Doolittle, R. F., The multiplicity of domains in proteins, *Annu. Rev. Biochem.* 64:287–314 (1995).

Dreuth, J., *Principles of Protein X-Ray Crystallography,* Springer-Verlag, New York (1994).

Eaton, W. A., and Hofrichter, J., Sickle cell hemoglobin polymerization, *Adv. Protein Chem.* 40:63–279 (1988).

Eisenberg, D. S., and Richards, F. M., Protein stability, *Adv. Protein Chem.* 46:1–311 (1995).

Huang, C. S., He, W., Meister, A., and Anderson, M. E., Amino acid sequence of rat kidney glutathione synthetase, *Proc. Natl. Acad. Sci. USA* 92:1232–1236 (1995).

Karger, B. L., and Hancock, W. S. (eds.), High resolution separation and analysis of biological macromolecules, Part B, *Methods in Enzymology,* Vol. 271, Academic Press, San Diego (1996).

Liu, Y., Van Heeswijck, R., Høj, P., and Hoogenraad, N., Purification and characterization of ornithine acetyltransferase from *Saccharomyces cerevisiae, Eur. J. Biochem.* 228:291–296 (1995).

Olson, J. S., and Phillips, G. N., Jr., Kinetic pathways and barriers for ligand binding to myoglobin, *J. Biol. Chem.* 271:17593–17596 (1996).

Ozols, J., Amino acid analysis, in *Methods in Enzymology,* Vol. 182, pp. 587–601, Academic Press, San Diego (1990).

Pain, R. H. (ed.), *Mechanisms of Protein Folding,* IRL Press, Oxford (1994).

Pauling, L, Corey, R. B., and Branson, H. R., The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chains, *Proc. Natl. Acad. Sci. USA* 37:205–211 (1951).

Perutz, M. F., *Mechanisms of Cooperativity and Allosteric Regulation in Proteins,* Cambridge University Press, Cambridge (1990).

REVIEW QUESTIONS

A. Define each of the following terms:

Random coil	Denaturation
Globular protein	Cooperative interactions
Prosthetic group	Reverse turn
Peptide map	Unit evolutionary period (UEP)
Bohr effect	Subunit

B. Differentiate between the two terms in each of the following pairs:

α-Helix / β-pleated sheet	Oligomer / monomer
Domain / supersecondary structure	Simple protein / conjugated protein
Intrachain disulfide bond / interchain disulfide bond	Sequence homology / amino acid composition
Ligand / binding site	Protofibril/microfibril

C. (1) What is meant by the terms primary, secondary, tertiary, and quaternary structure of proteins? What bonds or forces are involved at each level and how can these bonds be disrupted? How are solubility and optical rotation related to protein structure?
(2) What changes occur in deoxyhemoglobin as it undergoes oxygenation and how do these changes explain the positive cooperativity observed?
(3) Outline the steps typically used in the isolation, purification, and characterization of a protein.
(4) Describe the stages involved in determining the primary structure of a protein.
(5) Studies of the primary structure of proteins have led to a number of important generalizations or principles. What are they?
(6) What are the constraints on the formation of an α-helical segment and why do they have this effect?

PROBLEMS

3.1. Draw a graph, plotting solubility as a function of pH, for two hypothetical proteins, A and B. Protein A consists of 60 glutamic acid and 20 lysine residues; protein B consists of 20 glutamic acid and 60 lysine residues. On the basis of your graph, suggest a way of separating these two proteins.

3.2. The two least frequent amino acids in a protein are valine and threonine. The protein contains 0.400 mol % valine and 0.500 mol % threonine (mol % = no. of moles/100 moles). What is the minimum molecular weight of the protein, assuming that the molecular weight of an amino acid residue is 110?

3.3. Assuming that 2,3-bisphosphoglycerate binds *nonspecifically* to proteins, which of the following might be likely to exhibit such binding at pH 7.0?

Urease (pI = 5)
Myoglobin (pI = 7)
Lysozyme (pI = 11)

3.4. The following amino acids are internal residues in a polypeptide chain and are located such that they can interact as indicated. Which amino acid pairs in (a) could produce ionic interactions? Which pairs in (b) could form H-bonds? Which pairs in (c) could produce hydrophobic interactions?

	(a)	(b)	(c)
1.	Leu . . . His	Cys . . . Cys	Arg . . . Thr
2.	Asn . . . Gln	Ala . . . Pro	Ile . . . Ile
3.	Arg . . . Asp	Tyr . . . Glu	Val . . . Leu
4.	Cys . . . Met	Gly . . . Ala	Pro . . . Phe

3.5. Consider a typical globular protein that is titrated twice: (a) using the native protein and (b) using the denatured protein. Would you expect to find any differences with respect to the *number* of side-chain functional groups accessible to titration in (a) and (b)? If you were able to determine the pK'_a of a given functional group of a given amino acid residue, would you expect it to be the same in (a) and (b)?

3.6. An oligomeric protein consists of a number of *nonidentical* subunits; each subunit is a single polypeptide chain. DNP end-group analysis of 1.8 g of protein yields 0.050 millimoles of DNP-Gly and 0.040 millimoles of DNP-Trp. What is the minimum average molecular weight of the subunits?

3.7.* Ten grams of a protein (MW = 20,000) is dissolved in 100 ml of water at pH 7.0. To this solution is added 2.0 ml of 6.0*M* HCl, resulting in a new pH of 6.0. Calculate the number of protons that have become bound per molecule of protein upon addition of the HCl, on the assumption that all of the added protons have become bound to the protein.

3.8.* What is the sequence of a tetrapeptide (4 AA) for which the following have been ascertained?

(a) It contains an unhydroxylated, helix-breaking amino acid.
(b) It yields an optically inactive DNP-amino acid when treated with the Sanger reagent.
(c) It yields the most basic amino acid when treated for a very brief period with carboxypeptidase.
(d) It contains an imidazole group but no sulfur.
(e) It can be cleaved to two dipeptides, both of which, at pH 4, migrate toward the cathode in electrophoresis (see Table 2.1).

3.9. Ten milligrams of a protein is completely precipitated by adding 1.0 ml of $2.0 \times 10^{-4}M$ $HgCl_2$. What is the minimum molecular weight of the protein?

3.10. Incubation of 1-g samples of a protein with the enzyme aminopeptidase yields the following results:

Incubation time (s)	Amino acids released (micromoles)	
	Phe	Leu
10	5.0	—
20	8.0	2.0
40	8.0	4.0

What is the minimum molecular weight of the protein?

3.11.* Deduce the sequence of a hexapeptide (6 AA) that yields:

(a) DNP-Asp when treated with 1-fluoro-2,4-dinitrobenzene, followed by acid hydrolysis
(b) PTH-Met during the second cycle of an Edman degradation
(c) the amino alcohol of glycine when reduced with $LiBH_4$, followed by hydrolysis
(d) free glycine and histidine when treated for a brief period with carboxypeptidase
(e) two spots, corresponding to dipeptides, when subjected to fingerprinting *after* treatment with cyanogen bromide

3.12.* A heptapeptide (7 AA) yields two peptides when treated with performic acid. The smaller peptide yields:

(a) DNP-Cys when treated with the Sanger reagent
(b) cysteine and alanine when treated for a brief period with aminopeptidase
(c) two amino acid derivatives and free proline when treated with hydrazine

The larger peptide yields:

(d) PTH-Phe after one cycle of the Edman degradation
(e) the amino alcohol of aspartic acid when treated with $LiBH_4$, followed by hydrolysis
(f) aspartic acid and leucine when treated for a brief period with carboxypeptidase

Deduce the sequence of the original heptapeptide.

3.13. One millimole of a pure protein (MW = 60,000) reacts with exactly 6.0 millimoles of 2-mercaptoethanol to produce a pure, single macromolecular component having a molecular weight of 20,000. What can you conclude from these data?

3.14.* Deduce the sequence of a pentapeptide (5 AA) from the following information. The peptide:

(a) consists of five different amino acids
(b) contains arginine and tyrosine
(c) yields DNP-Ser when treated with the Sanger reagent
(d) yields glutamic acid and cysteine when treated briefly with carboxypeptidase
(e) is cleaved into two peptides. The smaller peptide yields PTH-Ser when treated with the Edman reagent. The larger peptide yields the amino alcohol of cysteine when treated with $LiBH_4$, followed by hydrolysis. At pH 6.0, the two peptides carry opposite charges; one fragment moves toward the anode, and the other moves toward the cathode.

3.15. Consider the ionic interaction between the side-chain functional groups of aspartic acid 94 and histidine 146 in hemoglobin (Figure 3.24). Assuming that all other conditions remain constant, how would this interaction be affected if: (a) the pH is changed from 7.0 to 11.0; (b) the pH is changed from 7.0 to 3.0; (c) the distance between the two functional groups is increased?

3.16. Calf thymus histone H3 (a basic protein found in association with DNA) has 135 amino acid residues and a unit evolutionary period of 330×10^6 years. How many years would it take to bring about two amino acid changes per molecule as a result of mutation?

3.17. An oligomeric protein (MW = 72,000) consists of identical subunits. The oligomer is fully dissociated into its subunits and then reacted with 1-fluoro-2,4-dinitrobenzene. A total of 5.56 micromoles of DNP-Gly is obtained from 100 mg of protein. How many subunits are there per molecule of oligomer?

3.18. Based on Figure 3.26, estimate the factor by which the partial pressure of oxygen has to be increased in order to raise the saturation of (a) myoglobin and (b) hemoglobin (at pH 7.2) from 10 to 90%.

3.19. List the peptide fragments that you would obtain by hydrolysis of human gastrin (Table 2.5) with: (a) trypsin; (b) chymotrypsin.

3.20. A peptide's composition is: Asp Gly_2 His_2 Leu_2 Phe Tyr. Partial hydrolysis of the peptide yields:

1. Gly-Leu
2. His-Tyr
3. Gly-Leu-Asp
4. Gly-His-Phe
5. Asp-His-Tyr
6. Leu-Gly-Leu
7. Tyr-Gly-His-Phe

Deduce the sequence of the original peptide.

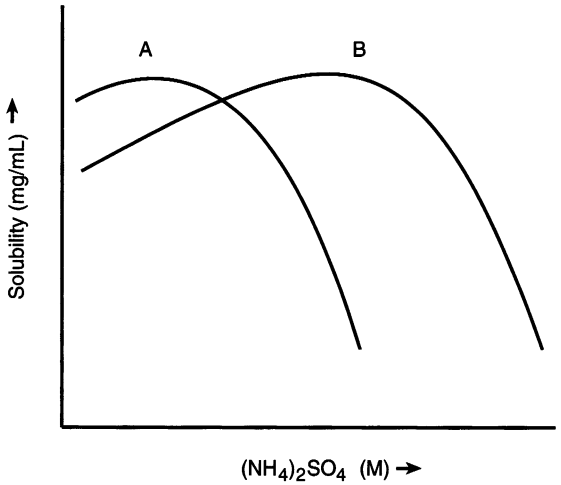
3.21. An α-helical segment consists of 180 amino acid residues. How many turns are there in this segment, and how long is it?

3.22.* Given that denaturation of an α-helical segment usually involves cooperative interactions and usually leads to the specific rotation (see Appendix B) becoming more levorotatory (negative), draw a graph, plotting specific rotation as a function of pH, for polylysine and polyglutamic acid.

3.23. After how many turns will the α-helix of polyalanine repeat itself exactly, that is, have all of the atoms in the same precise orientations?

3.24. How many different forms of tetrameric hemoglobin could occur if the α and β subunits associated, not as indicated in the chapter, but rather in a linear fashion?

3.25. Two proteins, A and B, have the solubility behavior shown. Suggest a way of separating the two proteins.



3.26.* What is the approximate growth rate in a protofibril of human hair that is increasing in length by 20 cm per year in terms of (a) amino acids incorporated and (b) peptide bonds synthesized per second?

3.27.* An unknown polypeptide was divided into two parts. One part was hydrolyzed with trypsin, and the other part was treated with cyanogen bromide. The fragments obtained are listed below. What was the amino acid sequence of the original polypeptide?

I. Trypsin hydrolysis
   1. Asp-Met-Lys
   2. Ala-Met-Glu-Ser-Lys
   3. Leu-His-Arg
   4. Phe-Gly-Met-Lys

II. Cyanogen bromide treatment
   5. Glu-Ser-Lys-Asp-Met
   6. Leu-His-Arg-Ala-Met
   7. Lys-Phe-Gly-Met
   8. Lys

Enzymes

4

Enzymes are the catalysts of biochemical reactions. They serve to make reactions in living systems proceed at accelerated rates. Without the action of enzymes, reactions would proceed too slowly, and life, as we know it, would not be feasible.

The discovery of enzymes occurred around the middle of the 19th century and is generally attributed to Anselme Payen and Jean-François Persoz, working in a sugar factory in Paris. In 1833, these chemists reported on a reusable factor, prepared from a malt extract, that was capable of converting starch into sugar. Payen and Persoz called the factor *diastase,* but we now term it *amylase*. The name *diastase* ultimately led to the current naming of enzymes by use of the suffix "-ase," first proposed by E. Duclaux in 1898.

Shortly after the discovery of diastase, Pierre Berthelot successfully separated a fraction from a cell-free extract of yeast that catalyzed the conversion of sucrose to glucose and fructose (1860). During the second half of the nineteenth century, controversy raged around the views of Justus von Liebig, who held that yeast fermentation and similar processes were due to the action of chemical substances, and those of Louis Pasteur, who maintained that fermentation was inseparable from living cells. The two scientists referred to the active components as "unorganized ferments" (extracted enzymes) and "organized ferments" (microorganisms), respectively. The Pasteur–Liebig controversy came to an end when Eduard Buchner, in 1897, obtained a cell-free extract from yeast that was capable of carrying out fermentation.

The name *enzyme* (from the Greek, meaning "in yeast") was proposed by Wilhelm Kühne in 1878, but the protein nature of enzymes was not fully established until J. B. Sumner, in 1926, crystallized the enzyme urease from jack beans. He showed that despite multiple crystallizations of the preparations, the ratio of enzymatic activity to protein mass remained constant. This provided strong evidence that the protein molecules, rather than a contaminant, were responsible for the catalytic activity observed. By coincidence, urease, the first purified enzyme, happened to catalyze the breakdown of urea, the first organic compound synthesized

in the laboratory. Shortly after Sumner's work, John Northrop and his colleagues isolated crystalline proteolytic enzymes. By about 1940, isolation and purification of intracellular enzymes was in full swing.

4.1. GENERAL PROPERTIES OF CATALYSTS

We can characterize all **catalysts,** including enzymes, by four general properties:

1. A catalyst is a substance, other than a reactant or a product, that alters the rate of a chemical reaction. In common usage, the term catalyst is synonymous with *positive catalyst* and refers to a substance that *increases* the reaction rate.

2. A catalyst accelerates both the forward and the reverse reaction. A catalyst lowers the *energy of activation* (see Section 4.2) for both the forward and the reverse reactions. Accordingly, the reaction proceeds faster in either direction and equilibrium is established more rapidly. Thus, a catalyst shortens the *time* to reach equilibrium; it does not, however, alter the *position* of the equilibrium (value of the equilibrium constant). Nor does a catalyst alter a reaction's *overall free energy change,* which is determined by the initial and final states (energies of reactants and products).

3. A catalyst can only function in a reaction that would occur without it. For a catalyst to be effective, the reaction must proceed in its absence, if only at an infinitesimally slow rate. An **endergonic** reaction that has a positive free energy change ($\Delta G > 0$) is thermodynamically not feasible and does not proceed spontaneously. Adding a catalyst, be it an enzyme or some other substance, cannot make such a reaction go. By the same token, if a reaction is not feasible because of chemical considerations, adding a catalyst cannot make that reaction proceed either.

4. A catalyst appears unchanged at the end of the reaction. A catalyst can be recovered in its original form and is not permanently altered during the reaction. Most enzymes behave in this fashion, but some undergo inactivation after catalyzing a reaction with certain toxic analogs of their natural **substrates** (a substrate is a reactant acted upon by an enzyme). These toxic analogs are called "suicide substrates" because they cause the enzyme to "commit suicide."

In addition to these four general properties, enzymes have other characteristics that set them apart as special kinds of catalysts. First of all, enzymes (excluding catalytic RNAs; see Section 7.1) are proteins and have all of the attributes of these biomolecules (charge, solubility, denaturation, and so on). Additionally, enzymes have unique properties that we will discuss under three broad headings related to the great *specificity* of enzymes, their unusual *efficiency,* and their detailed *regulation.*

4.2. ENZYME SPECIFICITY

Enzyme action comprises three essential parts. First, the substrate must bind to the enzyme. Next, the enzyme catalyzes chemical alterations of the substrate, including breaking and/or formation of chemical bonds. Lastly, the products of the reaction are released from the enzyme.

An enzyme catalyzes only certain, more or less related, reactions. We call this selectivity in the type of reaction catalyzed **enzyme specificity.** The degree of specificity varies. Some enzymes catalyze a reaction in which many compounds can serve as substrates. Other enzymes show greater specificity, being effective with only a few substrates. *Absolute specificity* exists when an enzyme catalyzes only one reaction, involving a single substrate in the forward reaction (a different substrate is involved in the reverse reaction).

The proteolytic enzymes aminopeptidase, carboxypeptidase, trypsin, and chymotrypsin have varying degrees of specificity (Table 3.3). *Urease* is an enzyme that has absolute specificity. As far as we know, this enzyme catalyzes only the hydrolysis of urea to ammonia and carbon dioxide:

$$NH_2CONH_2 + H_2O \rightarrow 2NH_3 + CO_2$$

4.2.1. The Active Site

4.2.1A. Lock-and-Key Theory. Enzyme specificity derives from the unique binding interaction between enzyme and substrate. Two models describe this interaction. According to one model, the substrate binds to the enzyme at a certain part of the enzyme, the **active site** (Figure 4.1), but only if there exists *structural complementarity* between the substrate and the active site. The substrate must fit stereochemically into the active site, thereby forming an **enzyme–substrate complex** as an intermediate in the reaction sequence. This complex is generally, but not always, a noncovalent complex. Only compounds having the proper stereochemical configuration of the substrate, or a configuration reasonably close to it, can bind at the active site and undergo reaction.

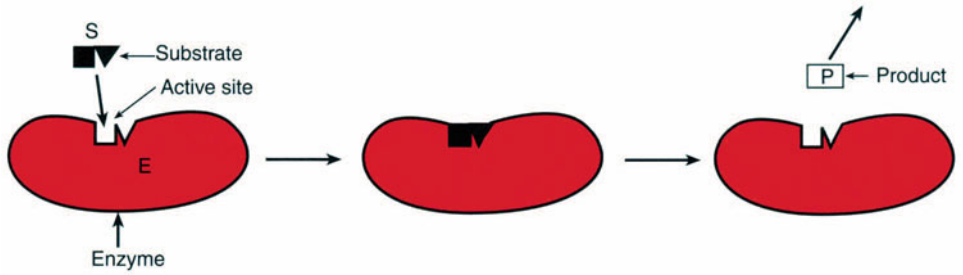


Figure 4.1. Lock-and-key theory. The substrate fits into a preformed active site on the enzyme.

In this model of enzyme action, substrates either fit or do not fit into a *preformed* active site. We refer to this concept as the **lock-and-key theory.** The substrate fits into the active site much as a key fits into a lock. On this basis, we write the fundamental reaction in enzymology as

$$E + S \rightleftarrows ES \rightleftarrows E + P$$

where E is an enzyme, S is a substrate, ES is the enzyme–substrate complex, and P is the product.

4.2.1B. Induced Fit Model. Some enzymes interact with their substrates differently. In these cases, the enzyme does not possess a preformed active site. Instead, exposing the enzyme to the substrate induces a conformational change in the enzyme that causes the active site to be formed or exposed. Only after this has occurred can the enzyme bind the substrate in a manner that results in catalysis. We say that such enzymes have a **flexible active site** (Figure 4.2) and that substrate binding involves an **induced fit.** We can explain the action of many enzymes by the lock-and-key theory, but for others we must invoke the induced fit model.

Regardless of whether an enzyme possesses a preformed or an inducible active site, that site is often formed by two or more segments of the convoluted polypeptide chain(s) coming close together in the tertiary structure of the enzyme. The structure and function of the active site depend on the interplay of amino acid residues at different locations along the chain(s). This means that even a small conformational change in the enzyme may be enough to disrupt the precise architecture of this site and cause the enzyme to lose its biological activity. Knowing the structure of the active site makes it possible for us to design drugs that activate or inhibit specific enzymes.

4.2.1C. Polyaffinity Theory. One aspect of the interaction between substrate and enzyme relates to the stereospecificity of enzymes as explained by the **polyaffinity theory.** Frequently, an organic compound, though lacking in chiral carbons, has the *potential* to react

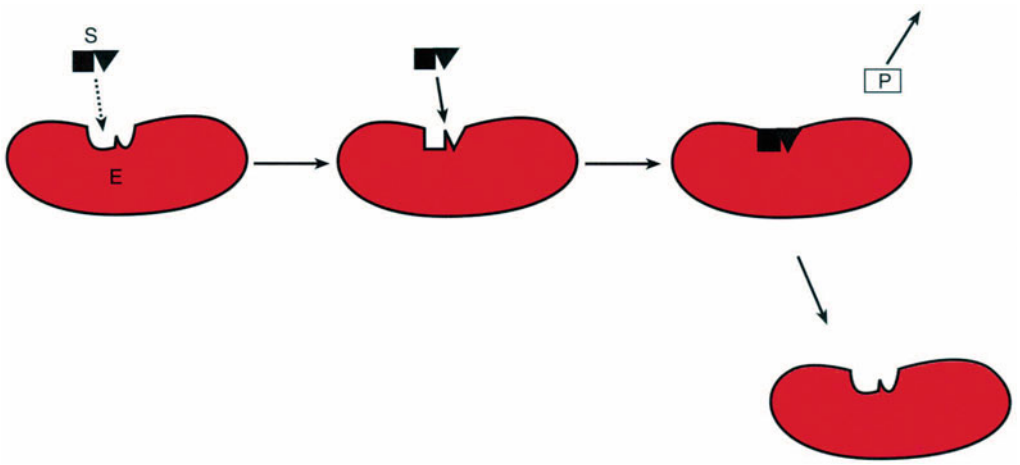
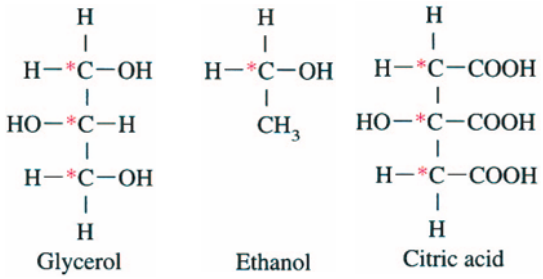


Figure 4.2. Induced fit model. The substrate causes a conformational change in the enzyme, resulting in formation of the active site.

asymmetrically if the enzyme acting on it has an asymmetric active site. We call organic compounds that can act in this fashion **prochiral compounds,** and the carbons capable of reacting asymmetrically **prochiral carbons.** A prochiral carbon represents a carbon atom that has *two identical and two different substituents* attached to it; replacing one of the identical groups with a different one results in formation of a chiral center. In the following formulas, prochiral carbons are designated by an asterisk:



In ordinary chemical reactions, which involve the two identical groups of a prochiral carbon, 50% of the reaction proceeds with one of these identical groups and 50% with the other group. Equal amounts of two products result unless other factors operate.

An enzyme, however, *generally reacts with only one of the two identical substituents* so that only one product results. Historically, this phenomenon was discovered in 1948 when A. Ogston studied the conversion of citric acid in the citric acid cycle. Ogston found that only one of two possible chiral products formed from citric acid. This led him to propose that the enzyme's active site was *asymmetric* and contained a *minimum* of three sites to which the citric acid molecule must bind in a specific manner.

With the requirement of at least three points of contact between the substrate and the enzyme, the two identical groups of a prochiral carbon become fixed in space and occupy unique positions (Figure 4.3). One group is positioned correctly and undergoes enzyme-catalyzed reaction; the other group is positioned incorrectly and does not undergo reaction.

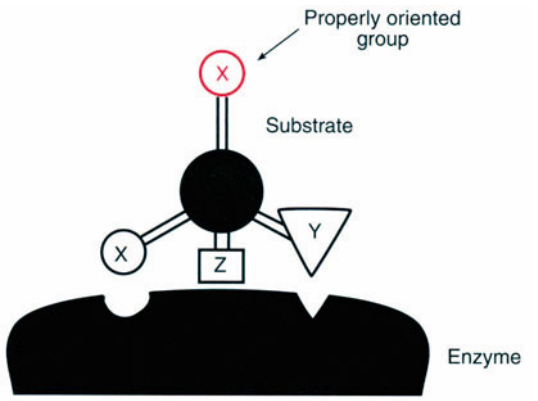


Figure 4.3. Polyaffinity theory. Enzyme and substrate have a minimum of three points of contact so that only one of two identical groups of a prochiral carbon is oriented properly to undergo reaction.

4.2.2. Enzyme Classification

We name and classify enzymes according to the reaction that they catalyze. A special Enzyme Commission (EC) of the International Union of Biochemistry (IUB) assigns each enzyme a *recommended* or *common name* and a *systematic name.* The recommended name may be the previously used trivial name (e.g., trypsin, chymotrypsin) or a name that includes the substrate (e.g., urease, sucrase) or the reaction catalyzed (e.g., methionine decarboxylase, lactate dehydrogenase). We use systematic names when ambiguity must be minimized. Systematic names consist of the name(s) of the substrate(s) followed by a word with the suffix "ase" specifying one of the six main classes of enzymes (Table 4.1). In enzyme nomenclature, as well as in other biochemical terminology, we commonly name weak acids as their salts. Hence we use "lactate" rather than "lactic acid." This convention recognizes that at physiological pH, these acids exist in deprotonated forms, as the corresponding anions.

The Enzyme Commission has subdivided the six main classes further. Each enzyme is assigned four *classification numbers,* separated by periods and preceded by the letters EC. The first number represents the main class, based on reaction type; the second stands for the subclass, based on the substrate(s) for the reaction; the third, the sub-subclass, identifies unique properties of the enzyme; and the fourth number is an arbitrarily assigned serial number. A partial breakdown of hydrolases (Main Class 3) illustrates these principles:

Class 3:	Hydrolases
Subclass 4:	Acting on peptide bonds
Sub-subclass 17:	Metallo-carboxypeptidases
Serial no. 1	Carboxypeptidase A (EC 3.4.17.1)
Subclass 4:	Acting on peptide bonds
Sub-subclass 21:	Serine proteinases
Serial no. 4	Trypsin (EC 3.4.21.4)
Subclass 5:	Acting on nonpeptide C—N bonds
Sub-subclass 1:	In linear amides
Serial no. 5	Urease (EC 3.5.1.5)

Table 4.1. Enzyme Classification

Enzyme class and example[a]	Type of reaction and example
1. Oxidoreductases *Example:* Alcohol dehydrogenase (alcohol:NAD^+ oxidoreductase), EC 1.1.1.1	Oxidation–reduction reactions $CH_3CH_2OH + NAD^+ \rightarrow CH_3CHO + NADH + H^+$ Ethanol → Acetaldehyde
2. Transferases *Example:* Aspartate aminotransferase (L-aspartate:α-ketoglutarate aminotransferase), EC 2.6.1.1	Transfer of functional groups $R{-}CH(NH_3^+){-}COO^- + R'{-}CO{-}COO^- \rightarrow R{-}CO{-}COO^- + R'{-}CH(NH_3^+){-}COO^-$ Amino acid → Keto acid′ ; Keto acid → Amino acid′
3. Hydrolases *Example:* Dipeptidase (dipeptide hydrolase), EC 3.4.13.11	Hydrolysis reactions Amino acid–amino acid + $H_2O \rightarrow$ 2 amino acids
4. Lyases *Example:* Pyruvate decarboxylase (pyruvate carboxy-lyase), EC 4.1.1.1	Elimination of a group to form a double bond, or addition of a group to a double bond $CH_3{-}CO{-}COO^- + H^+ \rightarrow CH_3{-}CHO + CO_2$ Pyruvate → Acetaldehyde
5. Isomerases *Example:* Amino acid racemase (Amino acid racemase), EC 5.1.1.10	Isomerization reactions L-Amino acid → D-amino acid
6. Ligases *Example:* Pyruvate carboxylase (pyruvate:carbon dioxide ligase), EC 6.4.1.1	Bond formation by joining two molecules, or two parts of one molecule, coupled with hydrolysis of ATP or similar energy-rich pyrophosphate bonds $CH_3{-}CO{-}COO^- + CO_2 + ATP^{4-} + H_2O$ Pyruvate ↓ $^-OOC{-}CO{-}CH_2{-}COO^- + ADP^{3-} + P_i^{2-} + 2H^+$ Oxaloacetate

[a] For each enzyme, the common name is listed first, followed by the systematic name (in parentheses) and the classification number.

4.2.3. Mechanism of Action

According to **transition-state theory,** a chemical reaction proceeds when the reactants become activated by conversion to an energy-rich intermediate, called an **activated complex** or **transition state,** which subsequently breaks down to form products. The activated complex is a chemical entity even if it has only a fleeting existence; it constitutes a transient state, an unstable complex stabilized by noncovalent bonds. Conversion of reactants to activated complexes, and from there to products, can be depicted graphically by plotting *free energy (G)* as a function of a *reaction coordinate* that describes the *progress of the reaction* (Figure 4.4).

The change (Δ) in free energy, as for all thermodynamic functions, refers to the difference between *final* and *initial states.* In this case:

$$\Delta G_{\text{reaction}} = G_{\text{products}} - G_{\text{reactants}}$$

For a reaction to occur spontaneously, reactants must have a higher energy level than products. The reaction is **exergonic;** its *ΔG is negative.* However, despite the favorable energy difference between reactants and products, the reactants must be raised to a still higher energy level—that of the activated complex—before they can be converted to products. We call the additional amount of ener-

gy that needs to be channeled into the reactants **energy of activation.** The requirement of an energy of activation has a beneficial consequence. If reactants did not require activation, all substances capable of spontaneous reactions would be unstable, and life would be impossible.

As indicated in Figure 4.4, *enzymes function by lowering the energy of activation.* It takes less energy to raise the reactants to the level of an activated complex in the presence of an enzyme than in its absence. You can understand why lowering the energy of activation accelerates a reaction from inspection of a *Maxwell distribution,* a plot of the number of molecules as a function of their energy (Figure 4.5). Assume that the energy of activation for a given reaction has a value of E_a. In that case, only the few molecules having energies equal to, or greater than, E_a can immediately undergo reaction; all others must first be activated to that level. If the energy of activation is now lowered to E'_a, a much larger fraction of molecules can react immediately; the reaction is accelerated. Moreover, since less energy is required to activate molecules to E'_a than to E_a, any added energy will activate a *larger number* of molecules so that, again, the reaction will proceed more rapidly.

Even a small lowering of the energy of activation results in a significant increase in reaction rate, as you can see from Table 4.2. Typically, enzymes enhance the rate of an uncatalyzed reaction by a factor of about 10^6 to 10^{20} (10^7 for carbonic anhydrase, 10^{14} for urease, and 10^{15} for catalase). To put these numbers in perspective, consider that with a rate enhancement of 10^{15}, catalase can do in *one second* what it would take the uncatalyzed reaction *31.7 million years* to do!

The free energy of a reaction is related to two other

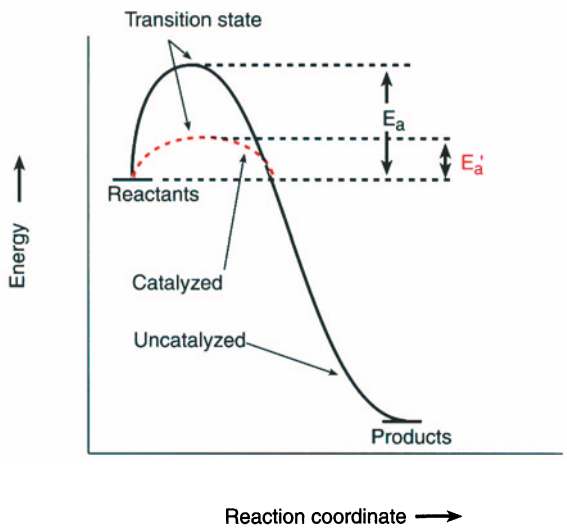


Figure 4.4. Activation energy of a catalyzed (E'_a) and an uncatalyzed (E_a) chemical reaction.

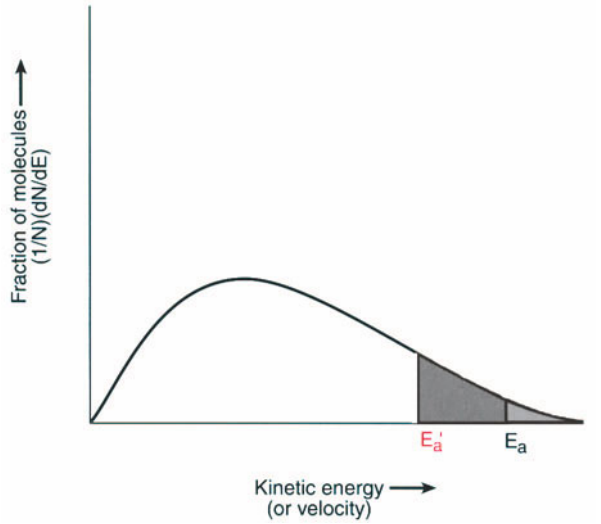


Figure 4.5. Maxwell distribution of the kinetic energy of molecules. N is the number of molecules of a gas or a liquid that have kinetic energy E. E_a and E'_a are the activation energies in the absence and presence of a catalyst, respectively.

thermodynamic functions, *entropy (S)* and *enthalpy* or *heat content (H),* by the equation

$$\Delta G = \Delta H - T\Delta S \tag{4.1}$$

where T is the absolute temperature. Because formation of the activated complex from the reactants is itself a chemical reaction, the energy of activation is identical to the free energy change for formation of the activated complex ($\Delta G^{\ddagger}$). Therefore, we can characterize the formation of an activated complex from reactants thermodynamically by

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger} \tag{4.1a}$$

where ‡ refers to the reaction that yields the activated complex. Consequently, a decrease in the energy of activation

Table 4.2. Energy of Activation (E_a) for the Decomposition of H_2O_2[a]

Catalyst	Rate[b] (mol liter^{-1} s^{-1})	E_a (kJ mol^{-1})
None	10^{-8}	71.1
HBr	10^{-4}	50.2
Fe^{2+}/Fe^{3+}	10^{-3}	41.8
$Fe(OH)_2$-triethylenetetramine	10^{3}	29.3
Catalase	10^{7}	8.4

[a]From *Physical Chemistry: Principles & Applications in Biological Sciences,* 3/e, Tinoco/Sauer/Wang © 1995. Reprinted by permission of Prentice-Hall, Inc., Upper Saddle River, NJ.

[b]Rate = change in $[H_2O_2]$ with time, calculated for 1M concentration of H_2O_2, 1M concentrations of nonenzymatic catalysts, and 1M concentration of active sites for catalase. Catalase catalyzes the decomposition of hydrogen peroxide according to the equation $2H_2O_2 \rightarrow 2H_2O + O_2$.

($\Delta G^{\ddagger}$) can be produced by changes in enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) for formation of the activated complex. Enzymes function by decreasing $\Delta H^{\ddagger}$, by increasing $\Delta S^{\ddagger}$, or by doing both.

4.3. ENZYME EFFICIENCY

As illustrated by the enzyme catalase (Table 4.2), enzymes are incredibly efficient. They catalyze reactions at rates that cannot be matched by simple catalysts used in organic chemistry, such as platinum, palladium, and nickel.

4.3.1. Turnover Number

We commonly describe enzyme efficiency in terms of a **turnover number,** defined as *the number of moles of substrate transformed into products per mole of enzyme per second.* If the enzyme has more than one active site, the turnover number refers to the transformations per mole of active site.

Catalase is one of the most active enzymes known. It has a turnover number of approximately 40×10^6. This means that one *molecule* of catalase catalyzes the decomposition of *40 million molecules* of hydrogen peroxide per second! Table 4.3 lists turnover numbers of several other enzymes. Most enzymes have a turnover number in the range of 1–10,000.

We compute the turnover number by dividing the *maximum velocity* of an enzyme, V_{max} (see Section 4.5), by the *total enzyme concentration,* $[E_t]$, and frequently report it as a number with dimensions of reciprocal seconds (E = enzyme, AS = active site):

$$\text{Turnover number} = \frac{V_{max}}{[E_t]} = \frac{\text{moles (S} \rightarrow \text{P)} \times \text{s}^{-1} \times \text{liter}^{-1}}{\text{moles (E or AS)} \times \text{liter}^{-1}} = \text{s}^{-1} \qquad (4.2)$$

Table 4.3. Turnover Numbers of Some Enzymes

Enzyme	Turnover number (moles substrate/mole enzyme per second)
Catalase	40,000,000
Carbonic anhydrase	600,000
Acetylcholinesterase	20,000
Urease	10,000
Lactate dehydrogenase	1,000
β-Galactosidase	208
Chymotrypsin	100
DNA polymerase I	15
Tyrosyl-tRNA synthetase	8
Tryptophan synthase	2
Pepsin	0.5

4.3.2. Enzyme Units

Because many enzymes have high turnover numbers, descriptions of enzyme solutions in terms of common measures of concentrations are frequently not very useful. Instead, we express enzyme concentrations in terms of defined units based on their activity. One **enzyme unit** is *the amount of enzyme causing the transformation of one micromole of substrate per minute under defined conditions.* By measuring both enzymatic activity and total protein concentration, one can derive the **specific activity,** defined as *the number of units per milligram of total protein.* In the course of an enzyme purification, one progressively removes more and more extraneous protein so that the specific activity of the desired enzyme increases. However, the **total activity,** the total number of enzyme units, frequently decreases at the same time owing to losses at various steps of the isolation. Table 4.4 illustrates a few steps in the purification of a bacterial enzyme. Enzymes and other proteins vary in the degree of purification they require. Some necessitate a hundredfold or even several hundredfold purification.

4.3.3. Factors Contributing to Catalysis

Biochemists believe that the tremendous efficiency of enzymes results from a combination of different factors. Several possible ones are outlined below.

4.3.3A. Proximity and Orientation. In order for a reaction to take place, two substrate molecules (or one substrate molecule and a catalytic group) must approach each other closely (*proximity*) and must be properly aligned with respect to the susceptible bonds (*orientation*). Occurrence of a reaction at the enzyme's active site facilitates both of these requirements. The probability that two substrate molecules are suitably close to each other and have their functional groups properly aligned is much greater when the substrates fit stereochemically into an active site than when they merely collide randomly in solution. Note, however, that the enzyme still must collide with and bind the substrate initially.

Binding of substrate molecules to the active site also increases the *time* during which the molecules are in suitable orientation for undergoing reaction as compared to that available during random collisions in solution. We term this concept *substrate anchoring.*

Lastly, the fact that substrate molecules have a tendency to become bound to the active site results in an increase in the local concentration of substrate. It has been calculated that the effective concentration of substrate near the active site might be some 10^5 times greater than its concentration in the bulk enzyme solution. Since the rate of a chemical reaction is proportional to the concen-

Table 4.4. Purification of DNA Polymerase III Holoenzyme from *E. coli*[a]

Step	Total protein (mg)	Total activity (units × 10^{-6})	Specific activity (units × 10^{-3}/mg)	Purification[b] (fold)	Yield[c] (%)
I. Cell-free extract[d]	293,000	—	—	—	—
II. Ammonium sulfate precipitation	2,040	102	50	1.0	100
III. Backwash	185	49	260	5.2	48
IV. Heparin-agarose chromatography	7.2	16	2,200	44	16
V. DEAE-cellulose chromatography	0.39	12	30,000	600	12
VI. Heparin-agarose chromatography	0.18	6	33,000	660	6

[a] Adapted with permission from S. Maki, and A. Kornberg, *J. Biol. Chem.* 263:6555–6560 (1988).
[b] Purification = specific activity at a given step/specific activity of the initial extract.
[c] Yield = (total activity at a given step/total activity of the initial extract) × 100.
[d] 710 g of *E. coli* were used for the purification. Frequently, purification and yield are computed on the basis of the cell-free extract, rather than on a subsequent fraction as done here.

trations of the reactants, a very large enhancement of rate can be expected in such a localized area of high concentration.

4.3.3B. Electrostatic Effects. Electrostatic interactions are stronger in organic solvents than in water because of lower dielectric constants (see Eq. 1.10). Scientists postulate that this property could be used by enzymes to stabilize polar transition states in sections of the molecule where the protein structure provides a low dielectric medium. Stabilization may involve fixed dipoles in the enzyme, such as the NH groups of the polypeptide chain. Other electrostatic interactions occur when enzymes use part of their own structure or bound ions, rather than bulk water, to solvate transition states.

4.3.3C. Entropic Factors. Biochemists think that one of the most important factors in enzyme catalysis is entropy. Catalyzed reactions in solution are slow because bringing together catalyst and substrate involves a considerable loss of entropy (an increase in order). Enzyme reactions take place in the confines of the enzyme–substrate complex, in which catalytic groups form an integral part. Because reactants, transition states, and products are all bound to the enzyme, the entropy losses are much smaller than for reactions in solution.

4.3.3D. Strain and Distortion. Binding of substrate to an enzyme's active site may be associated with a conformational change in the substrate, the enzyme, or both. This binding may distort the substrate, destabilize some of its bonds, and make it easier to break or form bonds during the reaction. Alternatively, substrate binding may induce a conformational change in the enzyme, thereby enhancing its catalytic properties. Recently, the modified concept of *transition-state stabilization* has gained favor. According to this concept, it is not so much that the substrate is destabilized, but rather that the transition state makes better contacts with the enzyme than the substrate.

4.3.3E. Acid–Base Catalysis. Interaction of enzyme and substrate frequently includes acid–base catalysis for one or more steps of the mechanism. *Specific acid–base catalysis* refers to rate enhancement produced by H^+ or OH^- ions. *General acid–base catalysis* refers to rate enhancement produced by various Brønsted acids and bases. Specific acid–base catalysis appears to be of minor importance in enzyme reactions, but general acid–base catalysis is very likely to be involved in many cases. Of the large number of acidic and basic groups that enzymes carry, those located at or near the active site may contribute significantly to the total catalytic efficiency of the enzyme.

4.3.3F. Covalent Catalysis. Some enzymes form an intermediate in which the substrate is linked covalently to the enzyme. Such **enzyme–substrate compounds** differ from ordinary enzyme–substrate complexes and may provide a more efficient mechanism for the reaction. For covalent catalysis to be effective, the enzyme–substrate compound must be unstable, so that it is rapidly converted to products. (A stable enzyme–substrate compound would serve to inactivate the enzyme.) The catalytic mechanism of chymotrypsin, described at the end of this chapter, illustrates proximity and orientation as well as acid–base and covalent catalysis.

4.4. REGULATION OF ENZYME ACTIVITY I: ENZYME PROPERTIES

Enzymes differ from ordinary catalysts of organic chemistry in that their activity is affected by many factors and controlled in various ways. All mechanisms of enzyme regulation fall into one of two basic categories (Figure 4.6). One category of regulatory modes includes those that affect the *amount of enzyme*—the actual number of enzyme molecules per cell at any given time. Enzyme amount reflects intracellular rates of protein synthesis and degradation, topics that we will cover in Chapter 19. The second category of regulatory modes comprises those that affect *enzyme activity*—the actual effectiveness of the enzyme molecules of any given cell. In this section, we will discuss the many different variables that influence enzymatic activity as grouped in Figure 4.6.

4.4.1. Rate, Concentration, and Reaction Time

4.4.1A. Rate. The **rate** or **velocity** of a reaction (v) refers to the speed with which it proceeds. The rate represents the change in concentration of substrate (S) or product (P) with time:

$$v = \frac{d[\mathrm{P}]}{dt} = -\frac{d[\mathrm{S}]}{dt} \tag{4.3}$$

where d denotes an infinitesimal change. The negative sign in front of the $d[\mathrm{S}]/dt$ term indicates that the concentration of S decreases with time. In practice, the rate may be expressed as

$$v = \frac{\Delta[\mathrm{P}]}{\Delta t} = -\frac{\Delta[\mathrm{S}]}{\Delta t} \tag{4.4}$$

where Δ denotes a finite small change:

$$\Delta[\mathrm{P}] = [\mathrm{P}]_2 - [\mathrm{P}]_1; \quad \Delta[\mathrm{S}] = [\mathrm{S}]_2 - [\mathrm{S}]_1; \quad \Delta t = t_2 - t_1$$

In a plot of substrate or product concentration as a function of time (Figure 4.7), the rate is given by the slope of the curve at any given time. The **initial velocity** is the rate at the very earliest stages of the reaction—the slope of the curve measured at the origin.

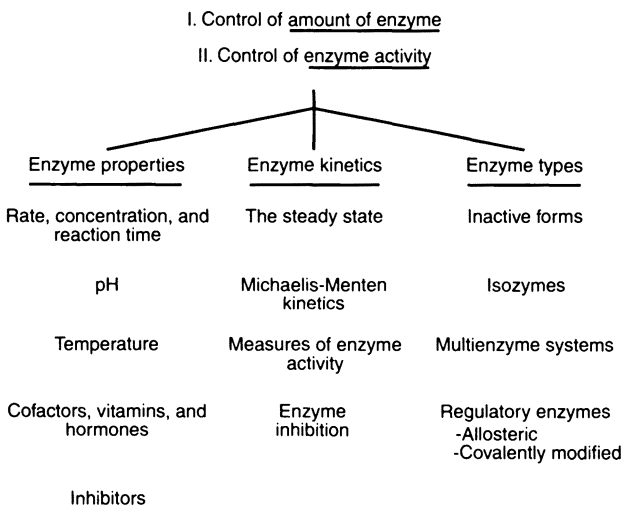


Figure 4.6. An overview of the multifaceted regulation of enzymes.

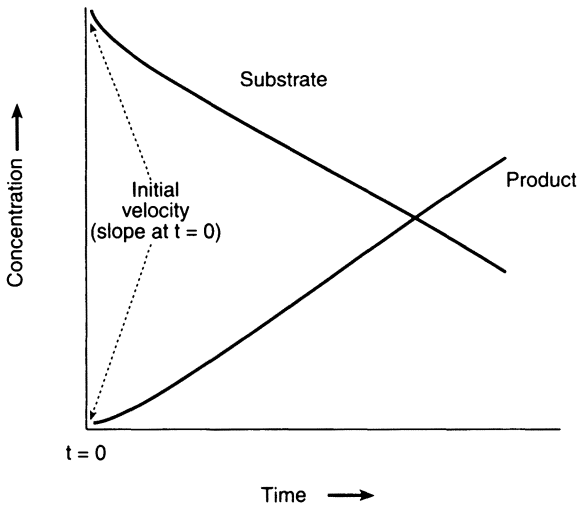


Figure 4.7. Time course of the change in substrate and product concentrations in a typical enzyme-catalyzed reaction.

4.4.1B. Concentration. Rates of enzymatic reactions depend on substrate concentration (Figure 4.8). A plot of rate (v) as a function of substrate concentration yields a hyperbolic curve, analogous to that describing the binding of oxygen to myoglobin (Figure 3.22).

Historically, it was the determination of v versus [S] curves that led to development of the concept of an enzyme–substrate complex. In an attempt to explain the observed hyperbolic shape of the curves, investigators postulated that the enzyme must combine with the substrate to form an enzyme–substrate complex. In that event, as the substrate concentration increases, more enzyme–substrate complexes can form per unit time and the rate should increase. At a certain substrate concentration, however, all available enzyme molecules have substrate bound per unit time; all exist as enzyme–substrate com-

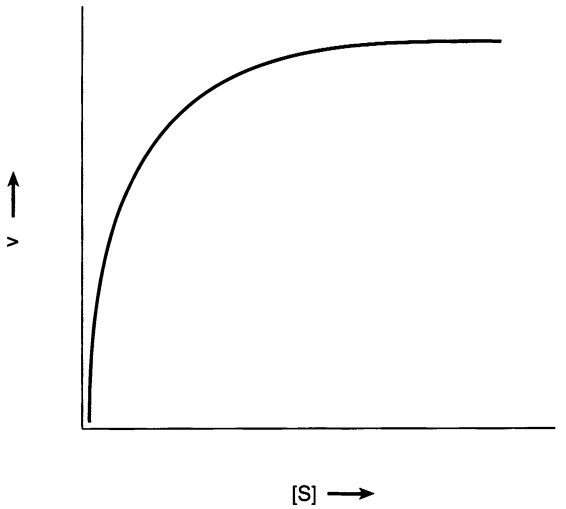


Figure 4.8. The effect of substrate concentration on the rate of an enzyme-catalyzed reaction.

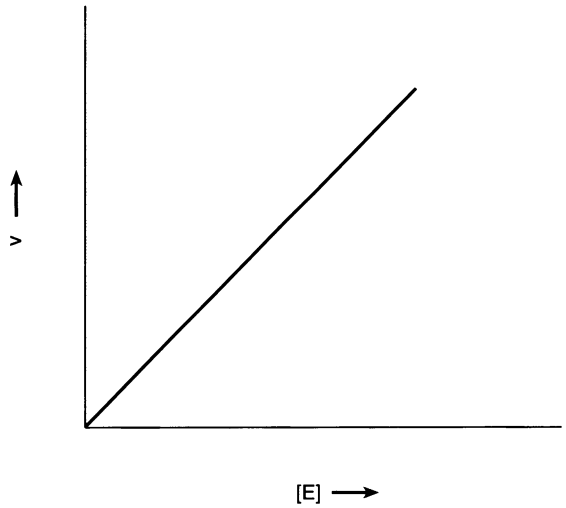


Figure 4.9. The effect of enzyme concentration on the rate of an enzyme-catalyzed reaction.

plexes. At this substrate concentration, the rate should level off as it becomes independent of substrate concentration. In this range the reaction proceeds at its *maximum velocity,* and the enzyme is *saturated* with substrate, much as oxymyoglobin is saturated with oxygen. You can see that the concept of an enzyme–substrate complex qualitatively explains the shape of the v versus [S] curve. (A mathematical treatment of this concept was developed by Leonor Michaelis and Maude Menten in 1913, building on earlier work by Victor Henri. We will discuss the Michaelis–Menten equation in Section 4.5.)

Rates of enzyme reactions also depend on enzyme concentration when the substrate is present in excess. Under those conditions, an increase in enzyme concentration leads to an increase in the number of enzyme–substrate complexes formed per unit time and, consequently, to an increase in rate (Figure 4.9).

4.4.1C. Reaction Time. The rate of an enzyme reaction changes as a function of the time the enzyme is *incubated* with substrate. The rate decreases with time and ultimately levels off (Figure 4.10). Leveling off results from a number of factors such as substrate depletion, product inhibition, enzyme denaturation, coenzyme inactivation, and increase of the reverse reaction as the product concentration builds up. Accordingly, you must determine the rate not at any arbitrary time but only at the very beginning of the reaction, when these factors are generally insignificant. In other words, it is essential to determine the *initial velocity* of enzyme reactions.

4.4.2. pH

An enzyme's catalytic activity is often limited to a relatively narrow pH range. On either side of this range, the activity may drop off so that a plot of rate as a function of pH frequently yields an approximately bell-shaped curve (Figure 4.11).

We call the pH at which an enzyme shows maximum activity the **optimum pH.** This pH does not represent a fixed molecular property of the enzyme, like amino acid sequence or molecular weight. An optimum pH is fixed only as long as all reaction variables (temperature, ionic strength, buffer type, incubation time, etc.) have specified values. The optimum pH is frequently close to neutrality, though some enzymes have unusually high or unusually low pH optima. Many reactions that occur without enzymes only at extreme pH can proceed at physiological pH in the presence of enzymes. For example, protein hy-

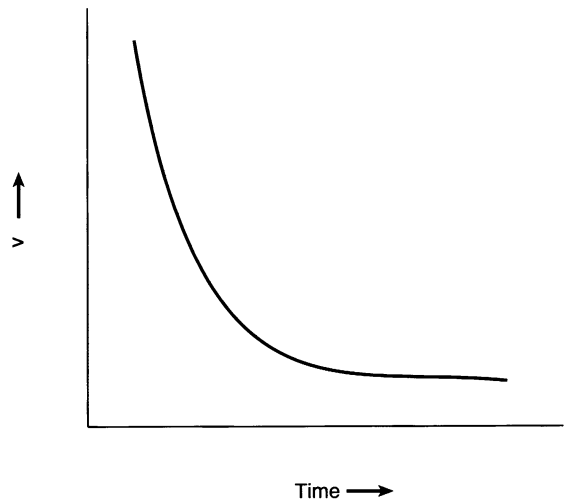


Figure 4.10. The effect of incubation time on the rate of an enzyme-catalyzed reaction.

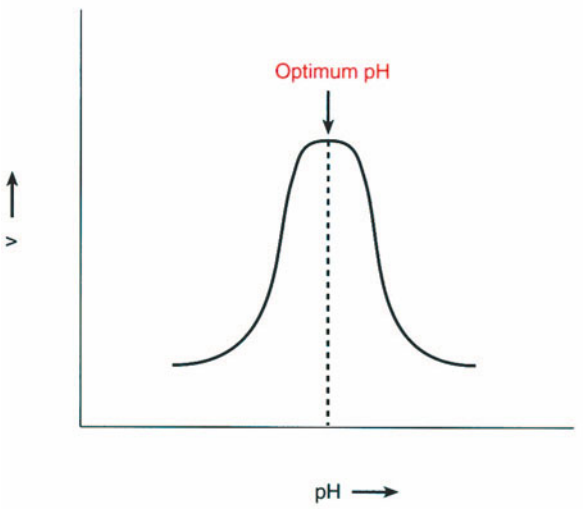


Figure 4.11. The effect of pH on the rate of an enzyme-catalyzed reaction.

drolysis in test tubes requires incubation with 12*M* HCl at 110°C for 1–2 days. By contrast, the same reaction is carried out in our digestive tract at physiological pH in a short time. Enzyme location, optimum pH, and surrounding pH are well matched in this system. Pepsin (in the stomach) has an optimum pH close to gastric pH, and trypsin (in the intestine) has an optimum pH close to intestinal pH.

As discussed below, dependence of rate on pH may result from changes in the ionization or the conformation of the enzyme or the substrate.

4.4.2A. Changes in the State of Ionization. A pH change may affect a reaction by altering the state of ionization of the enzyme, the substrate, or an intermediate. Assume, for example, that forming an enzyme–substrate complex requires electrostatic attraction between oppositely charged groups of the enzyme and the substrate. Within a relatively narrow pH range, the charges of these groups remain essentially the same so that the rate, a function of substrate binding, does not vary greatly. At a sufficiently low or high pH, however, some of these groups lose their charge, thereby impairing binding. Additionally, previously uncharged groups may acquire a charge as the pH increases or decreases. At some pH values, complementary regions of the enzyme and the substrate may carry charges of the same sign and repel each other, leading to a drop in rate.

4.4.2B. Changes in Conformation. A pH change may also affect a reaction by causing a change in the conformation of the enzyme or the substrate, or both. Altering the charges of functional groups leads to altered ionic interactions involving these groups. These new interactions may produce a conformational change at the active site, affecting the site's capacity to bind substrate. Alternatively, protein and nucleic acid substrates may undergo conformational changes and lose their capacity to bind to the enzyme.

4.4.3. Temperature

A plot of reaction rate as a function of temperature frequently yields a skewed curve, as shown in Figure 4.12. We call the temperature at which an enzyme shows maximum activity the **optimum temperature.** This temperature, like the optimum pH, is fixed only as long as all other reaction variables have specified values. The optimum temperature for enzymes falls typically in the range of 20–40°C, though some enzymes have unusually high temperature optima (*thermostable enzymes*), and some have unusually low temperature optima (*cold-stable enzymes*). Dependence of enzymatic reactions on temperature results from two opposing effects, an increase in rate and an increase in denaturation.

4.4.3A. Effect of Temperature on Reaction Rates. Rates of enzyme reactions, like those of most other chemical reactions, increase as the temperature goes up. The extent to which the rate changes depends on the magnitude of the energy of activation, as is apparent from the **Arrhenius equation:**

$$\ln k = \ln A - E_a/RT \qquad (4.5)$$

where k is the rate constant (see Section 4.5), A is a constant, E_a is the energy of activation, R is the gas constant, and T is the absolute temperature. A plot of the logarithm

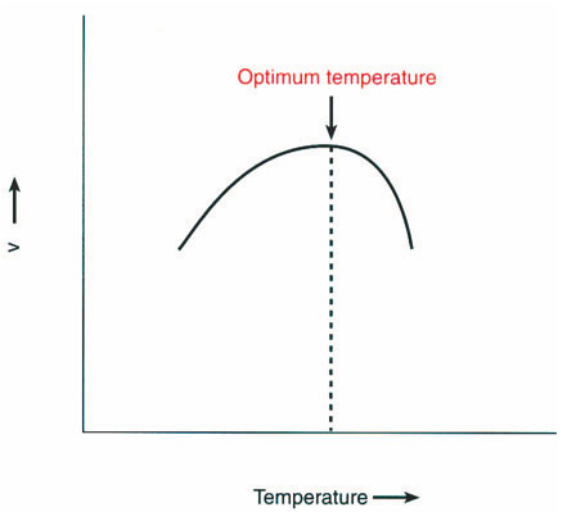


Figure 4.12. The effect of temperature on the rate of an enzyme-catalyzed reaction.

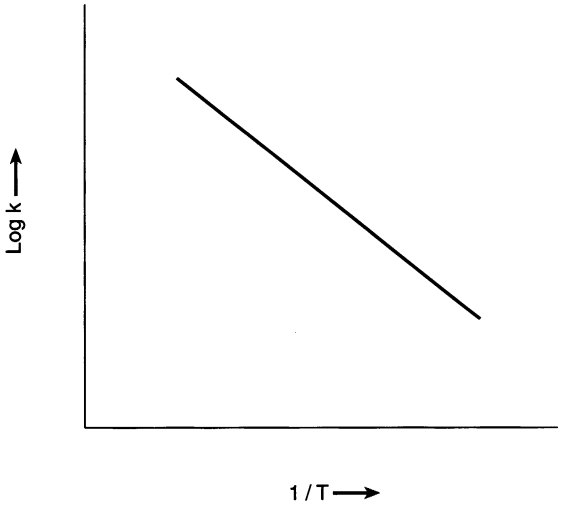


Figure 4.13. An Arrhenius plot.

of the rate constant (ln k or log k) versus the reciprocal of the absolute temperature constitutes an **Arrhenius plot** (Figure 4.13). The energy of activation can be calculated from the slope ($-E_a/R$) of such a plot.

We call the ratio of two reaction rates at two temperatures, 10°C apart, the **temperature coefficient** (Q_{10}) of the reaction:

$$Q_{10} = \frac{\text{rate at } (T + 10)}{\text{rate at } T} = \frac{10}{T_2 - T_1} \times \frac{v_2}{v_1} \qquad (4.6)$$

where v_1 and v_2 are reaction rates at temperatures T_1 and T_2, respectively. Typically, chemical reactions have a Q_{10} of about 2 so the rate doubles for every 10°C rise in temperature. Enzyme reactions have a Q_{10} of about 1.7–2.5.

The rate of an enzyme reaction will increase exponentially with temperature over the range in which Q_{10} has a fixed value (Figure 4.14A).

4.4.3B. Effect of Temperature on Enzyme Stability. If the enzyme is a protein, as most enzymes are, an increase in temperature leads to denaturation of the molecule. The higher the temperature is, the more extensive the denaturation and the smaller the amount of native, active enzyme that remains (Figure 4.14B).

The two opposing effects of temperature typically combine to produce a skewed curve for the variation of rate with temperature. Below the optimum temperature, the predominant effect of a temperature increase is a steady rise in reaction rate. Above the optimum temperature, the major effect is extensive enzyme denaturation. Despite the continued increase in reaction rate, so little active enzyme remains that the rate decreases sharply.

4.4.4. Cofactors, Vitamins, and Hormones

Frequently, an enzyme requires a non-amino acid component for catalysis (Figure 4.15). We refer to such components as **cofactors;** they may consist of either a metal ion (*activator*) or an organic compound (*coenzyme*). As for proteins in general, we use the term *prosthetic group* to describe a tightly bound cofactor, one that cannot be removed by dialysis. For enzymes that require cofactors, we term the protein component *apoenzyme* and the combination of apoenzyme and cofactor *holoenzyme.* The holoenzyme generally has full activity whereas the apoenzyme has little or no activity.

Water-soluble *vitamins* (Section 8.4) are structural components of coenzyme molecules. A vitamin deficien-

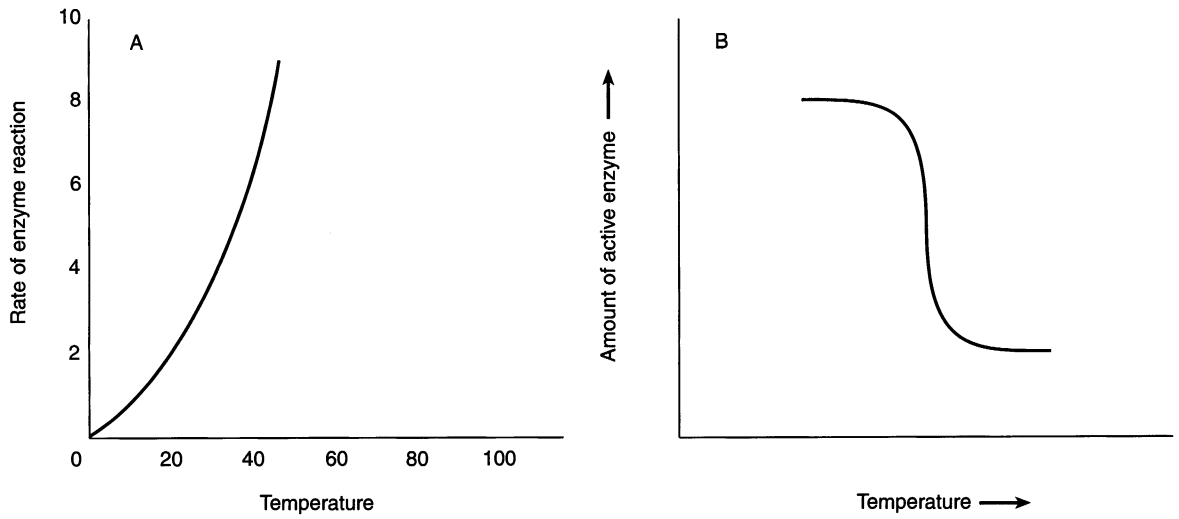


Figure 4.14. The effect of temperature on enzyme reactions. (A) Variation of reaction rate with temperature, assuming a Q_{10} of 2.0. (B) Variation of enzyme denaturation with temperature, assuming a sigmoidal curve.

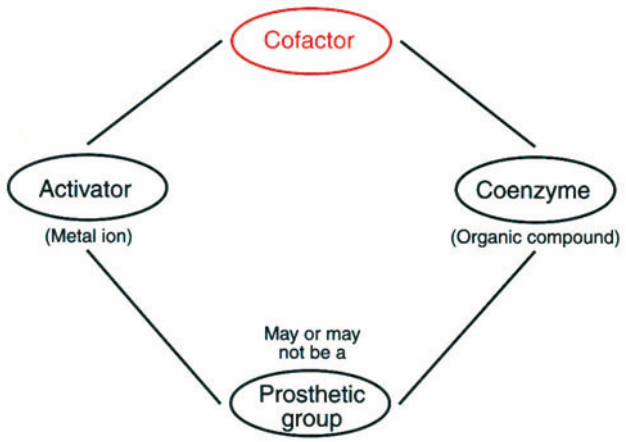


Figure 4.15. Nonprotein components of enzymes required for full catalytic activity.

cy leads to decreased synthesis of the corresponding coenzyme, impairment of the enzyme reactions requiring this coenzyme, and development of a characteristic deficiency disease.

Hormones (Section 8.2) can affect either the amount or the activity of an enzyme. *Steroid hormones* act primarily at the level of gene expression by stimulating synthesis of specific enzymes and other proteins. *Epinephrine* activates or inactivates target enzymes such as the membrane-bound adenyl cyclase. *Insulin* functions at both control levels; it stimulates protein synthesis and acts on target enzymes in carbohydrate metabolism.

4.4.5. Inhibitors

Much as enzymatic activity can be enhanced by the action of some substances, it can be decreased by the action of others. **Enzyme inhibition** refers to the decrease in rate produced by specific ions or molecules—the *inhibitors*—that combine with the enzyme. Inhibitors fall into two broad classes, reversible and irreversible inhibitors.

A **reversible inhibitor** binds to the enzyme *noncovalently,* at the active site or elsewhere. We will discuss three important groups of reversible inhibitors, called *competitive, noncompetitive,* and *uncompetitive,* in Section 4.5. For a reversible inhibitor, there exists a dynamic equilibrium between the enzyme, the inhibitor, and the *enzyme–inhibitor complex.* Affecting this equilibrium, through either physical or chemical means, may displace the inhibitor from the enzyme and restore full enzymatic activity. Removing an inhibitor by dialysis and displacing a competitive inhibitor by increasing the substrate concentration illustrate physical and chemical means, respectively, for eliminating the effect of a reversible inhibitor.

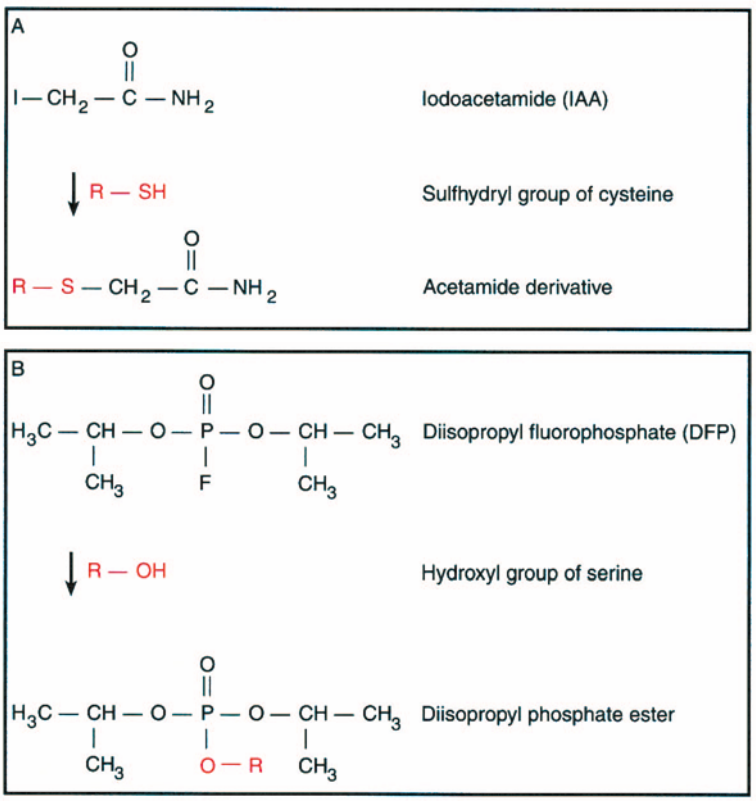


Figure 4.16. Examples of irreversible inhibitors. (A) Iodoacetamide (IAA); (B) diisopropyl fluorophosphate (DFP).

An **irreversible inhibitor** binds to the enzyme *covalently,* at the active site or elsewhere. An irreversible inhibitor cannot be removed by physical means but may be taken off by chemical means. Depending on the chemical treatment required, the enzyme may regain its full initial activity, have partial activity, or be completely inactive. Iodoacetamide (IAA) and diisopropyl fluorophosphate (DFP) are two irreversible inhibitors (Figure 4.16). IAA combines with the sulfhydryl group of cysteine; DFP combines with the hydroxyl group of serine, an amino acid that frequently occurs at or near the active site. DFP is a potent inhibitor of acetylcholinesterase, the enzyme that catalyzes the hydrolysis of acetylcholine to choline and acetate during transmission of nerve impulses. This accounts for DFP's use in nerve gas.

4.5. REGULATION OF ENZYME ACTIVITY II: ENZYME KINETICS

4.5.1. Rate Equations

The rate behavior or *kinetics* of reactions constitutes a second major way for regulating enzyme activity. Rates are proportional to reactant concentration because reactants must collide in order to undergo reaction, and increasing their concentrations increases the number of their collisions. Rates may also depend on product concentration if the reverse reaction occurs to an appreciable extent. Expressing the rate as a function of the concentrations of reactants, or reactants and products, yields a **rate equation.** Table 4.5 lists several examples of rate equations. Rate equations may be simple or complex depending upon the mechanism of the reaction. The sum of the exponents of the concentration terms provides the **reaction order** (zero order, first order, etc.). A *zero-order reaction* is one whose rate is independent of concentration (v is constant). The proportionality constant k is called the **rate constant.**

Table 4.5. Rate Equations and Reaction Orders for Various Reactions[a]

Stoichiometric reaction	Rate equation	Reaction order
$2N_2O_5 \rightarrow 4NO_2 + O_2$	$v = k[N_2O_5]$	1
L-Ile → D-Ile	$v = k$[L-Ile]	1
$Hb{\bullet}3O_2 + O_2 \rightarrow Hb{\bullet}4O_2$[b]	$v = k[Hb{\bullet}3O_2][O_2]$	2
$2NO_2 \rightarrow 2NO + O_2$	$v = k[NO_2]^2$	2

[a] From *Physical Chemistry: Principles & Applications in Biological Sciences,* 3/e, Tinoco/Sauer/Wang © 1995. Reprinted by permission of Prentice-Hall, Inc., Upper Saddle River, NJ.
[b] Hb = hemoglobin.

4.5.2. The Steady State

To be able to make quantitative calculations for a given enzyme reaction, you have to derive the appropriate rate equation. To do this, you must assume a mechanism based either on *equilibrium* conditions or on those of a *steady state.*

Often, we cannot describe biochemical reactions by equilibrium conditions. First of all, there may not be enough time for a reaction to achieve true chemical equilibrium. Second, many biochemical reactions are linked to other reactions; the product of one reaction serves as reactant for a second reaction, and so on. Consequently, a section of metabolism is typically described by a *pathway* consisting of a series of consecutive reactions, such as

$$A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow G$$

In this pathway, B and C, for example, cannot achieve true chemical equilibrium. As C is used up by being converted to D, it is replenished by conversion of B. Likewise, as B is converted to C, it is replenished by conversion of A. Components of such a pathway may, however, attain a **steady state,** a condition in which *a given component is produced as rapidly as it is consumed.* If C forms from B at the *same rate* at which it undergoes conversion to D, C attains a steady state. Despite the fact that both reactions may proceed at full speed, the concentration of C remains constant because of the balance

$$\text{Rate of C formation} = \text{Rate of C consumption}$$

A classroom with students entering and leaving provides a simple analogy to the steady-state condition. Consider such a room that has 30 students and in which every 20 seconds 10 students go out through one door and leave the building, while 10 students arrive from elsewhere to enter the classroom through a second door. The total number of students in the classroom (like the concentration of C) remains constant as a function of time because the rate of students entering equals the rate of students leaving. The steady-state concept is central to enzyme kinetics and other aspects of metabolism.

4.5.3. Michaelis–Menten Equation

The Michaelis–Menten equation represents an important rate equation in enzymology. It applies to the following enzyme-catalyzed reaction in which a substrate S is converted to product P by forming an enzyme–substrate complex, ES:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P \qquad (4.7)$$

Note that we designate rate constants for forward and reverse reactions by positive and negative subscripts, respectively.

4.5.3A. Underlying Assumptions. The Michaelis–Menten equation, as derived by George Briggs and John Haldane, involves three assumptions:

1. The enzyme–substrate complex is in a steady state; the rate at which the complex forms from E and S equals the rate at which it is consumed by being broken down to E and P in the forward reaction and to E and S in the reverse reaction.

2. Formation of product is proportional to the concentration of enzyme–substrate complex, so that $v = k_2[ES]$. This assumption implies that the breakdown of ES to E and P constitutes the **rate-determining step** (a step significantly slower than all other steps in a reaction sequence; a reaction cannot proceed faster than at the rate of this step). The assumption that $v = k_2[ES]$ further implies that the reverse reaction, from E + P to ES and governed by k_{-2}, can be ignored: no or very little product has been formed so that $k_{-2}[E][P]$ is negligible. In other words, measurements refer to those of *initial velocity.*

3. Substrate concentration greatly exceeds the enzyme concentration. Hence the decrease in substrate concentration resulting from formation of the ES complex can be ignored, and the substrate concentration is taken to be constant.

4.5.3B. Derivation of the Rate Equation. The steady-state assumption (assumption 1) can be expressed mathematically as

$$\frac{d[ES]}{dt} = 0 \tag{4.8}$$

The rate of change of [ES] with time ($d[ES]/dt$) represents the difference between the rate of ES formation and the rate of ES consumption. To obtain expressions for these rates, we must consider all of the ways in which ES may be formed and all of the ways in which it may be consumed:

$$\text{Rate of ES formation} = k_1[E_f][S] + k_{-2}[E_f][P] \tag{4.9}$$

where $[E_f]$ is the concentration of free enzyme.

Since we ignore the reverse reaction (assumption 2), the second term in Eq. (4.9) can be deleted. [S] is constant, based on assumption 3. Thus, the equation reduces to

$$\text{Rate of ES formation} = k_1[E_f][S] \tag{4.10}$$

Proceeding in a similar manner, we express the rate of ES consumption. In this equation, both terms must be retained:

$$\text{Rate of ES consumption} = k_{-1}[ES] + k_2[ES] = (k_{-1} + k_2)[ES] \tag{4.11}$$

By means of Eqs.(4.10) and (4.11), we rewrite the steady-state expression as follows:

$$\frac{d[ES]}{dt} = \underset{\text{Formation}}{k_1[E_f][S]} - \underset{\text{Consumption}}{(k_{-1} + k_2)[ES]} = 0 \tag{4.8}$$

so that

$$\underset{\text{Formation}}{k_1[E_f]\,S} = \underset{\text{Consumption}}{(k_{-1} + k_2)[ES]} \tag{4.8a}$$

Because there are two forms of the enzyme in this mechanism, *free enzyme* (E_f) and *substrate-bound enzyme* (ES), we can express the *total enzyme concentration* (E) as being equal to the sum of two terms:

$$[E] = [E_f] + [ES] \tag{4.12}$$

It follows that

$$[E_f] = [E] - [ES]$$

Substituting for $[E_f]$ into Eq. (4.8a) yields

$$k_1([E] - [ES])[S] = (k_{-1} + k_2)[ES]$$

or

$$([E] - [ES])[S] = \frac{k_{-1} + k_2}{k_1}[ES] \tag{4.13}$$

We now define a new constant, K_m, such that

$$K_m = \frac{k_{-1} + k_2}{k_1} \tag{4.14}$$

The constant $\boldsymbol{K_m}$ is called the **Michaelis constant.** Substituting this constant into Eq. (4.13) yields

$$([E] - [ES])[S] = K_m[ES]$$

or

$$[E][S] - [ES][S] = K_m[ES] \tag{4.15}$$

Since the reaction rate is proportional to the concentration of ES complex (assumption 2),

$$v = k_2[\text{ES}] \tag{4.16}$$

Substituting for [ES] from this equation into Eq. (4.15) leads to the rate equation

$$[\text{E}][\text{S}] - \frac{v}{k_2}[\text{S}] = K_m \frac{v}{k_2} \tag{4.17}$$

Rearranging Eq. (4.17) yields

$$k_2[\text{E}][\text{S}] - v[\text{S}] = K_m v \tag{4.18}$$

The **maximum velocity** (V_{max}) of the reaction occurs when all of the enzyme molecules are combined with substrate molecules per unit time. Under those conditions, the total enzyme concentration equals the concentration of enzyme–substrate complex; the concentration of free, uncombined enzyme is zero. When $[E_f] = 0$, Eq. (4.12) reduces to [E] = [ES]. Therefore, to obtain the maximum velocity, we replace [ES] with [E] in Eq. (4.16):

$$V_{max} = k_2[\text{E}] \tag{4.19}$$

Replacing the first term on the left-hand side of Eq. (4.18) with $V_{max}[S]$, we obtain

$$V_{max}[\text{S}] - v[\text{S}] = K_m v$$

which, upon rearranging, yields

$$v = \frac{V_{max}[\text{S}]}{K_m + [\text{S}]} \tag{4.20}$$

Equation (4.20) is the **Michaelis–Menten equation.**

4.5.3C. Description of the *v* versus [S] Curve. We saw earlier that a plot of rate as a function of substrate concentration yields a hyperbolic curve (see Figure 4.8). We can now show that the Michaelis–Menten equation describes a curve of this type. To do this, consider two conditions:

1. When [S] is very small compared to K_m, the [S] term may be deleted from the denominator of Eq. (4.20). The equation then reduces to

$$v = \frac{V_{max}[\text{S}]}{K_m} = k''[\text{S}] \tag{4.21}$$

Equation (4.21) is the rate equation for a first-order reaction, in which the rate is proportional to substrate concentration, because V_{max}/K_m is a constant (k''). This corresponds to the low concentration end of the v versus [S] curve (Figure 4.17).

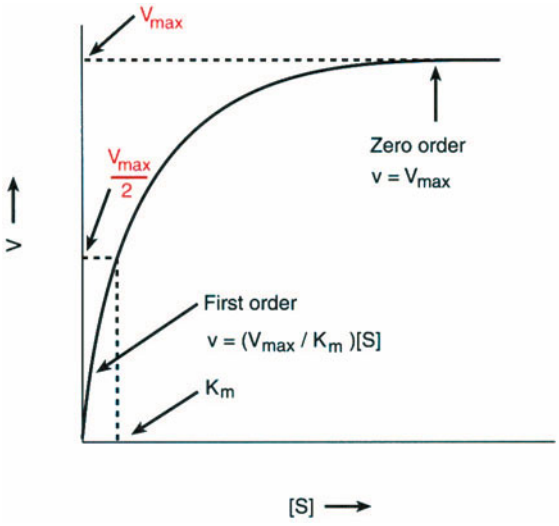


Figure 4.17. Variation of rate as a function of substrate concentration.

2. When [S] is much larger than K_m, the K_m term may be deleted from the denominator of Eq. (4.20). The equation then reduces to

$$v = \frac{V_{max}[\text{S}]}{[\text{S}]} = V_{max} \tag{4.22}$$

Equation (4.22) is the rate equation for a zero-order reaction, in which the velocity is independent of substrate concentration. This corresponds to the high-concentration end of the curve in Figure 4.17. The central part of the curve is described by the intermediate range of substrate concentrations. A special case in this range occurs when the substrate concentration equals K_m. At that point, Eq. (4.20) reduces to

$$v = \frac{V_{max}[\text{S}]}{[\text{S}] + [\text{S}]} = \frac{V_{max}}{2} \tag{4.23}$$

We can, therefore, define K_m as *the substrate concentration that yields one-half of the maximum velocity of the reaction.* You can verify that K_m *has units of concentration* by dimensional analysis of Eq. (4.20).

4.5.3D. Sample Calculations. We can perform several types of calculations with the Michaelis–Menten equation. Two examples follow.

(a) Given that $K_m = 1.00 \times 10^{-4}$ M and $V_{max} = 4.00$

$\times 10^{-2}$ mol liter^{-1} min^{-1}, calculate the rate at a substrate concentration of 1.00×10^{-2} M.

Solution:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$
$$= \frac{(4.00 \times 10^{-2}) \text{ mol liter}^{-1} \text{ min}^{-1} \times (1.00 \times 10^{-2})\, M}{(1.00 \times 10^{-4})\, M + (1.00 \times 10^{-2})\, M}$$
$$= 3.96 \times 10^{-2} \text{ mol liter}^{-1} \text{ min}^{-1}$$

(b) What is the K_m of an enzyme if, at [S] = 1.0×10^{-3} M, we obtain 25% of the maximum velocity of the reaction?

Solution:

$$v = \frac{V_{max}[S]}{K_m + [S]} = \frac{V_{max} \times (1.0 \times 10^{-3})\,\text{M}}{K_m + (1.0 \times 10^{-3})\, M} = 0.25 V_{max}$$

so that

$$V_{max}(1.0 \times 10^{-3}) = 0.25 V_{max}[K_m + (1.0 \times 10^{-3})]$$

and

$$K_m = 3.0 \times 10^{-3} M$$

4.5.3E. Lineweaver–Burk Transformation. The accuracy with which we can estimate K_m from the hyperbolic curve in Figure 4.17 depends on how clearly the curve levels off at high substrate concentrations. To get around this difficulty, scientists have proposed a number of linear transformations of the Michaelis–Menten equation. The most widely used one, proposed by Hans Lineweaver and Dean Burk, takes the following form:

$$\frac{1}{v} = \frac{K_m}{V_{max}} \frac{1}{[S]} + \frac{1}{V_{max}} \qquad (4.24)$$

A plot of $1/v$ versus $1/[S]$ yields a straight line having a slope of K_m/V_{max}, an intercept on the ordinate of $1/V_{max}$, and an intercept on the abscissa of $-1/K_m$ (Figure 4.18). The quantities K_m and V_{max} can be calculated from any two of these three parameters that determine a straight line. We term such a plot a *double-reciprocal plot.* Note that K_m, which has units of concentration, intercepts the $1/[S]$ axis, which has units involving concentration. Likewise, V_{max}, which has units of velocity, intercepts the $1/v$ axis, which has units involving velocity.

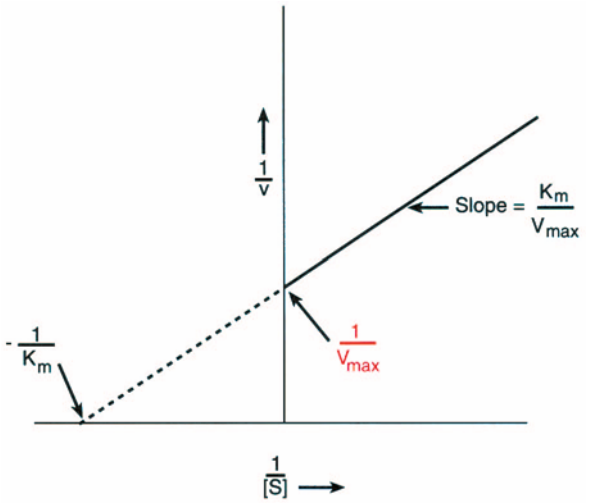


Figure 4.18. A Lineweaver–Burk plot of the Michaelis–Menten equation.

4.5.3F. Michaelis Constant (K_m). You have seen that the Michaelis constant, K_m, is a fundamental constant in enzymology. Let us summarize its key aspects:

- The K_m is the substrate concentration at which one-half of the maximum velocity of the reaction is observed. This is the most general definition of K_m and holds for all types of reaction mechanisms, even those not described by Eq. (4.7).
- The K_m has units of concentration. Small and large K_m values mean, respectively, that relatively low or high substrate concentrations are required to achieve one-half of the maximum velocity. K_m values of enzymes vary greatly and generally fall within the range of $10^{-8}M$ to $1.0M$.
- The K_m is a characteristic constant for a given enzyme but only under specified conditions of pH, temperature, ionic strength, and so on. As such, K_m is helpful in trying to determine whether enzymes isolated from two sources appear to be the same.
- The K_m is *not* an equilibrium constant, but rather a complex constant composed of various rate constants. For the mechanism described by Eq. (4.7),

$$K_m = (k_{-1} + k_2)/k_1$$

  For other reaction mechanisms, K_m is defined differently, involving other rate constants.
- The K_m is useful for comparing the activities of different enzymes and for comparing the "suitabilities" of alternate substrates for the same en-

zyme. A substrate with a lower K_m is a "better" substrate. The best substrate of an enzyme is one that provides the highest V_{max} and has the lowest K_m, that is, one yielding the highest V_{max}/K_m ratio.

4.5.4. Measures of Enzyme Activity

4.5.4A. Substrate Constant. We can evaluate the affinity of an enzyme for its substrate by considering the reverse of the first step in Eq. (4.7). We call the equilibrium constant for *dissociation* of the enzyme substrate complex a **substrate constant** and designate it K_s:

$$K_s = \frac{[E][S]}{[ES]} = \frac{k_{-1}}{k_1} \tag{4.25}$$

K_s represents an *inverse measure:* the greater the affinity of an enzyme is for its substrate, the greater the concentration of [ES] and the smaller the K_s.

4.5.4B. Catalytic Rate Constant. To assess an enzyme's efficiency, we use the rate constant of the rate-determining step, known as the **catalytic rate constant** (k_{cat}). In the above derivation, the catalytic rate constant is k_2. Hence, we can rewrite Eqs. (4.16) and (4.19), respectively, as

$$v = k_{cat}[ES] \tag{4.16a}$$

$$V_{max} = k_{cat}[E] \tag{4.19a}$$

It follows that $k_{cat} = V_{max}/[E]$ so that k_{cat} is identical to the turnover number of the enzyme. The turnover number measures enzyme efficiency when the enzyme is saturated with substrate; all of the active sites have substrate bound per unit time. However, most enzymes function *in vivo* under less than saturating conditions. For these situations, we need a different measure of efficiency. To derive it, recall that at low substrate concentrations, the Michaelis–Menten equation reduces to Eq. (4.21):

$$v = V_{max}[S]/K_m$$

Substituting into this equation from Eq. (4.19a) yields

$$v = (k_{cat}/K_m)[E][S] \tag{4.26}$$

4.5.4C. Specificity Constant. We refer to the ratio k_{cat}/K_m as the **specificity constant** of the enzyme. This constant amounts to an *apparent second-order rate constant* for the reaction between substrate and total enzyme when they combine to form the enzyme–substrate complex. This condition holds at low substrate concentrations when the concentration of free enzyme essentially equals that of total enzyme. Accordingly, the ratio k_{cat}/K_m measures what E and S can accomplish when abundant active sites are available for substrate binding. The larger this ratio is, the greater the enzyme's efficiency.

Specificity constants cannot be greater than the frequency with which enzymes and substrates collide. This *diffusion-controlled limit* falls within the range of 10^8–10^9 $M^{-1}\,s^{-1}$. You can see from the data in Table 4.6 that some enzymes have achieved virtual catalytic perfection; they catalyze a reaction essentially every time they collide with a substrate molecule.

4.5.5. Enzyme Inhibition

Reversible inhibitors function through a variety of mechanisms that we can sometimes distinguish by kinetic analysis. Three such groups of mechanisms are those of *competitive, noncompetitive,* and *uncompetitive inhibition* (Figure 4.19).

4.5.5A. Competitive Inhibitors. As the name implies, a **competitive inhibitor** competes with the

Table 4.6. Kinetic Parameters of Selected Enzymes

Enzyme	Substrate	K_m (M)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}\,s^{-1}$)
Chymotrypsin	*N*-Acetyl-Phe-Ala	1.5×10^{-2}	1.4×10^{-1}	9.3
Pepsin	Phe-Gly	3.0×10^{-4}	5.0×10^{-1}	1.7×10^{3}
Urease	Urea	2.5×10^{-2}	1.0×10^{4}	4.0×10^{5}
Carbonic anhydrase	HCO_3^-	2.6×10^{-2}	4.0×10^{5}	1.5×10^{7}
Fumarase	Fumarate	5.0×10^{-6}	8.0×10^{2}	1.6×10^{8}
Catalase	H_2O_2	2.5×10^{-2}	1.0×10^{7}	4.0×10^{8}

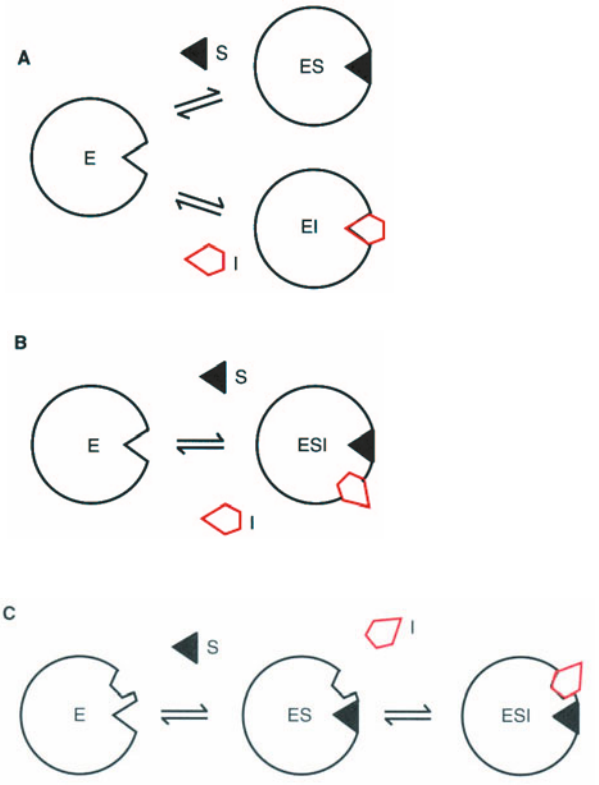


Figure 4.19. Types of reversible inhibitors. (A) A competitive inhibitor competes with the substrate for the same binding site. (B) A noncompetitive inhibitor binds to a site other than the active site. (C) An uncompetitive inhibitor binds only to the ES complex.

substrate for the same binding site (active site) on the enzyme. The inhibitor (I) usually resembles the substrate in its structure. When the substrate binds, the inhibitor cannot bind, and vice versa. Inhibitor binding is characterized by an equilibrium constant, called the **inhibitor constant** and designated K_i. Like the substrate constant, K_s, the inhibitor constant is a dissociation constant, specifically that of the enzyme–inhibitor complex:

$$K_i = \frac{[E][I]}{[EI]} \qquad (4.27)$$

A competitive inhibitor binds to the free enzyme but not to the enzyme–substrate complex. Since both the inhibitor and the substrate bind at the same active site, the effect of a competitive inhibitor can be overcome by simply increasing the substrate concentration, effectively displacing inhibitor molecules from the enzyme. Under these conditions, the enzyme achieves the full activity it would have had in the absence of inhibitor. The mechanism of competitive inhibition can be described as follows:

$$\begin{array}{ccccc} E + S & \rightleftharpoons & ES & \longrightarrow & E + P \\ + & & & & \\ I & & & & \\ K_i \updownarrow & & & & \\ EI & & & & \end{array}$$

We derive the rate equation for competitive inhibition in the same way as we derive the Michaelis–Menten equation, on the basis of appropriate assumptions. The derivation, based on a steady-state treatment for the ES complex, yields the following Lineweaver–Burk transformation:

$$\frac{1}{v} = \frac{K_m}{V_{max}}\left(\frac{1}{[S]}\right)\left(1 + \frac{[I]}{K_i}\right) + \frac{1}{V_{max}} \qquad (4.28)$$

Figure 4.20 shows a graphical plot of this equation. The slope of the line is greater than the slope of the line for the uninhibited reaction. However, the intercepts on the $1/v$ axis have the same values for both lines. This means that at high substrate concentrations ($1/[S] = 0$), one overcomes the inhibition and attains V_{max} of the uninhibited reaction. The intercept on the $1/[S]$ axis is smaller than that of the line for the normal system. The "apparent K_m" calculated from this intercept is larger than the "true K_m" of the enzyme. Specifically, "apparent K_m" = $(1 + [I]/K_i) \times$ "true K_m." The "apparent K_m" exceeds the

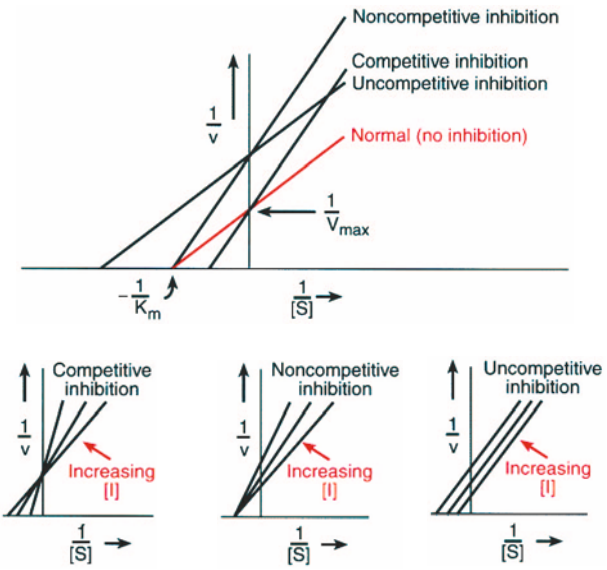


Figure 4.20. Lineweaver–Burk plots for different types of reversible enzyme inhibitors.

"true K_m" because fewer active sites are available for substrate binding.

4.5.5B. Antimetabolites. Competitive inhibitors are frequently **antimetabolites.** An antimetabolite is a synthetic compound that generally resembles a metabolite in its structure but has inhibitory effects. An antimetabolite inhibits a specific enzyme involved in utilization of the metabolite and may serve as a chemotherapeutic drug. *Sulfanilamide* (Figure 4.21) represents a classic antimetabolite. It belongs to the group of *sulfa drugs* which are lethal to a wide variety of pathogenic bacteria. Sulfanilamide resembles the metabolite *p-aminobenzoic acid (PABA)* in its structure.

PABA is a precursor of *folic acid.* Folic acid undergoes conversion to a number of different coenzyme forms, termed *folate coenzymes,* which are utilized by enzymes catalyzing reactions involving *one-carbon fragments* ($-CH_3$, HCHO, etc.). Sulfanilamide functions as a competitive inhibitor of *dihydropteroate synthetase,* an enzyme that catalyzes the incorporation of PABA in the pathway of folic acid synthesis (Figure 4.22). Sulfanilamide blocks the enzyme's active site and prevents folic acid synthesis. The resultant lack of folate coenzymes leads to impairment of one-carbon fragment metabolism and disruption of the entire network of metabolic reactions; the bacterium can neither grow nor divide, and it ultimately dies.

Because humans cannot synthesize folic acid, administration of sulfanilamide poses no danger at the doses that kill bacteria. Humans obtain folic acid through the diet and from *Escherichia coli* bacteria inhabiting the intestine. When sulfanilamide is administered, it is rapidly absorbed into the bloodstream so that it exerts its major effect on the infectious bacteria in the blood rather than on the useful organisms in the intestine.

4.5.5C. Noncompetitive Inhibitors. A **noncompetitive inhibitor** does not bind at the active site of the enzyme and usually bears no structural similarity to

p-Aminobenzoic acid (PABA, a metabolite) — Sulfanilamide (a sulfa drug)

Figure 4.21. Structural similarity of *p*-aminobenzoic acid (PABA) and sulfanilamide. Sulfanilamide (an antimetabolite) functions as a competitive inhibitor of PABA.

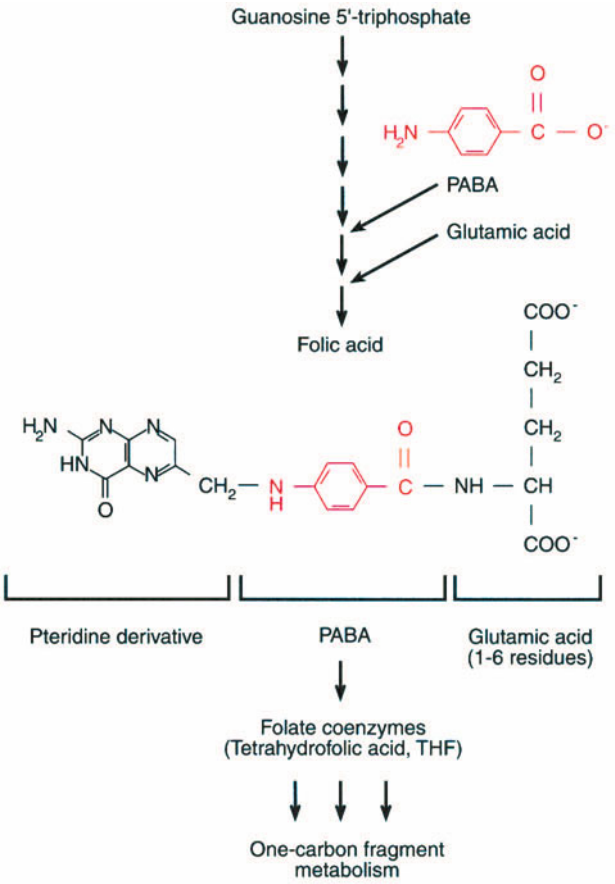


Figure 4.22. Schematic diagram depicting the synthesis and function of folate coenzymes.

the substrate. The inhibitor can bind to either the free enzyme or the enzyme–substrate complex, forming two types of inactive complexes, EI and ESI. In *pure noncompetitive inhibition,* the substrate and inhibitor bind independently; binding of one has no effect on binding of the other. Accordingly, binding of I to either E or ES is characterized by the same inhibitor constant, K_i. The mechanism of pure noncompetitive inhibition can be described as follows:

$$\begin{array}{ccc} E + S & \rightleftharpoons & ES \longrightarrow E + P \\ + & & + \\ I & & I \\ K_i \uparrow\downarrow & & \uparrow\downarrow K_i \\ EI & \rightleftharpoons & ESI \end{array}$$

Derivation of the rate equation for this type of inhibition differs from the derivations for uninhibited and competitively inhibited reactions in that we base it on the assumption of an *equilibrium between E, S, and ES.* In so

doing, we obtain a rate equation that in its Lineweaver–Burk transformation takes the form

$$\frac{1}{v} = \frac{K_m}{V_{max}}\left(\frac{1}{[S]}\right)\left(1 + \frac{[I]}{K_i}\right) + \frac{1}{V_{max}}\left(1 + \frac{[I]}{K_i}\right) \quad (4.29)$$

Figure 4.20 shows a graphical plot of this equation. Both the slope of the line and its intercept on the $1/v$ axis are greater than those for the uninhibited system. Hence, the maximum velocity attainable in the presence of a noncompetitive inhibitor is smaller than without the inhibitor; the inhibited enzyme has fewer functional catalytic sites. The maximum velocity attainable equals $V_{max}/(1 + [I]/K_i)$, where V_{max} represents the maximum velocity of the uninhibited reaction. The intercept on the $1/[S]$ axis has the same value as that for the normal reaction and yields the "true K_m" of the enzyme. Note that the inhibition by a noncompetitive inhibitor can never be completely overcome, not even at high substrate concentrations. Noncompetitive inhibitors have a significant impact on metabolism because they frequently function as *effectors* of *allosteric enzymes* (see Section 4.6).

4.5.5D. Uncompetitive Inhibitors. An **uncompetitive inhibitor** binds to the enzyme–substrate complex but does not bind to the free enzyme. It does not bind at the active site of the enzyme and usually bears no structural similarity to the substrate. In *pure uncompetitive inhibition,* the ESI complex cannot yield product, so the velocity of the reaction is proportional to the breakdown of the ES complex alone. The inhibitor constant is the dissociation constant of the ESI complex. The mechanism of pure uncompetitive inhibition can be described as follows:

$$\begin{array}{ccccc} E + S & \rightleftharpoons & ES & \longrightarrow & E + P \\ & & + & & \\ & & I & & \\ & & \Updownarrow K_i & & \\ & & ESI & & \end{array}$$

We derive the rate equation for uncompetitive inhibition in the same way as we derive the Michaelis–Menten equation. The derivation, based on a steady-state treatment of the ES complex, yields the following Lineweaver–Burk transformation:

$$\frac{1}{v} = \frac{K_m}{V_{max}}\left(\frac{1}{[S]}\right) + \frac{1}{V_{max}}\left(1 + \frac{[I]}{K_i}\right) \quad (4.30)$$

Figure 4.20 shows a graphical plot of this equation. The slope of the line is identical to the slope of the line for the uninhibited reaction, but both the $1/v$ intercept and the $1/[S]$ intercept are larger than those in the absence of inhibition. Both intercepts are increased by the factor $(1 + [I]/K_i)$. For the same inhibitor concentration, the maximum velocity attainable equals that for noncompetitive inhibition. Uncompetitive inhibition can never be completely overcome, not even at high substrate concentrations. The "apparent K_m" for uncompetitive inhibition is smaller than the "true K_m" of the enzyme. Specifically, "true K_m" = $(1 + [I]/K_i) \times$ "apparent K_m." Uncompetitive inhibition is rare for *single-substrate* reactions but common in *multisubstrate* reactions.

4.6. REGULATION OF ENZYME ACTIVITY III: ENZYME TYPES

In addition to regulation by properties and kinetics, enzyme activity is regulated by the occurrence of different "types" of enzymes. We include in this section enzymes kept in an inactive state that require specific activation before they can function as catalysts.

4.6.1. Inactive Forms

A number of mechanisms have evolved to maintain enzymes in an inactive state. Some enzymes occur in membrane-enclosed subcellular structures called *lysosomes.* Most lysosomal enzymes are hydrolytic enzymes that act on various types of biomolecules, particularly macromolecules. Encapsulation of these enzymes prevents their random hydrolysis of cytoplasmic substrates. Controlled release of enzymes from lysosomes helps regulate metabolism.

Other enzymes are prevented from being in a functional state at inappropriate times by a different mechanism. Proteolytic enzymes are prevented from digesting the very protein tissues that produced them by being synthesized in the form of larger, but inactive, precursors called **zymogens** or **proenzymes.** We name such inactive precursors using either an "ogen" suffix or a "pro" prefix:

Inactive		*Active*
trypsin*ogen*	→	trypsin
*pro*carboxypeptidase	→	carboxypeptidase

Converting a zymogen to an active enzyme requires hydrolysis of one or more peptide bonds. As an example, activating trypsinogen to trypsin requires removing a hexapeptide (Figure 4.23). The enzyme *enteropeptidase* catalyzes this activation, which occurs when trypsinogen enters the duodenum from the pancreas, where it is synthesized.

Enteropeptidase is a *serine protease,* a proteolytic enzyme having a serine residue at the active site, and its secretion from the mucosa in the duodenum is under hor-

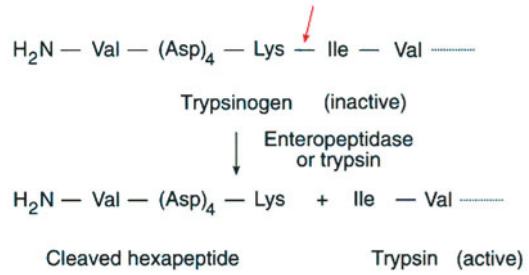


Figure 4.23. Trypsinogen activation. Hydrolytic removal of a hexapeptide is required for activation.

monal control. Because trypsinogen activation occurs at a trypsin-sensitive site (see Table 3.3), this reaction can also be catalyzed by trypsin itself. The small amount of trypsin formed by enteropeptidase acts on additional trypsinogen, thereby generating more active trypsin. As a result, trypsinogen activation is *autocatalytic;* trypsin serves as a catalyst for its own formation from trypsinogen.

A special group of serine proteases functions in blood clot formation. These enzymes undergo sequential activation, with each activated protease catalyzing the activation of a subsequent zymogen. The series of activation steps, called a **cascade mechanism** (Figure 4.24), provides tremendous *amplification* of the original signal. At each step, the number of active enzyme molecules increases greatly. Beginning with just a few active molecules, the cascade terminates with many, allowing for extensive catalysis. In the case of blood clotting, the capacity to form a clot efficiently is essential to prevent damage to the organism. Glycogen synthesis and breakdown, two processes that must be carried out efficiently, also feature control by means of cascade mechanisms (Figures 10.26 and 10.30).

4.6.2. Isozymes

Some enzymes occur in several forms that differ in their structures and properties but catalyze the same reaction. We

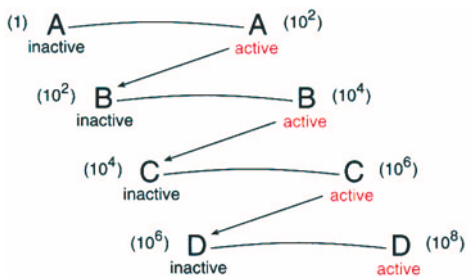


Figure 4.24. The principle of a cascade mechanism. If, at each step, one molecule catalyzes the activation of 100 molecules of the next compound, the short cascade from active A to active D amplifies a process by a factor of one million.

term such enzyme forms **isozymes** or **isoenzymes.** Their selective occurrence in specific tissues helps regulate metabolism. To illustrate the role of isozymes, consider *lactate dehydrogenase (LDH),* an oligomer composed of four subunits. Two types of subunits occur in living systems, and both have catalytic activity; each subunit has an active site for binding substrate. The enzyme occurs in five forms that differ in their quaternary structure:

M_4 (4 M subunits) M_3H M_2H_2 MH_3 H_4 (4 H subunits)

We designate the subunits as M and H because the M_4 isozyme predominates in skeletal muscles whereas the H_4 isozyme predominates in heart. LDH catalyzes the conversion of pyruvate to lactate, a key reaction in carbohydrate metabolism. The M_4 and H_4 isozymes differ in their catalytic properties as evidenced by their K_m values, the rate of pyruvate to lactate conversion, and the inhibition by pyruvate of the reverse reaction. Characteristics of overall carbohydrate metabolism in heart and skeletal muscles result in part from the predominance of a different LDH isozyme in each tissue.

4.6.3. Multienzyme Systems

Yet another "type" of enzyme occurs as part of a **multienzyme system.** A multienzyme system represents a structural and functional entity formed by association of several different enzymes, with each catalyzing a step in a reaction sequence. A given enzyme may occur as a single molecule or in multiple copies. In a linked system of this type, control of one enzyme automatically translates into control of the entire sequence. Multienzyme systems function at key steps in metabolism. Two examples are the *pyruvate dehydrogenase complex* and the *fatty acid synthase complex.*

The pyruvate dehydrogenase complex (Section 11.2) catalyzes five reactions and consists of three types of enzymes, designated E_1, E_2, and E_3. The mammalian complex contains 120 polypeptide chains of E_1, 60 chains of E_2, and 12 chains of E_3. It has a total molecular weight of 8.4×10^6. The same enzyme complex in *Escherichia coli* is smaller; it has a molecular weight of 4.6×10^6 and consists of 24 polypeptide chains of E_1, 24 chains of E_2, and 12 chains of E_3.

Fatty acid synthase (Section 13.5) catalyzes seven reactions and consists of seven enzymatic activities. The eukaryotic complex is a dimer of two large, identical polypeptide chains and has a total molecular weight of 520,000. Different parts of the chain have different enzymatic activities; the enzyme is *multifunctional.* The fatty acid synthase complex in yeast consists of 12 polypeptide chains of two types (A and B). The complex has the formula A_6B_6 and a total molecular weight of 2.4×10^6.

4.6.4. Regulatory Enzymes

Regulatory enzymes function at key locations in metabolism and are subject to special regulation, over and above that of enzymes in general. Control of these enzymes serves to regulate the vast network of interrelated metabolic reactions. Regulation of metabolism is analogous to controlling traffic in a large city during a power outage. It is clearly impossible to have police stationed at each and every intersection in the network of streets, avenues, and boulevards. Instead, officers regulate traffic only at major intersections, which automatically regulates the flow of traffic throughout the entire city. Regulatory enzymes control the network of metabolism in similar fashion. They frequently occur in one of two locations:

1. As the first enzyme in a series of linked reactions. A regulatory enzyme could be the first enzyme of a multienzyme system, a cyclic set of reactions, or some other set of linked reactions. The effect of this location becomes apparent when you consider a hypothetical sequence such as

$$A \xrightarrow{E_1} B \xrightarrow{E_2} C \xrightarrow{E_3} D \xrightarrow{E_4} F \xrightarrow{E_5} G \qquad (4.31)$$

By controlling the first step, catalyzed by enzyme E_1, you automatically control all of the other steps. If E_1 is activated, the concentrations of all products, B through G, will increase. If E_1 is inhibited, their concentrations will decrease. Controlling the first enzyme constitutes an efficient way to regulate a number of subsequent steps.

2. As the enzyme at a branch point in metabolism. To understand the effect of regulating an enzyme located at a branch point, consider a hypothetical pathway in which C has two metabolic fates; it can be converted to F or to M:

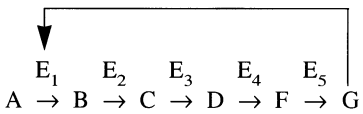
$$A \rightarrow B \rightarrow C \begin{cases} \nearrow D \rightarrow E \rightarrow F \\ \searrow K \rightarrow L \rightarrow M \end{cases} \qquad (4.32)$$

For such a system, it may be desirable to have *differential control* of both F and M at the same time. This can be achieved by regulating either the enzyme catalyzing C → D or that catalyzing C → K. If the enzyme catalyzing C → D is inhibited, then more of C will automatically be diverted to C → K and therefore toward production of M. Thus, one can *decrease F* and *increase M* at the same time. Likewise, one can *decrease M* and *increase F* by inhibiting the enzyme catalyzing C → K.

4.6.5. Feedback Mechanisms

Frequently, regulatory enzymes are controlled by means of a **feedback mechanism.** The term originated in engineering and refers to a process in which *part of the output of a system is returned to the input of the same system in such a fashion that it affects the subsequent output of the system.* Temperature regulation of a home furnace by a thermostat represents a familiar example. When sufficient heat has been generated (output of the system), the product (heat) turns off the source (furnace) of its own production. This system illustrates a negative feedback mechanism (inhibition).

Control of regulatory enzymes also frequently involves negative feedback mechanisms. To clarify how this works, assume that the reaction sequence of Eq. (4.31) is subject to feedback control. Typically, this means that the first enzyme in the series is inhibited by the last compound produced: enzyme E_1 is inhibited by compound G:



At the early stages of the reaction, the concentration of G is low, and its inhibitory effect minimal. However, as the concentration of G builds up, G's effect on E_1 becomes pronounced and leads to inhibition of E_1. G turns off its own source of production, much as the heat generated by the furnace turns off the furnace that produces heat. We call such a process **feedback inhibition** or **end-product inhibition.** It constitutes a reversible process that is typically noncompetitive in nature.

Assume next that the more complex pathway of Eq. (4.32) is also regulated via feedback (Figure 4.25). Here, a number of scenarios may be possible. In order to maintain a balance of F and M, high concentrations of F might inhibit the conversion of C to D. In that case, as F builds up, it would shut off its own production and automatically divert C toward the production of more M. In addition to inhibiting C → D, F might function as an activator of C → K. Alternatively, M could inhibit C → K and activate C → D. Lastly, a combination of both F and M might be required to inhibit A → B in order to achieve overall con-

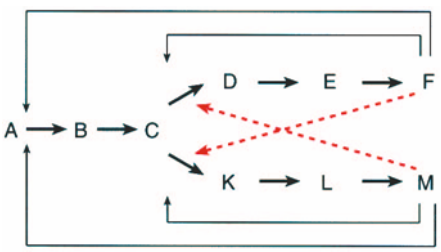


Figure 4.25. Complex feedback control. Dashed lines represent inhibition, and solid lines represent activation.

trol. In a real situation, all or only some of these controls may operate.

As just indicated, feedback mechanisms can also be positive, leading to an activation. Autocatalytic enzymes illustrate positive feedback. The initial conversion of inactive trypsinogen produces a small amount of active trypsin. As trypsin accumulates, it itself converts more trypsinogen to trypsin, resulting in greatly increased trypsinogen conversion.

4.6.6. Covalently Modified Enzymes

Regulatory enzymes are of two types: covalently modified enzymes and allosteric enzymes; some enzymes fit in both categories. A **covalently modified enzyme** is one whose catalytic activity can be changed by covalently altering its molecular structure. The alteration constitutes a chemical reaction distinct from the catalytic reaction of the regulatory enzyme and catalyzed by a second, different enzyme. Frequently, covalent modifications entail phosphorylation and dephosphorylation. In some cases, the phosphorylated form is active, and the dephosphorylated form inactive. In other cases, the reverse applies.

To illustrate covalent modification, consider *phosphorylase,* a key enzyme of carbohydrate metabolism. The enzyme, a dimer (Figure 4.26), catalyzes the breakdown of glycogen, the form in which carbohydrate is stored in animals. The dephosphorylated form (phosphorylase *b*) is inactive. Upon phosphorylation, a phosphate group becomes esterified to the hydroxyl group of serine 14 (i.e., the serine residue at position 14 in the amino acid sequence) in each subunit. The phosphorylation, catalyzed by *phosphorylase kinase,* converts phosphorylase *b* to active phosphorylase *a.* When glycogen breakdown becomes unnecessary, a different enzyme, *phosphoprotein phosphatase,* catalyzes the dephosphorylation of phosphorylase *a* to inactive phosphorylase *b.*

4.6.7. Allosteric Enzymes

Allosteric enzymes have binding sites that are distinct from the catalytic (active) sites for the binding of substrate. We call these additional sites **allosteric sites** (from the Greek *allos,* meaning "other," and *stereos,* meaning "space") or **regulatory sites.** The ligands that bind to these sites are known as **allosteric effectors,** or simply *effectors,* and can modulate the enzyme's catalytic activity. Some allosteric effectors are intermediates, coenzymes, or nucleotides that function in the pathway in which the allosteric enzyme catalyzes a specific step. Other effectors do not participate in the particular pathway and have different metabolic roles.

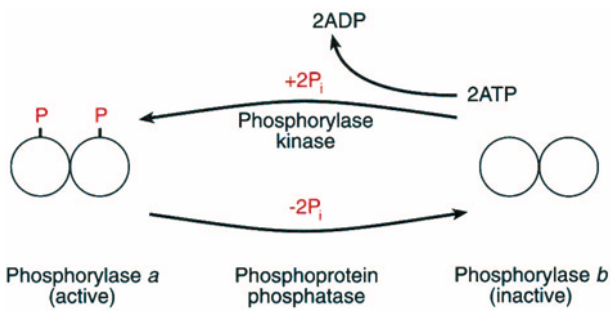


Figure 4.26. Phosphorylase, a covalently modified regulatory enzyme.

Allosteric enzymes are multisubunit enzymes with multiple binding sites. Upon binding substrates or effectors, these enzymes undergo conformational changes. Because of their subunit structure, they exhibit *cooperative effects* much like those shown by hemoglobin for the binding of oxygen (see Section 3.7).

Allosteric effectors can be positive or negative, depending on whether they increase (activate) or decrease (inhibit) the binding of substrate by the enzyme. Negative effectors typically function as noncompetitive inhibitors and frequently exert their effect via feedback mechanisms. Cooperative effects in allosteric enzymes and other proteins can involve identical ligands (*homotropic interactions*) or different ligands (*heterotropic interactions*). The effect on oxygen affinity of hemoglobin produced by O_2 binding is a homotropic effect; that produced by BPG binding is a heterotropic effect.

Scientists have proposed two major models to explain the action of allosteric enzymes. Both of these are equilibrium models in the sense that the interactions of allosteric effectors and enzymes can be described by various equilibria. One model, proposed by D. E. Koshland, G. Nemethy, and D. Filmer, is known as the *sequential model;* the other model, proposed by J. Monod, J. Wyman, and J.-P. Changeux, is known as the *concerted model.* We can explain the action of some enzymes best by the sequential model; other enzymes appear to function according to the concerted model.

4.6.7A. The Sequential Model. According to the **sequential model** (Figure 4.27), an allosteric enzyme consists of a number of subunits, each of which has a binding site for substrate. Binding of a substrate molecule to a subunit alters the conformation of that subunit; binding of the first substrate alters the first subunit, binding of the second substrate alters the second subunit, and so on (hence the term *sequential*). As each subunit undergoes a conformational change, the nature of the entire enzyme molecule is altered, and its affinity for binding substrate changes.

Binding of substrate to the enzyme is cooperative. The cooperativity can be positive or negative, depending on whether the binding of subsequent substrate ligands is

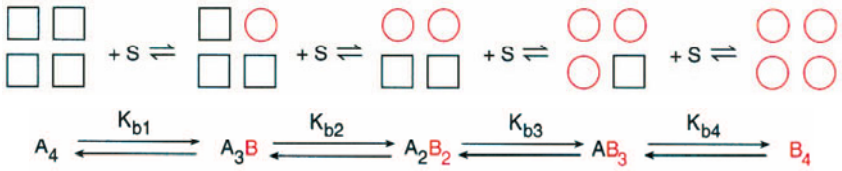


Figure 4.27. The sequential model of allosteric enzymes. S = substrate; A = □; B = ○.

enhanced or decreased. For an enzyme showing positive cooperativity, a plot of rate as a function of substrate concentration yields a *sigmoidal (S-shaped) curve* (Figure 4.28), analogous to the oxygen saturation curve of hemoglobin. Binding of allosteric effectors shifts the curve to the left for positive effectors and to the right for negative effectors. A positive effector lowers the K_m of the enzyme (the "apparent K_m" is smaller than the "true K_m") whereas a negative effector increases the K_m (the "apparent K_m" is larger than the "true K_m").

Each transition between two enzyme forms in Figure 4.27 is characterized by a complex equilibrium constant that has three components—one for the conformational change of the subunit, one for the binding of substrate to the subunit, and one for the interaction between the altered subunit and other subunits surrounding it. Mathematical treatment of these equilibria leads to an equation that describes the typical sigmoidal curve of cooperative interactions. In both physical and mathematical terms, substrate binding to multisubunit allosteric enzymes is analogous to oxygen binding to oligomeric hemoglobin.

4.6.7B. The Concerted Model. According to the **concerted model** (Figure 4.29), an allosteric enzyme exists as an equilibrium mixture of two conformational forms or states. The subunits, called *protomers,* differ in their affinity for substrate and effectors in the two states. In the *R* or *relaxed state,* the protomers bind substrate and activators but cannot bind inhibitors. In the *T* or *tensed state,* the protomers bind inhibitors but cannot bind substrate or activators.

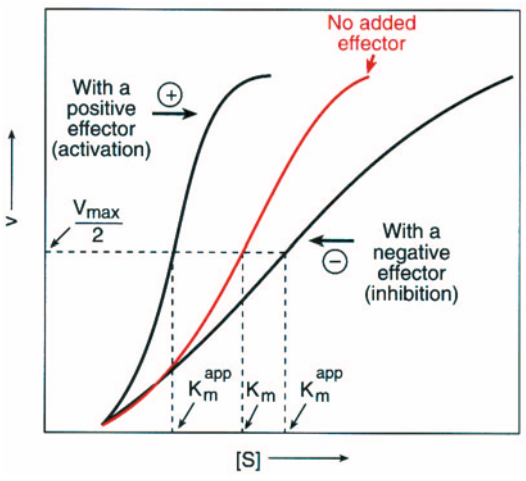


Figure 4.28. The effects of positive and negative effectors on cooperative substrate binding by allosteric enzymes.

When substrate adds to this equilibrium mixture, it binds only to the R-form. As it does, it effectively removes some substrate binding sites; in a sense, some of the R-form is "inactivated" since it cannot participate in further substrate binding. Because of the effective decrease in "active" R-form, some of the T-form must convert to the R-form. This establishes a new equilibrium mixture; the equilibrium of the R–T transition has shifted from right to left. The shift in equilibrium is analogous to that of any chemical reaction on removal of a product in the form of a precipitate or a gas.

Every time an enzyme molecule converts from the T- to the R-state, it generates *multiple substrate binding sites* (two in the example of Figure 4.29, four for a protein composed of four protomers, and so on). Hence the shift in equilibrium results in a large increase in the number of available substrate binding sites. Consequently, both substrate binding and reaction rate increase, and the system exhibits characteristics of cooperative interactions. As was true for the sequential model, mathematical treatment of these changes results in an equation that, when rate is plotted as a function of concentration, produces a sigmoidal curve.

Adding an activator, or positive effector, shifts the R–T equilibrium as if we were adding substrate. Conversely, adding an inhibitor, or negative effector, shifts the equilibrium in the opposite direction. Some allosteric enzymes have only substrate binding sites; the substrate and the allosteric effector are one and the same, and the cooperative interactions are *homotropic.* Other enzymes bind allosteric effectors in addition to binding substrate. They may have binding sites for one or more different activators and/or one or more different inhibitors; their cooperative interactions are *heterotropic.*

4.6.7C. Comparing the Models. An important difference between the two models is that the sequential model explains both positive and negative cooperativity, but the concerted model explains only positive cooperativity. In theory, sequential binding can result in subsequent binding that proceeds more readily (positive cooperativity) or less readily (negative cooperativity).

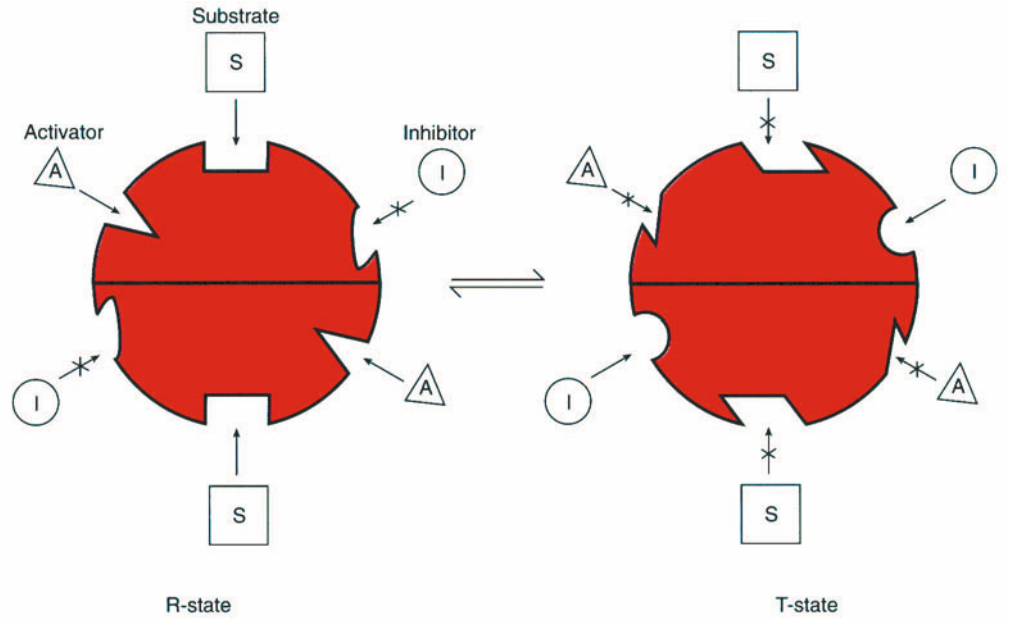


Figure 4.29. The concerted model of allosteric enzymes. S = substrate; A = activator; I = inhibitor.

Concerted binding, on the other hand, can lead only to more extensive binding. There is no way in which binding of substrate to an enzyme in the R-state, followed by conversion of molecules from the T- to the R-state, can result in *less* binding of substrate than before the conversion.

4.7. ENZYME FUNCTION—CHYMOTRYPSIN'S MECHANISM OF ACTION

A reaction mechanism describes how a chemical reaction might proceed based on experimental findings. The proposed mechanism may have to be modified or discarded at a later stage if new information comes to light that is not consistent with it. Based on the proposed mechanism, we can derive a rate equation, as we did for the Michaelis–Menten equation. The kinetics predicted by the rate equation must agree with experimental measurements. Agreement constitutes support for the proposed mechanism but does not prove that the mechanism is correct. In fact, we cannot prove that a particular mechanism constitutes the correct one, only that certain others cannot possibly apply.

The mechanism of action of *serine proteases* represents one of the best understood enzymatic mechanisms and has been studied thoroughly by chemical, kinetic, and physical techniques for well over 40 years. It appears to be firmly established. Serine proteases catalyze the hydrolysis of peptide bonds and have an exceptionally active serine residue at the active site. *Trypsin, chymotrypsin,* and *elastase* are the most extensively studied members of this family.

Elucidating an enzyme mechanism involves searching for answers to three basic questions:

1. Which amino acid residues are essential for enzymatic activity and are located at or near the active site?
2. What are the stereochemical relationships between the essential amino acid residues? What is the molecular architecture of the active site?
3. How do the essential amino acid residues function to cause the specific organic reaction? What is the actual mechanism of the reaction? How does the substrate bind to the site, and how is it converted to products?

Let us consider these three questions in turn by focusing on one member of the serine protease family—chymotrypsin.

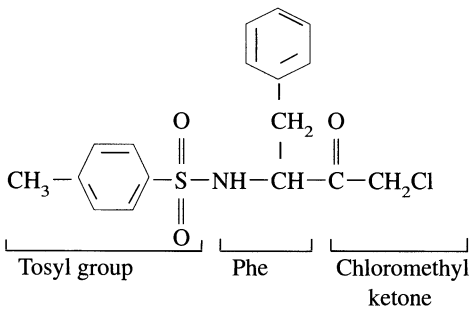
4.7.1. Amino Acids Essential for Activity

4.7.1A. Role of Serine. Treating chymotrypsin with *diisopropyl fluorophosphate (DFP;* see Figure 4.16) leads to complete inactivation of the enzyme. Since the reagent reacts with the hydroxyl group of serine, inactivation suggests the presence of serine at the active site. Additional work shows that DFP reacts with only one of the 27 serine residues in chymotrypsin, under conditions in which all the other serine residues are inactive. Hy-

drolysis of the treated enzyme, followed by analysis of the fragments, reveals that the reagent reacts specifically with residue 195. We conclude that Ser 195 appears to be located at the active site and to be exceptionally reactive.

4.7.1B. Role of Histidine. The importance of a second amino acid residue for chymotrypsin's catalytic activity can be demonstrated by **affinity labeling.** In this technique we react an enzyme with a reagent that resembles the substrate in its structure and possesses a reactive group. The reagent serves as a competitive inhibitor. We first allow the reagent to bind noncovalently to the enzyme's active site. Next, the reagent's reactive group is reacted with an amino acid residue at or near the active site. As a result of this reaction, the reagent becomes covalently linked to the enzyme at the active site and inhibits the enzyme completely.

In the case of chymotrypsin, affinity labeling was carried out with a reagent called *TPCK (tosyl-L-phenylalanyl chloromethyl ketone):*



TPCK resembles phenylalanine, one of the preferred amino acids in substrates of chymotrypsin (see Table 3.3). Additionally, TPCK contains the chloromethyl ketone group, a strong alkylating agent capable of reacting with the imidazole group of histidine.

When TPCK reacts with chymotrypsin, the enzyme is completely inactivated, implicating histidine as a second essential amino acid at the active site. Support for this conclusion comes from the fact that TPCK does not react with chymotrypsin in which the active site has been blocked by DFP. Hydrolysis of TPCK-chymotrypsin, followed by analysis of the fragments, reveals that the reagent has reacted specifically with His 57.

4.7.2. Molecular Architecture of the Active Site

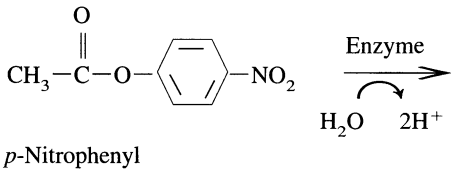
Results of X-ray diffraction studies of chymotrypsin agree with the tentative assignment of Ser 195 and His 57 to the active-site region. These two amino acids occur in close proximity in the three-dimensional structure of the molecule.

The X-ray diffraction pattern provides two additional clues to the reaction mechanism. First, an aspartic acid residue (102) is located close to Ser 195. Scientists believe that the aspartic acid residue helps to stabilize the intermediate in the reaction. Second, all three of the critical amino acids line a *hydrophobic pocket* in the enzyme molecule (Figure 4.30).

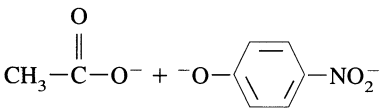
Recall from Table 3.3 that substrates of chymotrypsin contain the large hydrophobic groups of phenylalanine, tryptophan, or tyrosine. The X-ray pattern suggests that these groups fit into the hydrophobic pocket and thereby orient the substrate correctly on the enzyme. It is interesting to note that the pockets of trypsin and elastase differ from that of chymotrypsin. The altered pockets allow these enzymes to bind other substrates, according to their specificities.

4.7.3. Proposed Mechanism

4.7.3A. Two-Stage Reaction. We can often learn a great deal about a mechanism by studying an enzyme reaction, not with the natural substrate, but with related synthetic *model compounds.* In the case of chymotrypsin, esters can serve this purpose. Chymotrypsin possesses esterase activity in addition to its proteolytic function; the mechanisms of ester and amide hydrolysis are almost identical. One ester acted on by chymotrypsin is *p*-nitrophenyl acetate. Chymotrypsin catalyzes its hydrolysis to *p*-nitrophenol and acetate:



p-Nitrophenyl acetate



Acetate *p*-Nitrophenolate

Kinetic measurements of this reaction (Figure 4.31) indicate that it occurs in two stages: (a) a rapid "burst phase" in which *p*-nitrophenol forms at a fast rate and in stoichiometric quantities relative to the amount of enzyme present; and (b) a slow "steady - state phase" in which *p*-nitrophenol forms at a reduced but constant rate, independent of substrate concentration.

We can interpret these results in terms of a two-stage reaction sequence in which (a) the enzyme reacts rapidly with *p*-nitrophenol to form a *covalent acyl-enzyme com-*

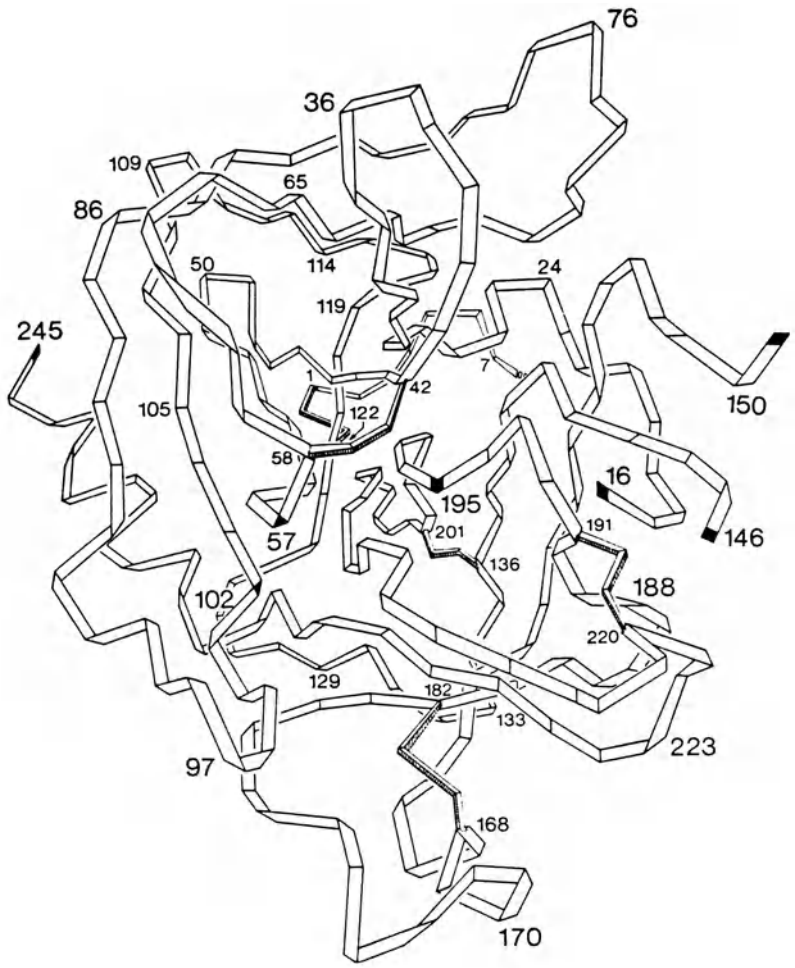


Figure 4.30. Main chain skeleton of the chymotrypsin molecule. Residues 57, 102, and 195 line a hydrophobic pocket. [Reprinted, with permission, from J. J. Birktoft and D. M. Blow, *J. Mol. Biol.* 68:187–240 (1972).]

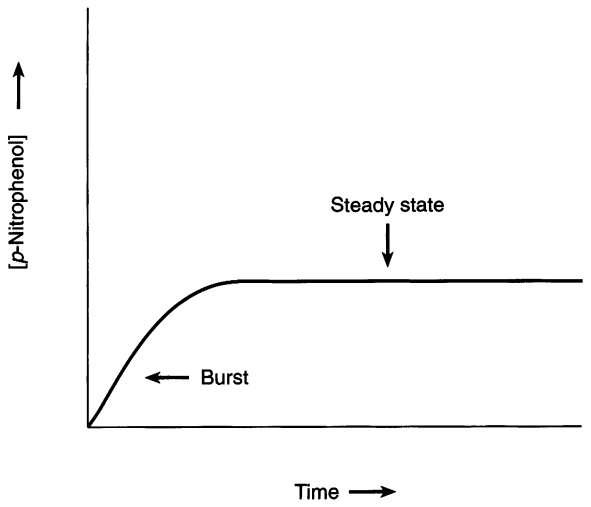
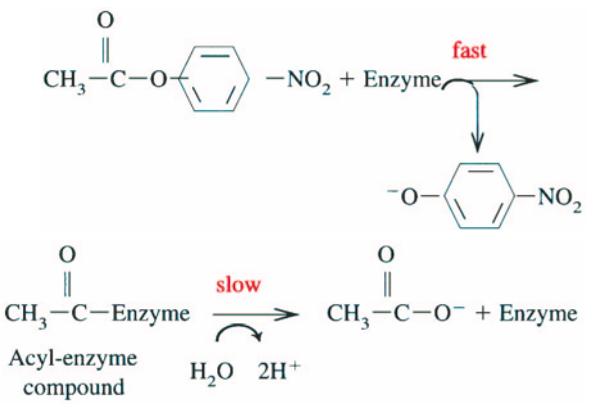


Figure 4.31. Formation of *p*-nitrophenol as a function of time in chymotrypsin-catalyzed hydrolysis of *p*-nitrophenyl acetate. A rapid initial rate is followed by a slower, steady- state rate.

pound and (b) the acyl-enzyme compound slowly hydrolyzes to release acetate:



The currently accepted mechanism for the proteolytic action of chymotrypsin postulates a similar two-stage process.

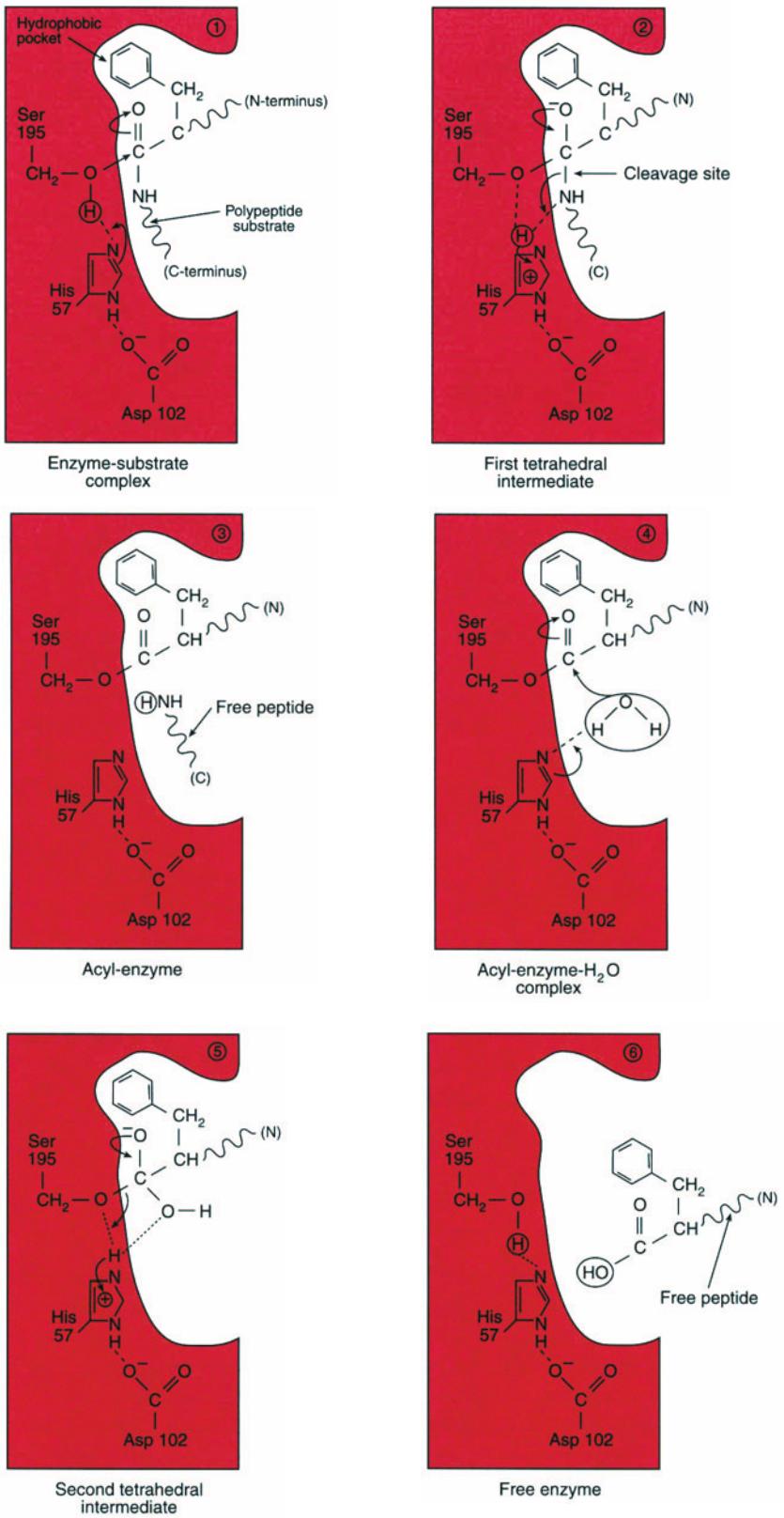
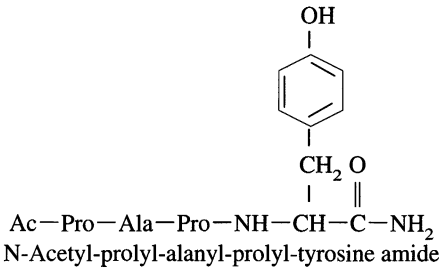


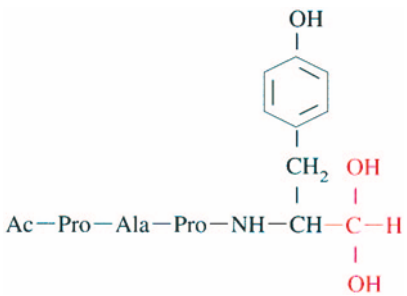
Figure 4.32. Chymotrypsin's mechanism of action. (1) Nucleophilic attack by serine's oxygen; (2) formation and stabilization of the first tetrahedral intermediate; (3) the first enzyme product forms; (4) nucleophilic attack by water's oxygen; (5) formation and stabilization of the second tetrahedral intermediate; (6) the second enzyme product forms.

4.7.3B. Tetrahedral Intermediate. Studies with inhibitors of chymotrypsin suggest the involvement of a *tetrahedral intermediate,* in which the four atoms bonded to the carbonyl carbon of the susceptible peptide bond are arranged as in a tetrahedron. Support for this concept comes from the chymotrypsin-catalyzed hydrolysis of



N-Acetyl-prolyl-alanyl-prolyl-tyrosine amide

to ammonia and *N*-acetyl-prolyl-alanyl-prolyl-tyrosine. The aldehyde form of the amide substrate, *N*-acetyl-prolyl-alanyl-prolyl-tyrosine aldehyde, inhibits this hydrolysis. Scientists believe that the aldehyde exists in aqueous solution predominantly as a hydrated molecule in which the aldehyde group has a tetrahedral structure



and that it functions as a competitive inhibitor for the tetrahedral intermediate of the transition state formed in the catalytic reaction.

4.7.3C. Postulated Mechanism. Based on the above information, we can propose the following mechanism to describe the action of chymotrypsin (Figure 4.32).

- The peptide is positioned at the active site; bulky nonpolar side chains of phenylalanine, tryptophan, or tyrosine fit into a hydrophobic pocket of the enzyme and are held in place by noncovalent hydrophobic interactions.
- The first stage of the reaction, the *acylation stage,* involves a nucleophilic attack by the oxygen of the hydroxyl group of Ser 195 on the carbonyl carbon of the susceptible peptide bond.
- A tetrahedral intermediate is formed and stabilized by Asp 102 and His 57. The tetrahedral intermediate leads to formation of an *acyl-enzyme.*
- The second stage of the reaction, the *deacylation stage,* is a reversal of the previous steps, but water hydrogen-bonds to His 57 and now acts as the attacking nucleophile.
- A tetrahedral intermediate forms as before. The tetrahedral intermediate regenerates the original, deacylated enzyme and forms a COOH group at the carbonyl carbon of the peptide bond.

This mechanism illustrates three factors believed to contribute to general catalysis (Section 4.3): *proximity and orientation, acid–base catalysis,* and *covalent catalysis.* The H-bonding of Ser 195 to His 57 increases the nucleophilicity of the Ser oxygen. The H-bond between Asp and His helps to aim the imidazole group at the serine hydroxyl. We refer to the H-bonded constellation of Asp, His, and Ser as the *catalytic triad.*

SUMMARY

Three unique properties—specificity, efficiency, and regulation—set enzymes apart from other catalysts. Enzymes are specific in the types of reactions they catalyze and in the types of compounds serving as their substrates. On the basis of the reactions catalyzed, we classify enzymes into six major classes, each divided into sub- and sub-subclasses.

Substrate binds to an active site on the enzyme, forming an enzyme–substrate complex. The active site may be preformed or induced by interaction with the substrate. Interaction of a prochiral substrate and its enzyme involves a minimum of three points of contact at the active site. Enzymes lower a reaction's energy of activation, allowing the activated complex to form more readily, and increase the reaction's rate.

Enzymes are regulated by controlling either the number of enzyme

molecules per cell or their level of activity. Regulation of enzyme activity is multifaceted. Reaction conditions of concentration, time, pH, and temperature all affect enzymatic activity. Inhibitors decrease enzymatic activity and act either reversibly or irreversibly. Reversible inhibitors may be competitive, noncompetitive, or uncompetitive. Noncompetitive inhibitors frequently function via a feedback mechanism.

The Michaelis–Menten equation constitutes a basic equation of enzyme kinetics. The Briggs–Haldane derivation of this equation assumes a steady state for the enzyme–substrate complex. The Michaelis constant (K_m), a fundamental constant in enzymology, equals the substrate concentration yielding one-half of the maximum velocity of the reaction.

Enzyme activity is also regulated by the diversity of enzyme "types." Enzymes occur as isozymes, inactive forms, and multienzyme systems. Regulatory enzymes function at key locations in metabolism and are subject to fine-tuned control. They occur in two forms—covalently modified and allosteric enzymes.

SELECTED READINGS

Carreras, W., and Santi, D. V., The catalytic mechanism and structure of thymidylate synthase, *Annu. Rev. Biochem.* 64:721–762 (1995).

Carter, P., and Wells, J. A., Dissecting the catalytic triad of a serine protease, *Nature (London)* 332:564–568 (1988).

Fersht, A. R., and Gani, D. (eds.), *Enzymic Catalysis,* Cambridge University Press, Cambridge (1991).

Kim, Y., Eom, S. H., Wang, J., Lee, D. S., Suh, S. W., and Steltz, T. A., Crystal structure of *Thermus aquaticus* DNA polymerase, *Nature (London)* 376:612–616 (1995).

Koleske, A. J., and Young, R. A., An RNA polymerase II holoenzyme responsive to activators, *Nature (London)* 368:466–469 (1994).

Koshland, D. E., Correlation of structure and function of enzyme action, *Science* 142:1533–1541 (1963).

Liu, Y., Van Heeswijck, R., Høj, P., and Hoogenraad, N., Purification and characterization of ornithine acetyltransferase from *Saccharomyces cerevisiae, Eur. J. Biochem.* 228:291–296 (1995).

Lohse, P. A., and Szostak, J. W., Ribozyme-catalyzed amino acid transfer reactions, *Nature (London)* 381:442–444 (1996).

Matthews, J. C., *Fundamentals of Receptor, Enzyme, and Transport Kinetics,* CRC Press, Boca Raton, Florida (1993).

Monod, J., Changeux, J.-P., and Jacob, F., Allosteric proteins and cellular control systems, *J. Mol. Biol.* 6:306–329 (1963).

Pinot, F., *et al.,* Molecular and biochemical evidence for the involvement of the Asp-333–His-523 pair in the catalytic mechanism of soluble epoxide hydrolase, *J. Biol. Chem.* 270:7968–7987 (1995).

Xue, Q., and Yeung, E. S., Differences in the chemical reactivity of individual molecules of an enzyme, *Nature (London)* 373:681–683 (1995).

REVIEW QUESTIONS

A. Define each of the following terms:

Initial velocity	Cascade mechanism
Rate-determining step	Polyaffinity theory
Affinity labeling	Enzyme specificity
Prochiral carbon	Catalytic rate constant

B. Differentiate between the two terms in each of the following pairs:

Steady state/transition state	Endergonic/exergonic
Specific activity/total activity	Cofactor/antimetabolite
Lock-and-key theory/induced fit	Activated complex/enzyme–substrate complex
Rate constant/rate equation	Zymogen/isozyme
Reversible inhibitor/irreversible inhibitor	Michaelis constant/substrate constant

C. (1) How could you determine whether compound X is a reversible or an irreversible inhibitor of a given enzyme? Assuming that X acts as a reversible inhibitor, how could you determine whether it functions in a competitive or a noncompetitive fashion?

(2) What is meant by each of the following? (i) Specificity constant; (ii) inhibitor constant; (iii) substrate constant; (iv) Michaelis constant.

(3) What properties do enzymes share with all other catalysts and what properties set enzymes apart?

(4) Write balanced equations for the reactions catalyzed by urease, alcohol dehydrogenase, catalase, and chymotrypsin (reaction with p-nitrophenyl acetate).

(5) Compare and contrast the sequential and concerted models for allosteric enzymes.

(6) How are the following involved in the regulation of enzyme activity? (i) Multienzyme systems; (ii) lysosomes; (iii) feedback mechanisms; (iv) covalently modified enzymes.

PROBLEMS

4.1. Based on Figures 4.8 and 4.28 (no added effector), estimate the factor by which the substrate concentration has to be increased in order to raise the rate of an ordinary and an allosteric enzyme reaction from 10% of V_{max} to 90% of V_{max}.

4.2. In order for a certain enzyme to show activity, an imidazole group of histidine ($pK_a' = 6.00$) at the active site must be in the protonated form so that it can interact electrostatically with a negatively charged γ-carboxyl of glutamic acid ($pK_a' = 4.25$) in the substrate. Considering only this interaction, what would you expect the optimum pH of the enzymatic reaction to be? Why?

4.3.* In an enzyme purification, fraction III contains a total of 50 mg of protein and has a specific activity of 1.5 units/mg of protein. Fraction IV has a volume of 20 ml, a protein concentration of 1.0 mg/ml, and a specific activity of 3.0 units/mg of protein. What is (a) the percent yield and (b) the degree of purification in going from fraction III to fraction IV?

4.4. An enzyme reaction has a rate of 14.7 and 20.8 mmol min^{-1} at 37.0°C and 45.0°C, respectively. What is the temperature coefficient of the reaction?

4.5.* By what factor does the rate constant of a reaction increase when the reaction's energy of activation decreases from 60.0 kJ/mol to 50.0 kJ/mol? (Hint: Use the Arrhenius equation in its exponential form: $k = Ae^{-E_a/RT}$. $R = 8.31$ J deg^{-1} mol^{-1}; $T = 298.2$ K.)

4.6. Express the total enzyme concentration for (a) competitive, (b) noncompetitive, and (c) uncompetitive inhibition by means of equations analogous to Eq. (4.12).

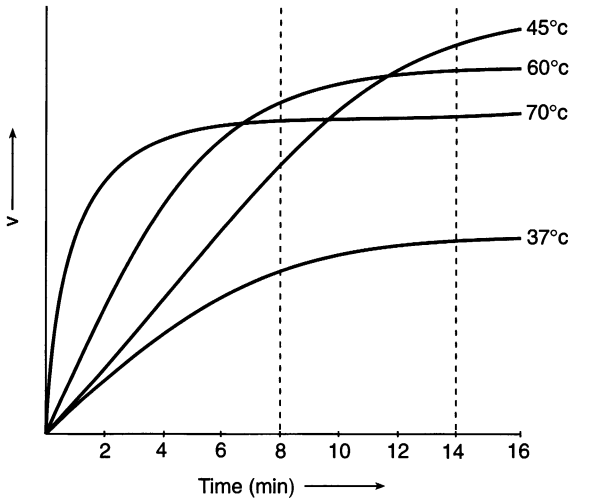
4.7. For enzyme A, $K_m = 3.0 \times 10^{-1} M$ and $k_{cat} = 4.2 \times 10^4$ s^{-1}; for enzyme B, $K_m = 2.1 \times 10^{-3} M$ and $k_{cat} = 8.4 \times 10^3$ s^{-1}. Which enzyme is the more effective catalyst?

4.8. What must be the ratio of K_m/[S] in order to attain a reaction rate equal to two-thirds of V_{max}?

4.9. The Michaelis constant of a dipeptidase is $2.8 \times 10^{-4} M$ for the dipeptide Ala-Gly and $3.5 \times 10^{-2} M$ for the dipeptide Leu-Ser. Which dipeptide is the better substrate of the enzyme? Two noncompetitive inhibitors of the enzyme have K_i values of 5.7×10^{-2} and 2.6×10^{-4}, respectively. Which of the two is the stronger inhibitor?

4.10. Write an equation for the steady-state rate of change of the ES complex for an uninhibited reaction in which the reverse reaction, from products to the ES complex, *cannot be ignored*.

4.11. Measuring the rate of an enzymatic reaction as a function of temperature yielded the set of curves shown below. What optimum temperature would you assign to the enzyme if you measured the rate (a) after 14 minutes, (b) after 8 minutes, or (c) at the origin (initial velocity)?



4.12.* For allosteric enzymes, *low* concentrations of competitive inhibitors frequently lead to an *increase* in the rate of the reaction. Why?

4.13. An enzymatic reaction yields *one-half* of the maximum velocity with a substrate concentration of $0.030M$. What substrate concentration is required in order to achieve *one-quarter* of the maximum velocity?

4.14. For a given enzyme reaction, the rate at a substrate concentration of $1.0 \times 10^{-2} M$ is four times that at a substrate concentration of $1.0 \times 10^{-3} M$. What is the K_m of the enzyme?

4.15. Two exergonic reactions have overall free energy changes of −21.8 kJ/mol and −40.6 kJ/mol, but the latter reaction proceeds at a slower rate than the former. How can you explain this?

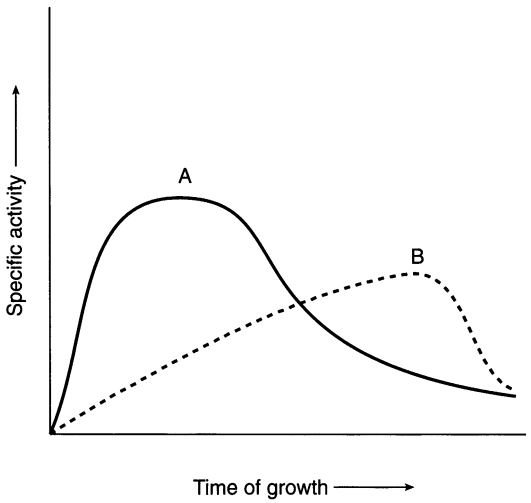
4.16. A pure enzyme preparation loses activity when dialyzed against water. Some investigators suggested that dialysis removes an essential cofactor; others proposed that the enzyme unfolds at low ionic strength. What experiments could you perform in order to decide between these two possibilities?

4.17. Fifty milligrams of an enzyme can be completely inhibited by adding *either* 2.0 micromoles of Ag^+ or 4.0 micromoles of Pb^{2+}. What is the enzyme's minimum molecular weight?

4.18.* What is the K_i of a noncompetitive inhibitor if a $3.0 \times 10^{-4} M$ concentration of inhibitor leads to 85% inhibition of the enzyme-catalyzed reaction?

4.19. Bacterial synthesis of compound C proceeds via a number of enzymatic steps. The pathway requires glycine in the first step and involves compounds A and B in later steps. An experimenter grew the bac-

teria in the presence of radioactive glycine and determined the specific activities (radioactive counts per milligram of compound). Plotting the data yielded the curves shown below. Which of the two compounds is a precursor of the other? When the experiment was repeated in the presence of added C, the incorporation of radioactive glycine into both A and B was drastically reduced. What can you conclude from this?



4.20.* Another linear transformation of the Michaelis–Menten equation results in the expression

$$v/[S] = -(1/K_m)v + V_{max}/K_m$$

so that a plot of v/[S] as a function of v yields a straight line. (This is known as an *Eadie–Scatchard plot.*) What are the intercepts on the ordinate and abscissa in such a plot?

4.21.* What must be the concentration of an *uncompetitive* inhibitor ($K_i = 1.0 \times 10^{-2}$ *M*) if, using Lineweaver-Burk plots, the intercept on the ordinate (1/v axis) produced by this inhibitor is identical to that produced by a $1.0 \times 10^{-4}M$ concentration of a *noncompetitive* inhibitor ($K_i = 1.0 \times 10^{-5}M$) of the same enzyme?

4.22. An enzyme loses 10% of its activity when heated for five minutes at 45°C. Repeating the same experiment in the presence of a large concentration of substrate leads to a loss in activity of only 3%. How do you explain these results?

4.23.* Methanol (wood alcohol) is highly toxic because of its oxidation to formaldehyde by liver alcohol dehydrogenase:

$$NAD^+ + \text{methanol} \rightarrow \text{formaldehyde} + NADH + H^+$$

Part of the medical treatment for methanol poisoning involves administering large doses of ethanol that would cause intoxication under ordinary conditions. Explain why this treatment is effective.

4.24.* Formation of the enzyme–substrate complex constitutes a chemical reaction involving noncovalent interactions and proceeding from reactants (E, S) to product (ES). Likewise, conversion of the ES complex to the EP complex, prior to release of the product, constitutes a second chemical reaction. Since each chemical reaction requires formation of an activated complex as an intermediate, draw a possible curve for an exergonic reaction, like that in Figure 4.4, but reflecting more accurately the true mechanism of the reaction.

4.25. The reaction

$$\text{pyruvate} + NADH + H^+ \rightarrow \text{lactate} + NAD^+$$

is catalyzed by lactate dehydrogenase and can be followed by measuring the decrease in absorbance at 340 nm (since NADH absorbs at this wavelength, but NAD^+ does not). After five minutes of incubation at 37°C, a reaction mixture showed a decrease in absorbance of 0.40. In the presence of an inhibitor, the same reaction mixture led to an absorbance decrease of only 0.10. Calculate the percent inhibition in the second reaction mixture.

4.26.* Amylase, present in saliva, catalyzes the degradation of starch and can be assayed by measuring the time required to reach the *achromic point.* We define the achromic point as that stage in starch hydrolysis at which addition of iodine fails to produce a blue color (iodine forms a blue complex with intact starch). In an experiment, 2.0 ml of 1:10 diluted saliva (1.0 ml enzyme + 9.0 ml water) yielded an achromic point of 6.0 min. If we define an enzyme unit as that amount of enzyme yielding an achromic point of 4.0 min, calculate the number of enzyme units in 1.0 ml of the original, undiluted saliva.

4.27. A Lineweaver–Burk plot of 1/v versus 1/[S] has a slope of 1.2×10^3 min and an intercept on the 1/v axis of 2.0×10^{-2} $nmol^{-1}$ ml min. Calculate V_{max} and K_m.

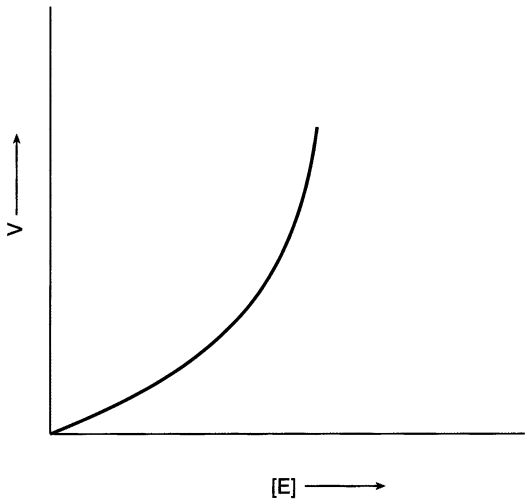
4.28.* An enzyme assay yields the following data:

[S] (mM)	v (no inhibitor) (mmol ml^{-1} min^{-1})	v (inhibitor present) (mmol ml^{-1} min^{-1})
1.50	0.167	0.115
2.00	0.204	0.143
2.50	0.232	0.167
5.00	0.313	0.250
10.00	0.385	0.333

Calculate K_m and V_{max} for the uninhibited and inhibited reactions. Is the inhibitor competitive, noncompetitve, or uncompetitive?

4.29. Reacting an enzyme with either diisopropyl fluorophosphate (DFP) or iodoacetamide (IAA) produces complete inactivation. After having been treated with either DFP or IAA, the enzyme does not react with the second inhibitor. What conclusions can you draw from these results?

4.30.* A student determines the variation of reaction rate as a function of enzyme concentration for a trypsin preparation. Unexpectedly, the results do not yield a straight line as in Figure 4.9 but rather a plot having an upward curvature as shown. Why does the curve have such a shape?



4.31.* The protein of muscle tissue in a normal human adult is in a steady state so that its rate of synthesis equals its rate of breakdown. On this basis, predict the change of muscle protein labeling as a function of time if the adult were given a single dose of radioactively labeled amino acids.

4.32.* A proteolytic enzyme is isolated from liver. The cell-free extract (300 ml) contains 150 mg of protein and has a total activity of 360 enzyme units. After a series of purification steps, the final preparation (4.0 ml) has a protein content of 0.080 mg of protein and a total activity of 288 enzyme units. What are the percent yield and the degree of purification achieved?

Carbohydrates

5

Carbohydrates, or **saccharides** (from the Greek, meaning "sugar"), are widespread in nature, especially in plants. More commonly known as *sugars,* they constitute, on the basis of mass, the most abundant class of biomolecules. More than half of all the organic carbon on Earth occurs in the form of two carbohydrates, starch and cellulose.

As implied by the name proposed about 100 years ago, **carbohydrates** were thought to be *hydrates of carbon,* having the general formula $C_n(H_2O)_n$. However, this formula correctly describes only simple sugars. Other carbohydrates have more complex compositions. The first carbohydrate structures elucidated were those of several simple sugars. Emil Fischer carried out these pioneering studies toward the end of the 19th century, launching what many consider to be the birth of organic chemistry.

Carbohydrates have a number of biological functions. In most organisms, carbohydrates, mainly in the form of the simple sugar glucose, constitute the *primary source of energy for metabolism* in general and for biosynthetic processes in particular. Carbohydrates form the basis of the energy cycle on Earth. They are synthesized from carbon dioxide and water in photosynthesis, thereby harnessing some of the solar energy striking the Earth. Subsequently, they are degraded in metabolism back to carbon dioxide and water, releasing their stored energy in the process. Complete degradation of carbohydrates yields approximately 17 kJ/g (4 kcal/g).

Carbohydrates also represent a *storage form of energy* in both animals and plants. While carbohydrates in solution provide readily available energy, their deposits constitute a stored form of energy. Starch and glycogen are the major stored carbohydrates in plants and animals, respectively. Large amounts of starch occur in corn, potatoes, rice, and other plant tissues. Animals store only small amounts of glycogen, primarily in muscle and the liver. Most of the stored energy of animals takes the form of lipid deposits.

In most organisms, carbohydrates serve as precursors, providing a *source of carbon atoms for biosynthesis.* Glucose and other simple sugars

provide carbon skeletons for the biosynthesis of proteins, lipids, nucleic acids, and complex polysaccharides.

Carbohydrates serve as *structural components.* Cellulose constitutes the main structural component of cell walls and fibrous tissues of plants. Peptidoglycan, a cross-linked polymer of carbohydrates and peptides, forms the major structural element of bacterial cell walls. Complex polysaccharides occur in the cell coat, a slimy layer covering most eukaryotic cell membranes.

Lastly, carbohydrates have *informational roles.* Their covalent attachment to proteins, forming glycoproteins, provides proteins with important recognition factors. These may aid in targeting proteins (dispatching them to specific intra- or extracellular locations), in orienting them in biological membranes, and in having them perform their varied functions. Cell-coat carbohydrates provide for cell specificity and function in cell–cell interactions. Surface carbohydrates of red blood cells help determine the foreign versus self aspects of living organisms.

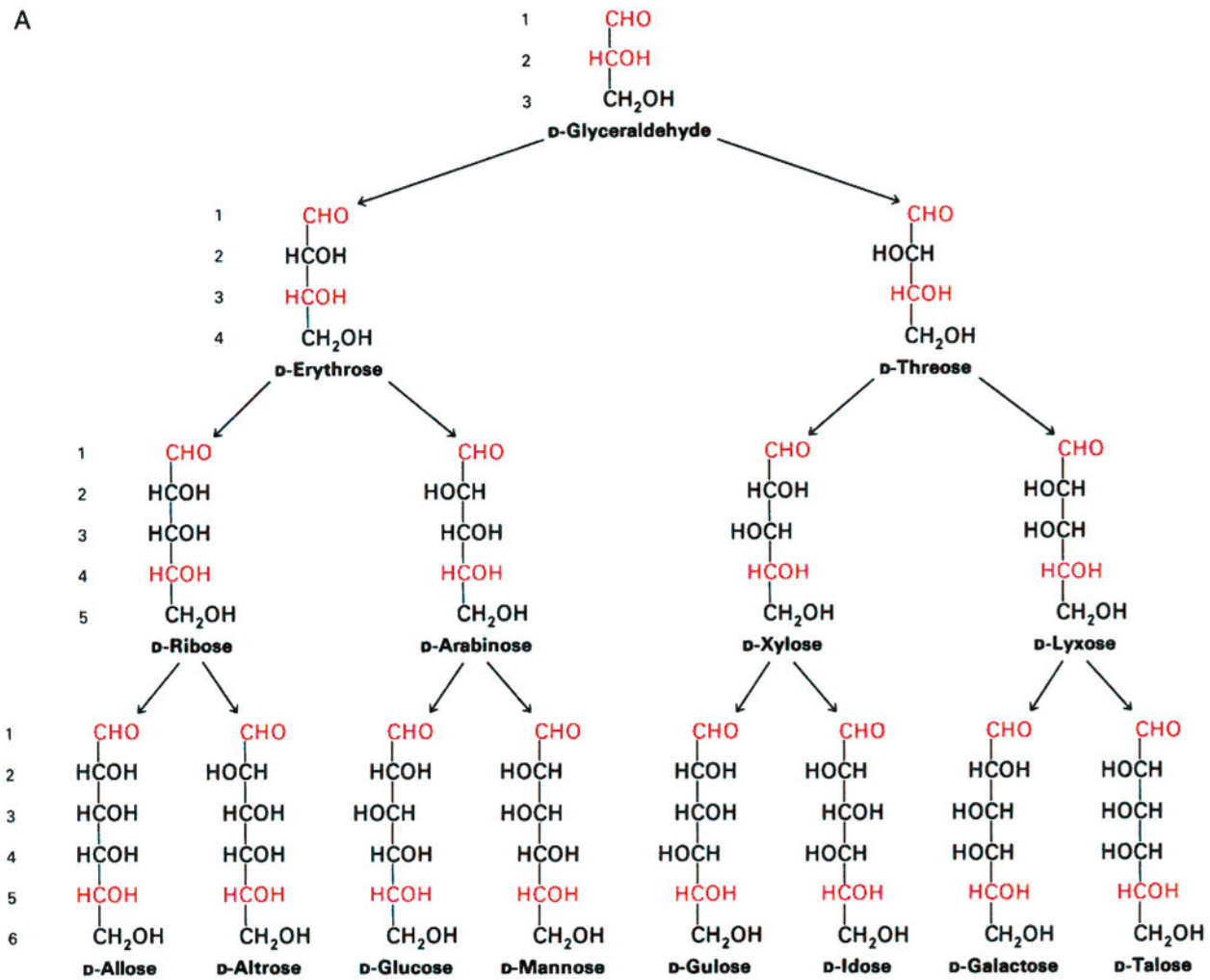


Figure 5.1. Stereochemical relationships among naturally occurring aldoses (A) and ketoses (B).

5.1 MONOSACCHARIDES

Chemically, carbohydrates are *aldehyde or ketone derivatives of polyhydroxy alcohols.* We can classify them on the basis of their functional carbonyl group (—CO—) as aldehydes or ketones. Additionally, we can group them based on their size into **monosaccharides, oligosaccharides,** and **polysaccharides.** These categories are analogous to those used for peptides and nucleotides: monosaccharides contain a single saccharide unit, oligosaccharides contain 2–10 units, and polysaccharides contain more than 10 units.

5.1.1. Some Basic Terminology

Monosaccharides, or *simple sugars,* are compounds that contain a carbonyl group and typically three to seven carbon atoms in an open chain or a cyclic structure. Every carbon atom, except the one that carries the carbonyl group, has an OH group attached. An aldehyde-containing sugar is an **aldose;** a ketone-containing sugar is a **ketose.** We indicate the number of carbon atoms by using Greek combining forms:

3 C	Triose
4 C	Tetrose
5 C	Pentose
6 C	Hexose
7 C	Heptose

Combining the terminology for the functional group and the number of carbon atoms provides a fuller description of a monosaccharide as an aldotriose, a ketotriose, and so on. Since a polyhydroxy alcohol must contain at least two hydroxyl groups, the two smallest monosaccharides (Figure 5.1) are *glyceraldehyde* (an aldotriose) and *dihydroxyacetone* (a ketotriose).

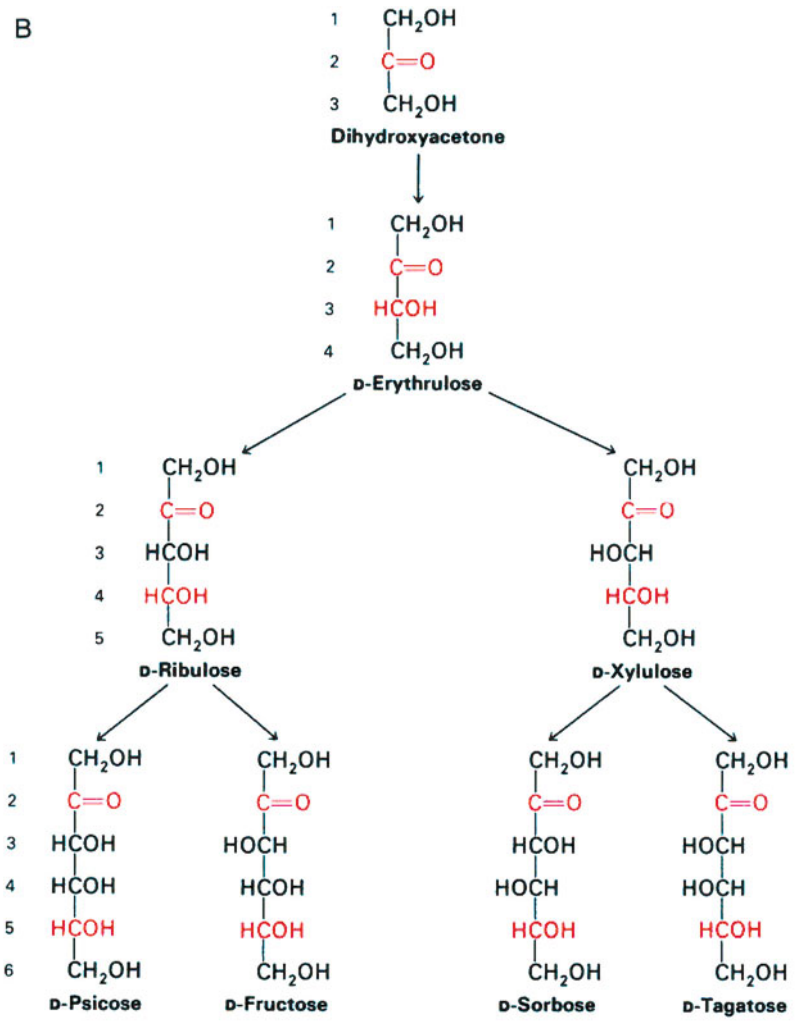


Figure 5.1. (*Continued*).

Monosaccharide configurations are assigned by reference to glyceraldehyde (see Appendix B). We designate a monosaccharide as having an L- or a D-configuration by comparing the *chiral center most remote from the carbonyl group* with the chiral center of glyceraldehyde. In so doing, we write the structure of the monosaccharide with the most highly oxidized carbon at or near the top (the carbonyl carbon in Figure 5.1). Most of the naturally occurring monosaccharides have the D-configuration.

Figure 5.1 illustrates the structural relationships among naturally occurring aldoses and ketoses. All the carbohydrates, except dihydroxyacetone, have one or more chiral centers and occur as optical isomers. We designate two monosaccharides that differ only in the configuration about a single chiral center as **epimers,** and the carbon atom of this chiral center as the **epimeric carbon.** *Glucose* and *mannose* are epimers with respect to C(2), and glucose and *galactose* are epimers with respect to C(4). We call the open-chain structures, shown in Figure 5.1, **Fischer projections.**

5.1.2. Ring Structures

5.1.2A. Mutarotation. The presence of one or more chiral centers in a carbohydrate causes optical activity (see Appendix B). Some D-carbohydrates are dextrorotatory whereas others are levorotatory.

D-Glucose is a dextrorotatory monosaccharide, hence its common name *dextrose.* When we analyze D-glucose carefully, we find that we can separate it into two optical isomers. One isomer, designated α-D-glucose, has a *specific rotation* of +112° (see Appendix B); the other isomer, designated β-D-glucose, has a specific rotation of +18.7°. When we dissolve either of the two pure isomers in water, the specific rotation changes as time passes until it reaches an equilibrium value of +52.7°:

α-D-glucose	⇌ equilibrium mixture	⇌ β-D-glucose
+112°	+52.7°	+18.7°

We call the change of optical rotation as a function of time **mutarotation.** The mutarotations of α- and β-D-glucose clearly must have some structural changes as their basis. What are they?

5.1.2B. Reaction of Carbonyl Groups with Alcohols. To understand what occurs in the glucose system, we must review the reaction of carbonyl groups with alcohols. Both aldehydes and ketones react readily with alcohols in two-step reactions. The first step involves a nucleophilic attack on the carbonyl group by the hydroxyl oxygen of the alcohol. The CO double bond breaks, and the alcohol components convert the aldehyde or ketone to a *hemiacetal* or *hemiketal,* respectively (Figure 5.2). No elimination of atoms occurs in this step; one functional group merely "*adds*" to the other.

In the presence of sufficient alcohol, a second step can occur in which *water is eliminated* and a new carbon–oxygen bond forms with the second alcohol residue. Reaction of a hemiacetal or hemiketal with an alcohol yields an *acetal* or *ketal,* respectively.

5.1.2C. Ring Formation. Because monosaccharides carry both carbonyl and hydroxyl groups within the same molecule, they can form hemiacetals or hemiketals by *intramolecular reactions;* the carbonyl group reacts with a hydroxyl group attached to another carbon atom. This reaction forms the basis of the mutarotation shown by glucose. Glucose, an aldose, forms a cyclic hemiacetal involving the aldehyde group of C(1) and the hydroxyl group of C(5) (Figure 5.3). Because C(1) now has four different substituents attached, it constitutes a new chiral center in the molecule.

With C(1) as a chiral center, the configuration of the attached H and OH can vary, resulting in two new optical isomers, designated α- and β-D-glucose, respectively. These are not mirror images because they differ only in the configuration about C(1); they represent *diastereomers.* We designate two monosaccharides that differ only in the configuration about the carbon of the carbonyl group as **anomers,** and the carbon of this carbonyl group as the **anomeric carbon.** Thus α- and β-D-glucose constitute anomers with respect to C(1). Dissolving either pure α- or pure β-D-glucose in water establishes an equilibrium in which all three forms—the two ring structures and the open-chain form—are present at fixed concentrations. As the reaction proceeds toward this equilibrium, the optical rotation of the mixture changes (mutarotation), finally reaching a value of +52.7°, which corresponds to the rotation of the equilibrium mixture.

5.1.2D. Potential Aldehyde Groups. Many monosaccharides can undergo intramolecular reactions like that described for glucose: anomeric carbons react with specific hydroxyl groups, forming cyclic hemiacetals or hemiketals. These monosaccharides exist in solution predominantly in the form of ring structures. However, because ring structures are in equilibrium with open-chain forms, cyclic carbohydrates possess **potential aldehyde groups.** The capacity to form free aldehyde groups is not limited to aldoses because many ketoses can be converted to aldoses. *In vitro* the conversion occurs in alkaline solution via the *Lobry de Bruyn–Alberta van Eck-*

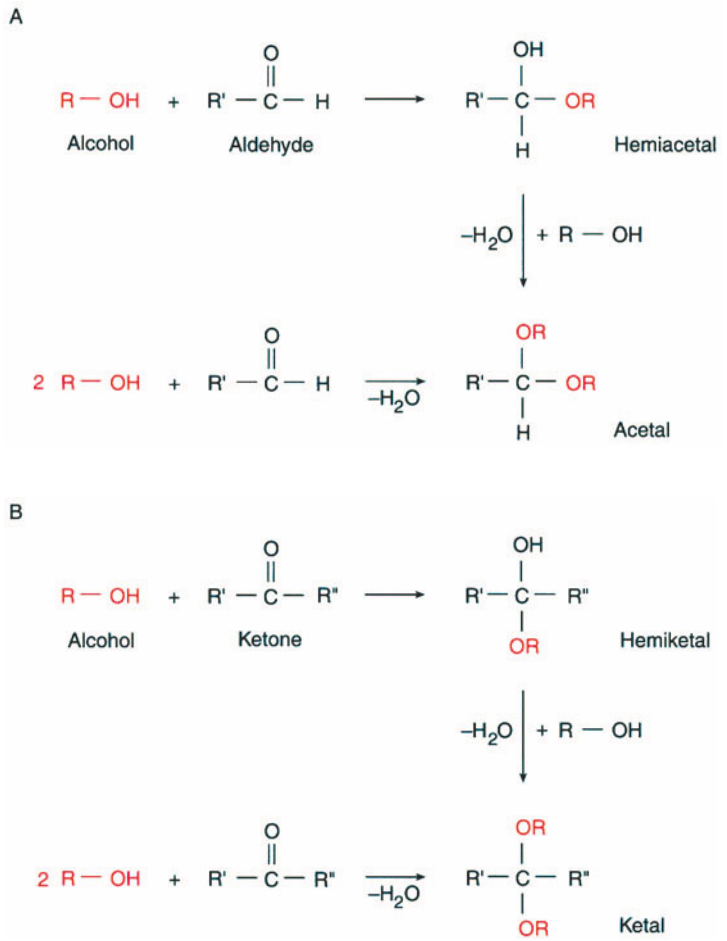


Figure 5.2. Two-step reactions of aldehydes (A) and ketones (B) with alcohols.

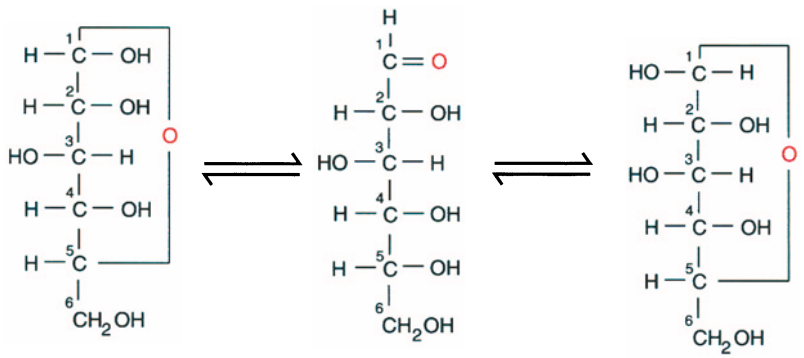


Figure 5.3. Equilibrium between straight-chain glucose and its two cyclic forms. Both ring structures are hemiacetals, and their specific rotations are indicated.

enstein transformation, and in biological systems it occurs due to the action of *isomerases.*

Any monosaccharide possessing a potential aldehyde group undergoes reactions characteristic of the free aldehyde. Adding a suitable reagent shifts the equilibria among the monosaccharide forms, converting ring to straight-chain forms which then react quantitatively with the reagent. Since aldehydes can be oxidized to acids, carbohydrates with potential aldehyde groups can be oxidized to acids as well. Such carbohydrates constitute **reducing sugars.**

5.1.2E. Haworth Projections. Monosaccharides of biological importance form either five- or six-membered rings, structurally related to two cyclic organic compounds, *furan* and *pyran* (Figure 5.4), whose names are used in systematic nomenclature.

We draw monosaccharide ring structures in perspective as oxygen-containing hexagons or pentagons, viewed nearly edge-on. In these formulas, called **Haworth projections,** the lower thick edge of the structure projects forward, at right angles to the paper, and toward the viewer; the opposite edge projects behind the plane of the paper (Figure 5.5). A carbon atom is located at each corner, except for that containing oxygen. Frequently, we abbreviate such projections by omitting either the hydrogens or both the hydrogens and the hydroxyl groups attached to the ring.

Neither the five- nor the six-membered rings are planar. Five-membered rings can take on a number of "pucker" conformations, and six-membered rings can exist in either a "boat" or a "chair" conformation (Figure 5.6), with the chair conformation usually being more stable. Drawing an axis perpendicular to the central plane of the chair conformation permits us to group bonds to substituents into two classes. Some bonds, called *axial bonds,* are roughly parallel to the axis. The remaining bonds, called *equatorial bonds,* are roughly perpendicular to the axis. As a rule, substituents linked via equatorial bonds have greater stability because of less steric interference with other substituents.

For our purposes, Haworth projections constitute adequate representations of monosaccharides, and we shall use them throughout this book. Biochemists refer to five- and six-membered Haworth projections as *furanose* and *pyranose,* respectively. In changing from a Fischer to a Haworth projection, any group to the left in the Fischer projection faces up in the Haworth projection, except for C(4) of pentoses and C(5) of hexoses, where inversions of configuration occur. By convention, we designate a Haworth projection as α when the *hydroxyl group* at the anomeric carbon is *down,* or *below the plane of the ring.* The projection is designated β when the hydroxyl group at the anomeric carbon is *up,* or *above the plane of the ring.*

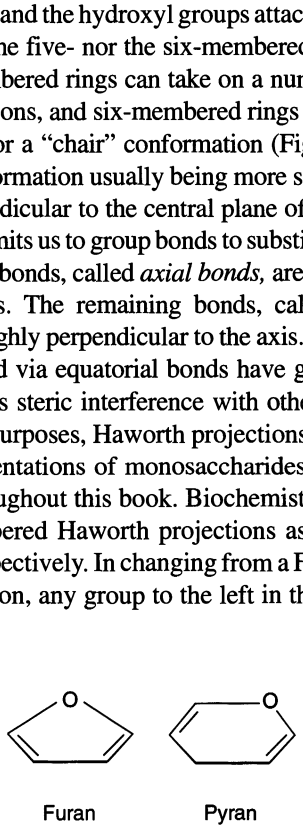


Figure 5.4. Reference compounds for carbohydrate structures.

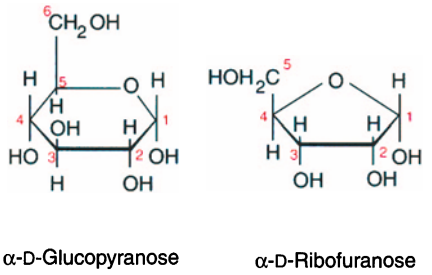


Figure 5.5. Haworth projections of glucose and ribose.

The bond formed between the anomeric carbon and some group or molecule other than the original H and OH is called a **glycosidic bond.** We designate these bonds as α or β, depending on the configuration at the anomeric carbon. When the glycosidic bond links an alcohol group, the carbohydrate forms an acetal or ketal. The anomeric carbon of an acetal or ketal has a "locked-in" configuration and cannot undergo mutarotation. To distinguish acetals (or ketals) from hemiacetals (or hemiketals), we denote the corresponding five- and six-membered rings of the former as *furanosides* and *pyranosides,* respectively. The glycosidic bond links monosaccharides together to form oligo- and polysaccharides; its role in carbohydrate chemistry is analogous to that of the peptide bond in protein chemistry.

5.1.3. Derived Monosaccharides

In addition to their occurrence as simple sugars, monosaccharides occur in the form of various derivatives. Biologically important compounds include sugar acids, sugar alcohols, amino sugars, deoxysugars, and phosphoric acid esters.

5.1.3A. Sugar Acids. **Sugar acids** form by oxidation of a carbonyl group and/or an alcohol group of a simple sugar to a carboxyl group. Oxidation of the terminal aldehyde group yields an *aldonic acid,* oxidation of the terminal $-CH_2OH$ forms a *uronic acid,* and oxidation of both of these groups produces an *aldaric acid.* The sugar acids formed from glucose are gluconic, glucuronic, and glucaric acid, respectively (Figure 5.7).

A special sugar acid, *vitamin C,* occurs in two active forms, *ascorbic acid* and *dehydroascorbic acid* (Figure 5.8). Both compounds derive from an aldonic acid by loss of two hydrogens and formation of an intramolecular es-

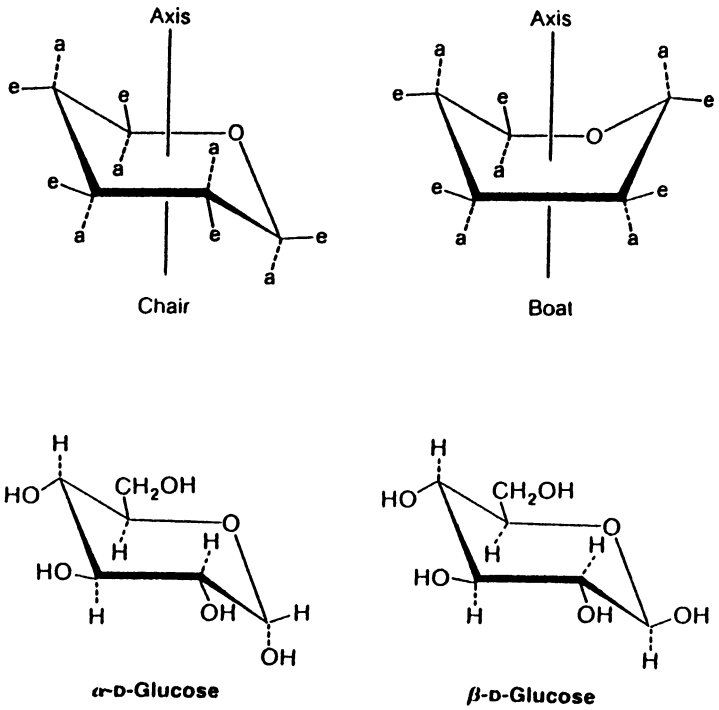


Figure 5.6. Conformational isomers of the pyranose ring. *Top:* Generalized formulas of chair and boat conformations: a, axial bond; e, equatorial bond. *Bottom:* Chair conformations of α- and β-D-glucose.

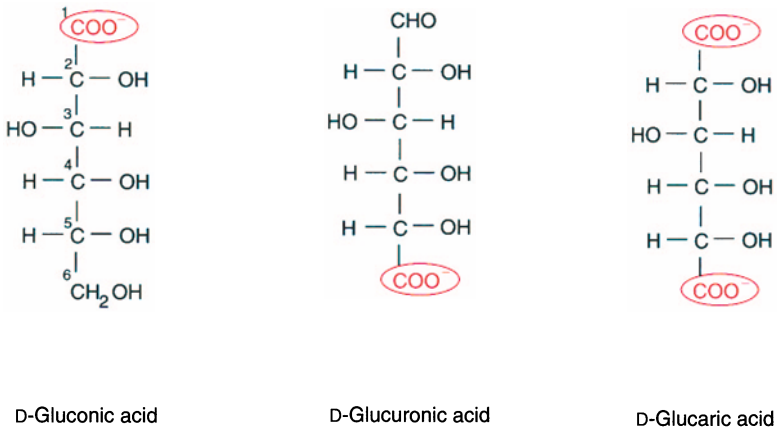


Figure 5.7. Sugar acids derived from glucose.

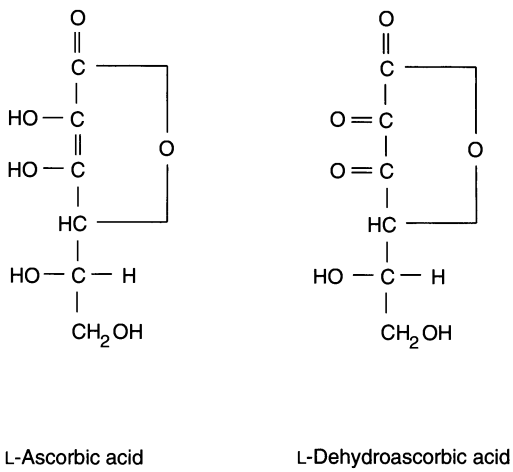


Figure 5.8. The two active forms of vitamin C.

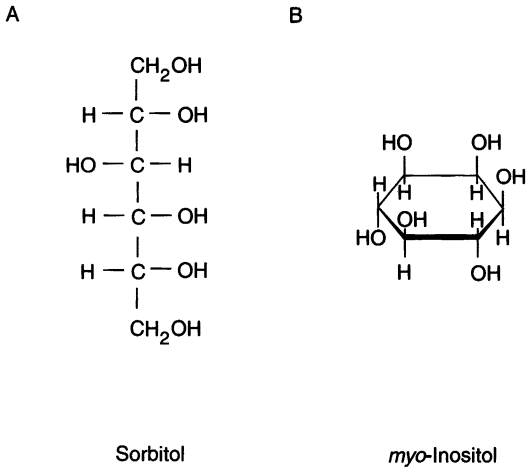


Figure 5.9. Examples of sugar alcohols.

ter (lactone). Vitamin C, the oldest therapeutically used vitamin, was provided to British sailors in the 1750s in the form of limes (hence the epithet "limey") but was only identified in 1930 as the essential dietary component of citrus juices. Humans require vitamin C for enzyme-catalyzed hydroxylations of proline and lysine to form hydroxyproline and hydroxylysine, which occur in collagen. A deficiency of vitamin C produces **scurvy,** a disease whose symptoms include skin lesions, blood vessel fragility, and poor wound healing.

5.1.3B. Sugar Alcohols. Reduction of the carbonyl group of a simple sugar to an alcoholic hydroxyl yields a **sugar alcohol** or *alditol.* The sugar alcohol derived from glucose, *sorbitol,* is used as a preservative (Figure 5.9A). Another alditol, *inositol,* occurs in the form of nine stereoisomers, and one of these, *myo-inositol,* serves as a component of phospholipids (Figure 5.9B).

5.1.3C. Amino Sugars. Replacement of one or more hydroxyl groups in a monosaccharide by amino groups yields an **amino sugar.** The amino group frequently occurs in acetylated form. Amino sugars serve as structural components of cell walls and cell coats. Figure 5.10 shows several amino sugars and related structures. Biochemists refer to *N*-acetylneuraminic acid and its derivatives as *sialic acids.* These compounds often form parts of glycoproteins.

5.1.3D. Deoxysugars. Replacement of a monosaccharide's hydroxyl group by hydrogen yields a **deoxysugar.** The most important deoxysugar in biological systems is *deoxyribose* (more precisely, D-2-deoxyribose), an essential building block of DNA (Figure 5.11). (Ribose, the parent compound of deoxyribose, is an essential building block of RNA.)

5.1.3E. Esters. Recall that an ester forms by reaction between an oxyacid and an alcohol, splitting out a

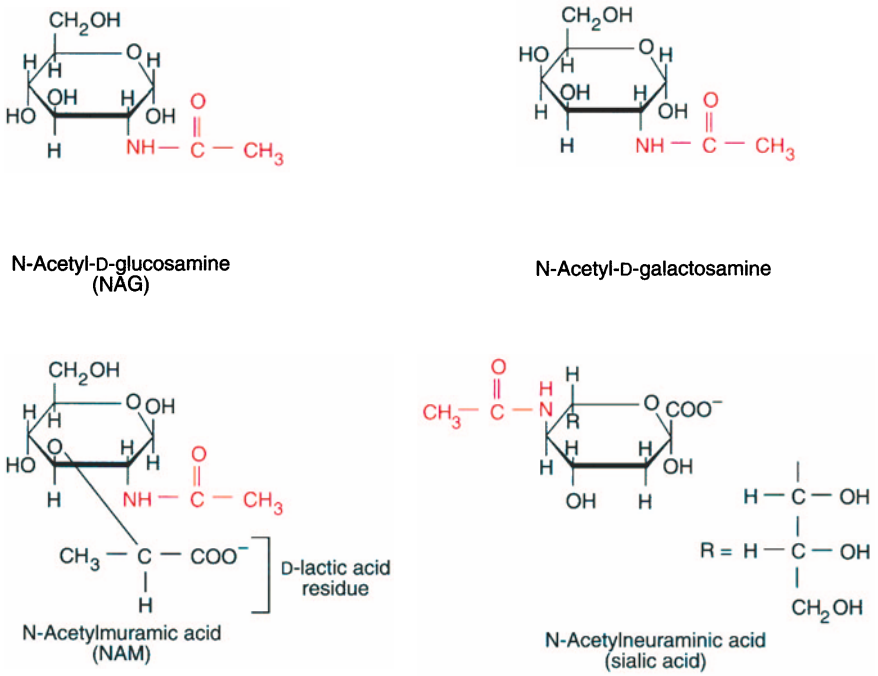
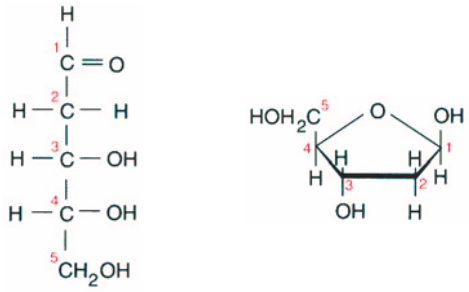


Figure 5.10. Some amino sugars and related structures.



D-2-Deoxyribose

Figure 5.11. Fischer (*left*) and Haworth (*right*) projections of D-2-deoxyribose.

molecule of water. Because carbohydrates carry multiple hydroxyl groups, they can form many different kinds of esters, depending on the type of acid esterified and on the position(s) where esterification takes place. Of particular importance in biochemical systems are esters formed with phosphoric acid.

5.2. CHEMICAL REACTIONS OF CARBOHYDRATES

5.2.1. Oxidation–Reduction Reactions

Carbohydrates can undergo various oxidation–reduction reactions. Typical reactions include oxidation of aldehyde groups to carboxyl groups or reduction of aldehyde groups to alcoholic hydroxyl groups. These reactions yield sugar acids and sugar alcohols, respectively.

Two simple laboratory tests have been used for many years as a qualitative test for reducing sugars. In one, the sugar reduces a silver ammonium complex, called *Tollens' reagent,* producing a silver mirror on the test tube wall:

$$RCHO + 2Ag(NH_3)_2^+ + 2OH^- \longrightarrow RCOO^- + 2Ag + 3NH_3 + NH_4^+ + H_2O$$

In the other test, the sugar reduces cupric ions in alkaline solution, called *Fehling's* or *Benedict's reagent,* yielding a red precipitate of cuprous oxide, Cu_2O:

$$RCHO + 2Cu^{2+} + 5OH^- \longrightarrow RCOO^- + Cu_2O + 3H_2O$$

Many carbohydrates participate in another set of oxidation–reduction reactions, *fermentation,* discussed in Section 8.1. Table 5.1 shows the susceptibility of some common carbohydrates to fermentation by bakers' yeast.

Table 5.1. Fermentability of Some Carbohydrates

Carbohydrate	Fermentable by bakers' yeast
Fructose	+
Galactose	+
Glucose	+
Glycogen	–
Lactose	–
Maltose	+
Ribose	–
Starch	–
Sucrose	+

All carbohydrates, containing only C, H, and O, can be completely oxidized to CO_2 and H_2O. In the laboratory, we perform such oxidations by means of a calorimeter. In living organisms, carbohydrates are oxidized to CO_2 and H_2O by means of a large number of enzyme-catalyzed steps.

$$\text{All carbohydrates (C, H, O)} \xrightarrow[\text{metabolism } [O_2]]{\text{laboratory } [O_2]} CO_2, H_2O$$

5.2.2. Phosphorylation

Reaction of carbohydrates with phosphoric acid produces **phosphoric acid esters.** These are of great importance in biochemical systems, both as intermediates in metabolism and as structural components of nucleic acids and other compounds. We designate the esterified portion of phosphoric acid (Figure 5.12), called a **phosphoryl group,** variously as

Dissociations of the first and second protons from a phosphoryl group are analogous to proton dissociations from H_3PO_4 and have similar pK_a' values. Hence, at physiological pH, we consider the phosphoryl group to be ionized as shown above. Note that we refer to the different ionic forms of free H_3PO_4, singly or as a mixture, as *inorganic phosphate* and denote them as P_i.

5.2.3. Etherification

Replacing the H of an OH group in a carbohydrate by an organic group, such as a methyl or an ethyl group, pro-

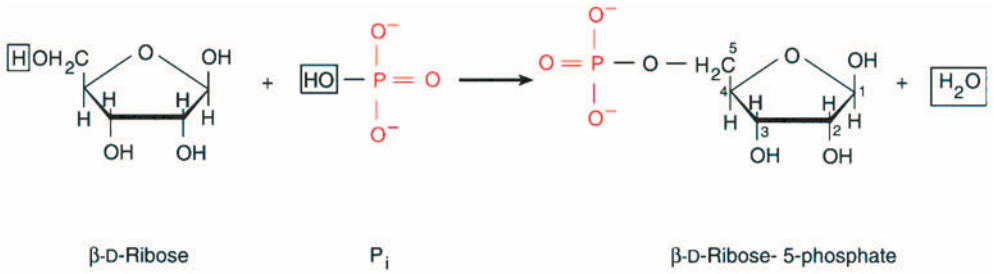


Figure 5.12. Formation of ribose 5-phosphate, a phosphoric acid ester of ribose.

duces an ether (R—O—R′). Naturally occurring ethers of carbohydrates are rare. One of the few biomolecules in which an ether linkage occurs is N-acetylmuramic acid (see Figure 5.10).

In the 1930s, N. Haworth pioneered a technique termed **exhaustive methylation** and useful for elucidating structural features of oligo- and polysaccharides. In this technique, we first convert all available –OH groups of a carbohydrate to $-O-CH_3$ groups. Following the methylation, we subject the compound to acid hydrolysis, which cleaves all of the glycosidic bonds between sugar residues and yields a mixture of methylated monosaccharides (Figure 5.13). We then determine the monosaccharides both qualitatively and quantitatively, typically by gas–liquid chromatography combined with mass spectrometry. The method may be used as an end-group analysis (Section 2.4). Depending on the carbohydrate studied, exhaustive methylation may permit partial or full characterization of the compound's structure, including the types of glycosidic bonds between monosaccharide units.

5.3. OLIGOSACCHARIDES

Oligosaccharides contain 2–10 monosaccharide units; those of greatest nutritional importance are **disaccharides.** We describe oligosaccharides structurally by specifying the *number, type, and sequence of the monosaccharide units, as well as the types of linkages between them.*

5.3.1. Maltose

Maltose, a disaccharide of D-glucose (Figure 5.14), serves as a building block for starch and glycogen. We consider and name its structure from left to right. The left glucose residue is an acetal and has a fixed configuration, namely, α-D-glucose. The hydroxyl group of the anomeric carbon of this residue has the α-configuration. The second glucose residue is a hemiacetal and has a variable configuration; the hydroxyl group of its anomeric carbon can have either an α- or a β-configuration, resulting in α- or

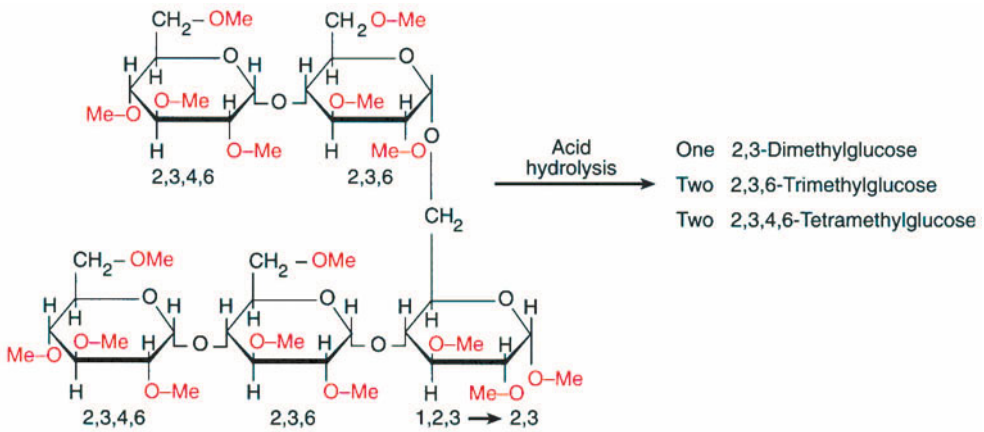


Figure 5.13. The technique of exhaustive methylation. Acid hydrolysis of one molecule of methylated pentasaccharide (Me = methyl group) yields five molecules of products. The methyl group at the anomeric carbon is lost in the process but may be retained if gentle enzymatic hydrolysis is used instead. In either case, determining 2,3-dimethylglucose or 1,2,3-trimethylglucose constitutes an end-group analysis of the pentasaccharide.

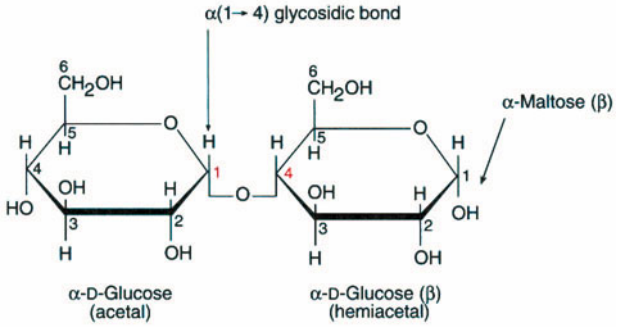


Figure 5.14. Maltose [*O*-α-D-glucopyranosyl-(1→4)-α-D-glucopyranose].

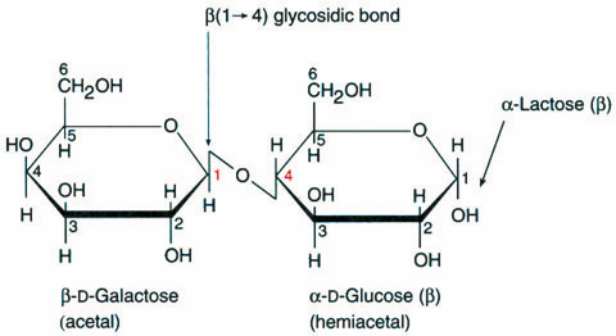


Figure 5.15. Lactose [*O*-β-D-galactopyranosyl-(1→4)-β-D-glucopyranose]. The β-glycosidic bond requires that one residue be turned "upside down" or rotated by 180° (not shown).

β-D-glucose. Thus, maltose occurs in two forms, designated α-maltose or β-maltose, based on the configuration at the anomeric carbon of the hemiacetal residue.

We term the bond linking the two glucose residues an α(1 → 4) glycosidic bond; (1 → 4) designates the linked carbons, and α refers to the configuration of the anomeric carbon, C(1), of the left glucose residue. Because linked acetals and ketals are named *furanosyl* for a five-membered ring and *pyranosyl* for a six-membered ring, maltose has the full systematic name of O-α-D-glucopyranosyl-(1 → 4)-α-D-glucopyranose, where "*O*" indicates the oxygen link between the two residues. The left glucose residue does not have a potential aldehyde group; its cyclic carbon chain cannot be opened at the anomeric carbon to produce a free aldehyde group. That can only occur *after* the glycosidic bond of the disaccharide has been hydrolyzed, converting the acetal to a hemiacetal. Consequently, the left glucose residue represents the **nonreducing end** of the disaccharide. The right glucose residue does have a potential aldehyde group and constitutes the molecule's **reducing end.** Because of this end's properties, the disaccharide as a whole has a potential aldehyde group, is a reducing sugar, and shows mutarotation.

5.3.2. Lactose

Lactose, the main carbohydrate in milk (Figure 5.15), is a disaccharide of galactose and glucose. It differs from maltose in two respects. First, its two monosaccharide units are nonidentical. Second, the two monosaccharides are linked in a different fashion. The bond between them is designated as a β(1 → 4) glycosidic bond, where β refers to the configuration of the anomeric carbon of the left residue (galactose, an acetal).

As for maltose, there exist two forms of lactose (α- and β-lactose), based on the configuration of the anomeric carbon of the right residue (glucose, a hemiacetal). Because of the properties of this residue, lactose has a potential aldehyde group, is a reducing sugar, and shows mutarotation.

Normally, *lactase,* an enzyme in the intestinal villi, catalyzes the hydrolysis of lactose to glucose and galactose. In some individuals, an inherited metabolic defect results in a deficiency of lactase; in others, the condition may develop during adult life. In either case, lactose, instead of being digested, moves unchanged through the digestive tract and into the colon. There it undergoes fermentation by intestinal bacteria with the production of hydrogen gas, carbon dioxide, and organic acids that lead to abdominal cramps, pain, and diarrhea. The syndrome is known as **lactose intolerance.**

5.3.3. Sucrose

Sucrose, a disaccharide composed of glucose (an aldose) and fructose (a ketose), is the sugar you commonly use at home (Figure 5.16). We derive sucrose from sugarcane, sugar beets, and maple syrup.

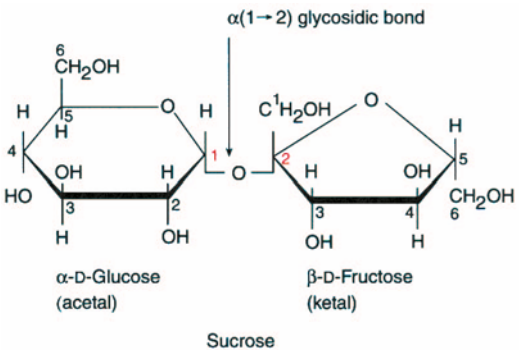


Figure 5.16. Sucrose [*O*-α-D-glucopyranosyl-(1→2)-β-D-fructofuranoside]. Fructose forms a five-membered ring because of its keto group at C(2).

Table 5.2. Relative Sweetness of Some Compounds

Compound	Relative sweetness (per gram of compound)
Lactose	0.2
Galactose	0.3
Maltose	0.3
Glucose	0.7
Sucrose	1.0
Fructose	1.7
Cyclamate	30
Aspartame	200
Saccharin	500
Monellin	2500

The link between the two monosaccharide units of sucrose constitutes an *α(1→2) glycosidic bond,* formed between the *anomeric carbons of both glucose and fructose.* In this structure, glucose exists as an acetal, and fructose as a ketal. Neither monosaccharide can have an alternate configuration about its anomeric carbon. Consequently, unlike maltose and lactose, sucrose occurs in only one form. Because both anomeric carbons are covalently linked via a glycosidic bond, sucrose does not have a potential aldehyde group, is not a reducing sugar, and does not show mutarotation.

Sweetness constitutes a property that is not limited to sucrose; other carbohydrates have a sweet taste, though to varying degrees (Table 5.2). In addition to carbohydrates, synthetic compounds and even proteins can serve as sweeteners. Fructose has about twice the sweetness of sucrose, which accounts for its use in "diet food" items; to attain the same sweetness as obtainable with sucrose, one can use fewer grams of fructose (and thus take in fewer joules or calories). There appear to be no harmful effects from the use of "high-fructose" products, but there exists lingering concern regarding synthetic sweeteners. Some investigators suspect that *aspartame* (see Section 2.5) may cause neurological problems, especially in individuals who cannot metabolize phenylalanine readily. Other workers have shown that *saccharin* (Figure 5.17) can cause cancer in laboratory animals. The government banned *cyclamate* (Figure 5.18) from prepared foods after research demonstrated its carcinogenicity in animals. *Monellin* is a carbohydrate-free protein derived from the sap of serendipity berries, the fruit of a West African plant. The protein consists of two polypeptide chains, containing 44 and 50 amino acids, respectively, and has an extremely high specificity for sweet-taste receptors.

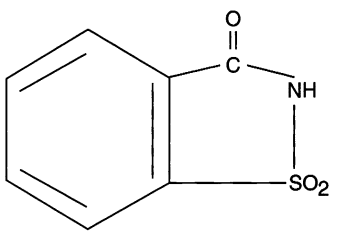


Figure 5.17. Saccharin.

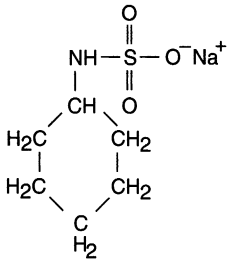


Figure 5.18. Cyclamate (sodium cyclohexyl sulfamate).

When we hydrolyze sucrose to glucose and fructose, either enzymatically or nonenzymatically, the optical rotation of the solution changes from *dextro* to *levo.* Because of this *inversion* of optical rotation, researchers call hydrolyzed sucrose *invert sugar;* the enzyme catalyzing this hydrolysis has the trivial name *invertase.*

5.4. HOMOPOLYSACCHARIDES

Homopolysaccharides—polysaccharides composed of a single type of monosaccharide unit—occur both as storage forms of food and energy and as structural components of cells. Because these large molecules have relatively simple structures, they possess little information content but lend themselves well to storage in large quantities and to formation of fibers for structural frameworks.

5.4.1. Storage Polysaccharides

5.4.1A. Starch. Starch represents the storage carbohydrate of plants. It occurs in special *starch granules* that also contain enzymes involved in starch synthesis and degradation. Starch occurs both as a linear polymer, **α-amylose,** and as a branched variety, **amylopectin** (Figure 5.19). Both types of starch consist entirely of D-glucose units, hence the term *glucans;* polysaccharides in general are called *glycans.*

In α-amylose, glucose units are linked by means of *α(1→4) glycosidic bonds.* Thus, maltose constitutes the

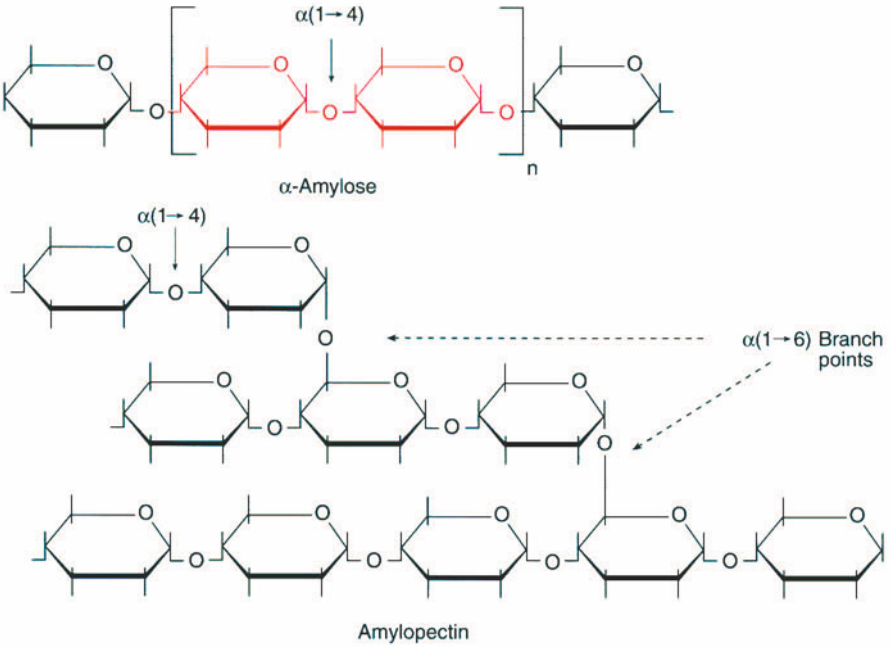


Figure 5.19. Linear (α-amylose) and branched (amylopectin) forms of starch.

disaccharide repeating unit of amylose. Because of the α(1→4) glycosidic bonds, each glucose residue is angled with respect to the next, causing a natural turning in the chain and favoring a regular helical configuration. This leads to amylose assuming the structure of a left-handed helix (Figure 5.20). The large helix core can accommodate a string of iodine atoms, producing a starch–iodine complex that has an intense blue color. Formation of this complex has long been used as a qualitative test for starch. A minimal amylose length, equivalent to six helical turns (36 glucose residues), is required to produce the color.

In amylopectin, glucose residues are linked by α(1 → 4) glycosidic bonds within each chain and by α(1→6) glycosidic bonds at branch points. Branch points occur about every 12–25 residues along the α(1 → 4) chain.

Naturally occurring starch consists of a mixture of amylose and amylopectin. Typically, starches contain 10–30% amylose and 70–90% amylopectin. Variations in properties of corn starch, potato starch, rice starch, and the like arise from differences in the relative amounts and chain lengths of amylose and amylopectin and in the number and location of branch points in amylopectin.

5.4.1B. Glycogen. Animals store their carbohydrate in the form of **glycogen,** often called *animal starch.* Glycogen's structure resembles that of amylopectin but is more highly branched, with branch points occurring about every 8–10 residues along the α(1→4) chain. This extensive branching allows for rapid generation of usable energy when glycogen is degraded in catabolism (Section 10.5). Glycogen is stored in *glycogen granules* in the cytoplasm and occurs primarily in two tissues, muscle and liver.

5.4.1C. Dextrans. Dextrans constitute a group of branched-chain polysaccharides of D-glucose found in yeast and some bacteria. In dextrans, glucose residues are linked α(1→6) and α(1→3), and the chains vary in their length and extent of branching. Dextrans serve as storage carbohydrates and as components of bacterial capsules.

5.4.1D. Degradation of Starch and Glycogen. Enzymes called **amylases** catalyze the hydrolysis of α(1→4) glycosidic bonds in both starch and glycogen. *α-Amylase,* an *endoenzyme,* catalyzes random hydrolysis and yields a mixture of glucose and maltose; *β-Amylase,* an *exoenzyme,* catalyzes sequential removal of maltose residues, beginning at the nonreducing end (Figure 5.21). Both enzymes can lead to complete digestion of amylose, but neither enzyme can catalyze the hydrolysis of α(1→6) bonds at branch points in amy-

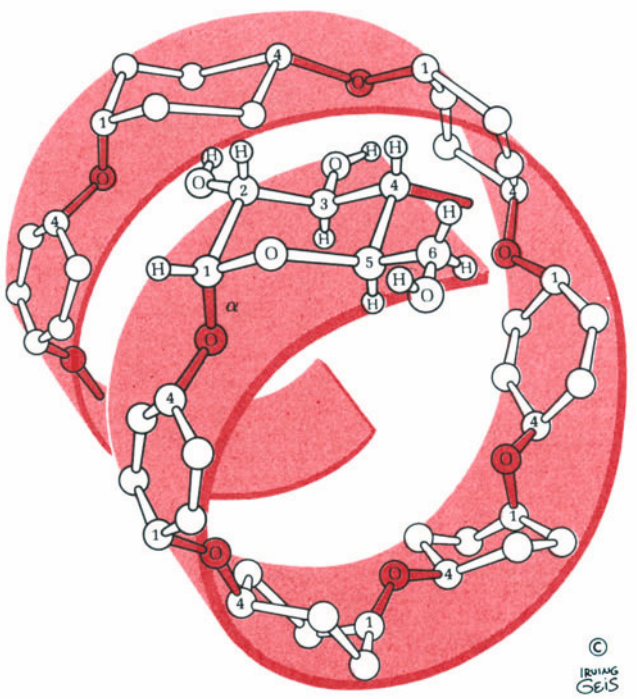


Figure 5.20. The secondary structure of α-amylose. A chain of glucose residues forms a left-handed helix having six residues per turn. (Figure copyrighted © by Irving Geis. Reproduced with permission.)

lopectin or glycogen. Treating amylopectin or glycogen with α- or β-amylase leaves an undigested, branched *core* behind. Hydrolysis of the α(1→6) branch points requires the action of a third enzyme called *debranching enzyme*.

Both starch and glycogen yield a heterogeneous group of substances, called **dextrins,** when subjected to partial degradation by heat, acid, or the action of enzymes. We use dextrins in the manufacture of glues and mucilage, as inert fillers of pills, and in the production of matches, fireworks, and explosives. *Limit dextrin* represents the branched core of either amylopectin or glycogen that remains after the polysaccharide has been digested with either α- or β-amylase.

5.4.2. Structural Polysaccharides

5.4.2A. Cellulose. **Cellulose** is the most abundant single polymer in the biosphere and the major polysaccharide in woody and fibrous plants. Like α-amylose, cellulose is a linear polymer of D-glucose. It differs from amylose in that the linkage between two glucose residues consists of a β(1→4) glycosidic bond (Figure 5.22). A disaccharide called *cellobiose* [*O*-β-D-glu-

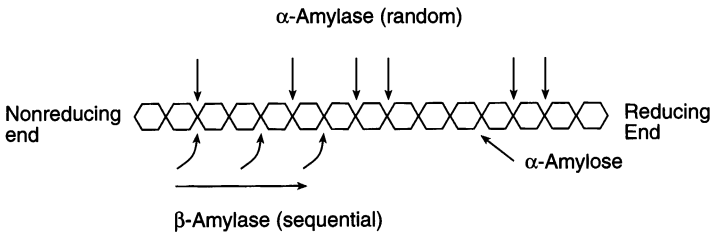


Figure 5.21. The action of α- and β-amylase on α-amylose.

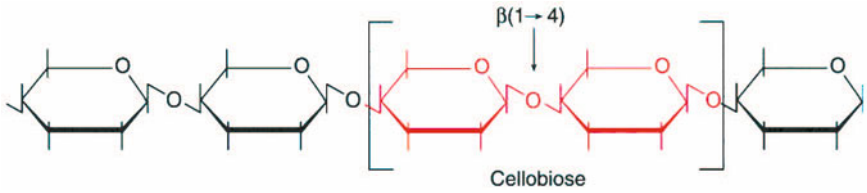


Figure 5.22. Cellulose. The 180° flip of successive glucose residues has been omitted for simplicity.

copyranosyl-(1→ 4)-β-D-glucopyranose] constitutes the repeating unit of cellulose. The difference between α-amylose and cellulose may appear to be minor, but that is not the case. Because of the two types of glycosidic bonds, starch and cellulose have significantly different properties.

The β-glycosidic bond causes a 180° flip of successive glucose residues in cellulose and results in the polymer assuming an easily packed, fully extended configuration. Cellulose fibers are insoluble in water and consist of some 40 parallel chains, with each chain containing up to 15,000 residues. A network of intra- and interchain H-bonds holds the chains together (Figure 5.23). Bundles of some 2000 chains constitute a *microfibril.*

Humans possess enzymes that catalyze the hydrolysis of α(1→4) glycosidic bonds of starch and glycogen (*a-glycosidases*, such as α-amylase) but lack the enzyme capable of recognizing the β(1→4) bond of cellulose. Consequently, we readily digest starch and use it as a major source of metabolic energy, but we cannot digest cellulose and must excrete it. Cows and other ruminants have symbiotic bacteria in their intestinal tract that produce the necessary β-glycosidase, also called *cellulase.* Hence they can feed on grass and other plants. Termites use wood as a nutrient because protozoans in their guts can digest cellulose. Mushrooms, too, can grow on wood since they produce cellulase.

5.4.2B. Chitin. **Chitin** resembles cellulose in both structure and function. It, too, consists of monosaccharide units, linked via β(1→4) glycosidic bonds to form a linear polymer. However, the repeating disaccharide unit does not contain simple sugars but rather derived monosaccharides, specifically *N-acetyl*-D-glucosamine (Figure 5.10). Much as they function in cellulose, hydrogen bonds hold bundles of chitin chains together to form strong fibers. Chitin is a major constituent of the hard, horny exoskeleton of insects and crustaceans, often forming a matrix on which mineralization takes place.

5.5. HETEROPOLYSACCHARIDES

Heteropolysaccharides or **complex polysaccharides** are polysaccharides composed of different types of monosaccharide units. Heteropolysaccharides occur

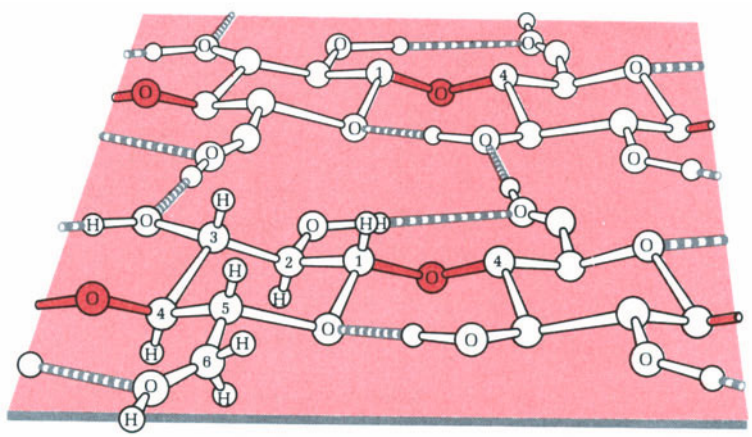


Figure 5.23. A structural model of cellulose showing the 180° flip of successive glucose residues. Parallel chains are linked by hydrogen bonds. (Figure copyrighted © by Irving Geis. Reproduced with permission.)

widely in nature and differ substantially from the homopolysaccharides discussed in the previous section. While many homopolysaccharides have relatively simple molecular forms, most heteropolysaccharides have a complex structure. Homopolysaccharides have two major functions, serving as either storage material or structural components, but heteropolysaccharides have many different functions. Lastly, heteropolysaccharides occur frequently in combination with noncarbohydrate material to which they are linked covalently or noncovalently.

5.5.1. Glycosaminoglycans

Glycosaminoglycans consist of repeating disaccharide units, linked to form linear polymers. The disaccharide always contains a hexose amine (*N-acetylglucosamine* or *N-acetylgalactosamine*) and generally a residue of a uronic acid and a sulfate group. Chemists recognize several distinct classes of glycosaminoglycans (Figure 5.24).

A major function of these glycans consists of forming a matrix that holds together protein components of skin and connective tissue. At physiological pH, glycosaminoglycans exist as highly ionized, polyanionic compounds. Because of their many charges, they are extensively hydrated and give rise to very viscous and gelatinous solutions (hence the older term *mucopolysaccharide*). These viscous properties make them such a good cement for skin and connective tissue.

Hyaluronic acid occurs in connective tissue, the synovial fluid of joints, the vitreous humor of the eye, and cartilage. It helps block the spread of invading microorganisms and toxic substances. It also serves as a lubricant and shock absorbent. *Hyaluronidase,* an enzyme present in snake venom, catalyzes the hydrolysis of hyaluronic acid. This enzyme greatly diminishes the capacity of an organism to defend itself against the spread of pathogens and toxins at the site of a snake bite.

Chondroitin sulfate is found in bone and cartilage and occurs in two forms that differ in their sulfation sites. *Dermatan sulfate* occurs in skin, blood vessels, and heart valves. *Keratan sulfate* is present in the cornea of the eye, cartilage, and intervertebral disks.

Heparin occurs in intracellular granules that line arterial walls. It is released on injury and enters the blood-

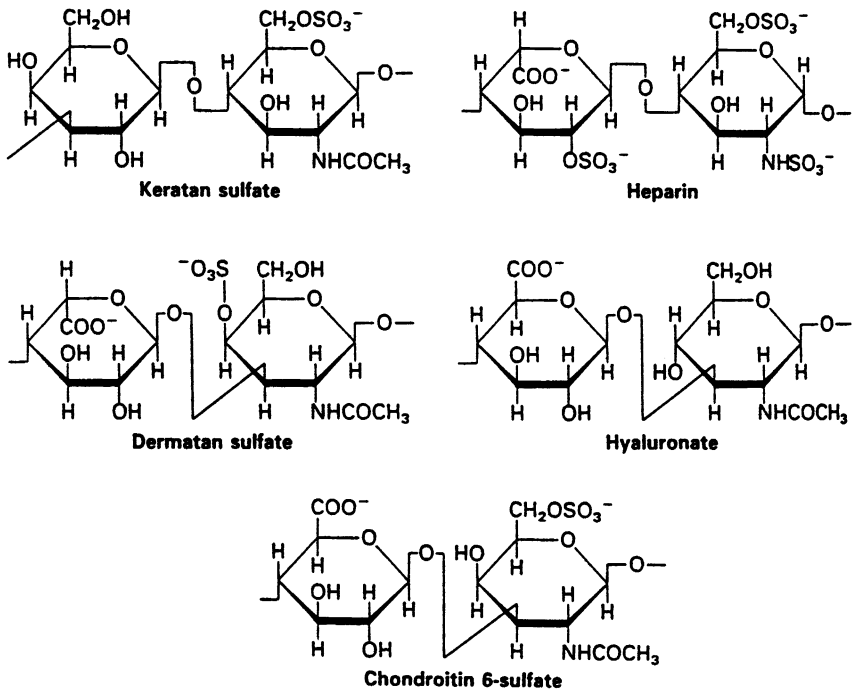


Figure 5.24. The disaccharide repeating units of some glycosaminoglycans. All repeating units, except those of heparin, are linked in alternating fashion, involving β(1→3) and β(1→4) glycosidic bonds. A variant form of chondroitin sulfate has the SO_3 group linked at C(4) as in dermatan sulfate.

stream, where it functions to prevent clot formation. Because of these properties, we use it medicinally as an *anticoagulant.* Heparin contains predominantly the repeating units shown in Figure 5.24, but also some other sugar derivatives, and is sulfated to varying degrees. *Heparan sulfate,* a widespread component of cell surfaces and arterial walls, closely resembles heparin but has a more variable composition; it contains some *N*-acetyl groups and fewer *N*- and *O*-sulfate groups.

5.5.2. Glycoproteins

Conjugated proteins having a covalently linked carbohydrate component are called **glycoproteins.** We now know that most proteins exist as glycoproteins and that glycoproteins cover the entire gamut of protein function, occurring as enzymes, hormones, antibodies, structural proteins, and transport proteins.

We refer to the covalent attachment of carbohydrates to proteins as **glycosylation.** The process requires formation of one of two types of glycosidic bonds (Figure 5.25). In one, the bond forms between the carbohydrate (usually, *N*-acetylgalactosamine) and the hydroxyl group of either serine, threonine, or hydroxylysine, resulting in an *O-linked oligosaccharide (O-linked glycan).* Alternatively, the bond can form between the carbohydrate (usually *N*-acetylglucosamine) and the amino group of asparagine, resulting in an *N-linked oligosaccharide (N-linked glycan).*

In glycoproteins, the covalently linked carbohydrate may be either N- or O-linked and may be either a monosaccharide or a short oligosaccharide fragment. Some glycoproteins contain one or a few carbohydrate groups; others contain numerous oligosaccharide side chains, which

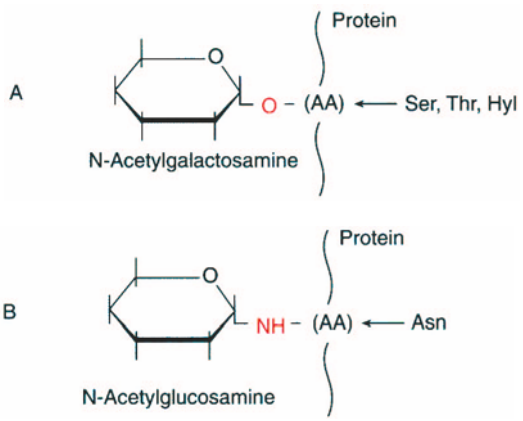


Figure 5.25. Glycosylation, the covalent linking of carbohydrates to proteins. (A) O-linked glycoprotein; (B) N-linked glycoprotein.

may be linear or branched. As a result, the carbohydrate content of glycoproteins varies from less than 1% to more than 90% by weight. Glycoproteins frequently contain *N*-acetylneuraminic acid (Figure 5.10) as one of their components.

Glycoproteins have an extracellular location and function; they are generally secreted into body fluids or serve as membrane proteins. In many cases, the specific function of the attached carbohydrate chain still needs to be elucidated. Scientists feel confident, however, that a major role of these carbohydrates is an informational one. Carbohydrates provide *recognition factors* to a protein, thereby guiding the protein's interaction with other molecules. A protein with a multitude of small carbohydrate chains on its surface has much greater information content than one without these additional tags. The presence of carbohydrate on an antibody molecule probably serves precisely this function. Researchers believe that carbohydrates in membranes and cell coats play a role in cell–cell interactions and in the orientation of proteins or lipids on membranes.

5.5.3. Proteoglycans

Proteoglycans are high-molecular-weight substances that contain large amounts (typically, 90–95%) of glycosaminoglycans linked covalently to a polypeptide backbone. Proteoglycans constitute massive aggregates (*proteoglycan aggregates*) rather than macromolecules. The aggregates have polyanionic character, and their properties resemble those of polysaccharides more than those of proteins. Proteoglycans form the ground substance in the extracellular matrix of connective tissue and serve as lubricants and support elements.

We can illustrate the structure of proteoglycans by the proteoglycan aggregates found in cartilage (Figure 5.26). The aggregate forms by noncovalent attachment of many long polypeptide chains (MW 200,000–300,000), termed *core proteins,* to a long strand of *hyaluronic acid.* A different protein, termed *link protein* (MW 40,000–60,000), stabilizes the attachment of each core protein. Lastly, numerous glycosaminoglycans, specifically *chondroitin sulfate* and *keratan sulfate,* become linked covalently to each core protein.

The core protein of cartilage appears to have three distinct binding regions. Its N-terminal region, near the hyaluronic acid binding site, contains a few N-linked oligosaccharides. This is followed by a region rich in O-linked keratan sulfate and a C-terminal region rich in O-linked chondroitin sulfate.

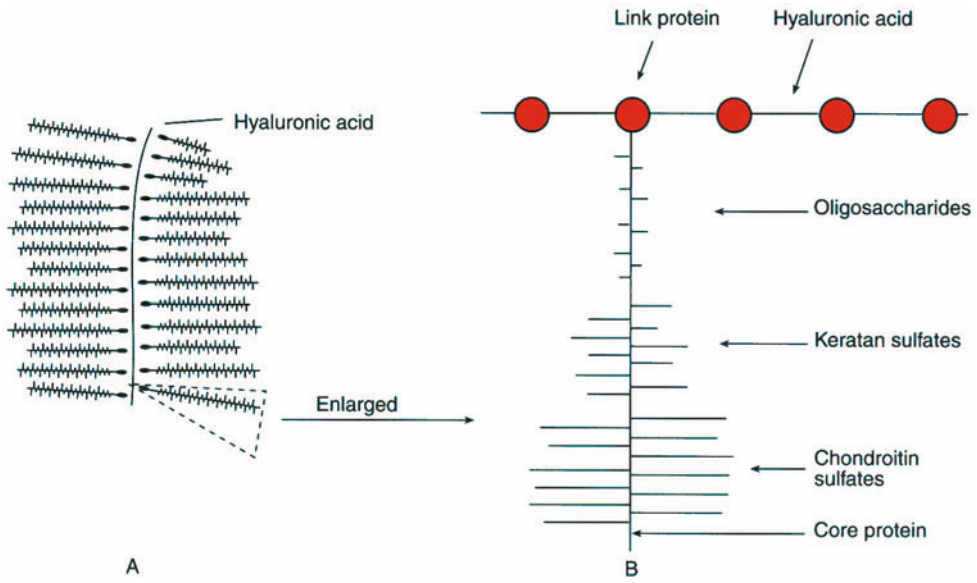


Figure 5.26. Proteoglycan structure in cartilage. (A) The overall brushlike structure, (B) enlarged view of one "bristle."

5.5.4. Lipopolysaccharides

Lipopolysaccharides are water-soluble, lipid–polysaccharide complexes that consist of a heteropolysaccharide chain linked covalently to a *glycolipid* (Section 6.3). Lipopolysaccharides serve as components of the membrane of gram-negative bacteria. The glycolipid (*lipid A*) consists of a disaccharide substituted with long-chain fatty acids. The heteropolysaccharide consists of a core polysaccharide (similar or identical in closely related bacteria) and an O-specific chain that determines the identity of cell surface antigens. We can represent the overall structure of a lipopolysaccharide as

[glycolipid]—[heteropolysaccharide]

or

[lipid A]—[core polysaccharide—O-specific chain]

5.5.5. Cell Surface Structures

5.5.5A. Plants. Cell walls constitute the rigid structures, external to the cell membranes, that enclose plant and prokaryotic cells. In the case of plants, cell walls achieve their strength from having long, tough fibers held together by a matrix of protein and polysaccharide. The fibers and matrix are cross-linked by covalent bonds and also held together by noncovalent interactions. The underlying structural principles resemble those found in such common building materials as fiberglass (glass fibers embedded in a resin) and reinforced concrete (steel rods embedded in concrete).

The long, tough fibers consist of cellulose, and the carbohydrate components of the matrix consist primarily of hemicellulose and pectin. **Hemicellulose** denotes a heterogeneous group of branched polysaccharides, each of which has a long linear backbone of D-*xylose* in β(1→4) linkage from which protrude short side chains of other sugars. **Pectin** denotes a heterogeneous group of polysaccharides consisting of polymers of D-*galacturonic acid* in α(1→4) linkage and variously methylated at C(6) to form methyl esters. Pectins are highly hydrated polymers, abundant in fruits, and used as gelling agents. The major nonpolysaccharide component of plant cell walls is *lignin,* a polymer of the phenolic compound coniferyl alcohol.

5.5.5B. Prokaryotes. Prokaryotic cells have either of two types of cell wall, depending on whether the cell is *gram-positive* or *gram-negative.* The capacity of cells to retain or lose a specific dye, when subjected to the *Gram stain* procedure, reflects fundamental differences in cell wall structure.

In both gram-positive and gram-negative organisms, **peptidoglycan,** also called *murein* (from the Latin, meaning "wall"), serves as the basic structural framework. Peptidoglycan consists of a network of polysaccharides and peptides. Repeating disaccharide units, composed of *N-acetylglucosamine (NAG)* linked β(1→4) to *N-acetylmuramic acid (NAM)* (Figure 5.10), form the polysaccharide chains. A tetrapeptide, *contain-*

ing D-*amino acids,* is linked to the lactic acid residue of muramic acid to form the peptidoglycan repeating unit (Figure 5.27). Adjacent chains of peptidoglycan are then cross-linked by means of their tetrapeptide side chains.

In the gram-positive bacterium *Staphylococcus aureus, pentaglycine connectors* link the tetrapeptide side chains. Each connector forms two peptide bonds, one with the α-COOH group of alanine at the C-terminus of one tetrapeptide, and the other with the ϵ-NH_2 group of lysine of a second tetrapeptide (Figure 5.28).

Gram-positive organisms have a thick and rigid cell wall (about 250 Å) that consists of many peptidoglycan layers, probably cross-linked additionally in three dimensions. Action of the enzyme *lysozyme* leads to cell wall digestion. Lysozyme, found in tears, secretions, mucus, and eggs, catalyzes hydrolysis of the β(1→4) glycosidic link between NAG and NAM; hydrolysis of this bond causes the cell wall to "dissolve."

Gram-negative organisms have a much thinner cell wall (about 30 Å). The wall has the same polysaccharide structure, but peptide chains and their linkages differ slightly. In these organisms, an outer membrane consisting of lipopolysaccharide, protein, and lipoprotein covers a thin peptidoglycan layer. The space between the cell wall and the cell membrane forms an aqueous compartment, called *periplasmic space* (Figure 5.29). Lysozyme does not cause digestion of these cell walls.

The antibiotic *penicillin,* produced by the mold *Penicillium notatum,* inhibits bacterial cell wall synthesis by specifically binding to and inhibiting enzymes that catalyze cross-linking of peptidoglycan strands. Frequently, a slimy, polysaccharide-rich layer, called a *capsule,* covers bacterial cell walls. Composition of the capsule varies, and

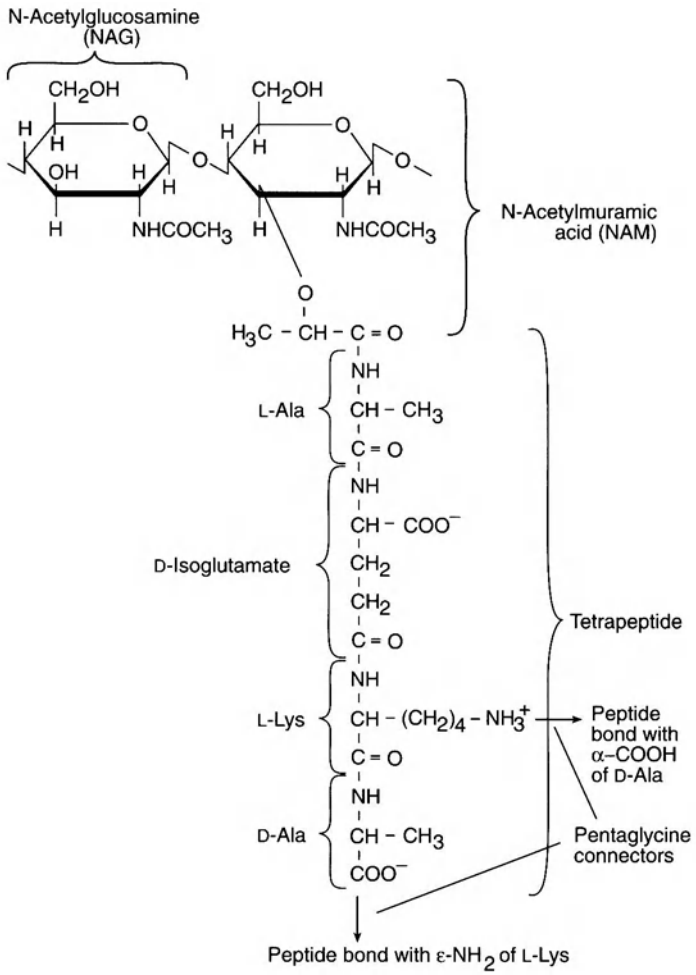


Figure 5.27. The repeating unit of peptidoglycan. Units are joined by means of pentaglycine connectors that cross-link tetrapeptides in repeating units. Each pentaglycine connector forms two peptide bonds, linking two tetrapeptides.

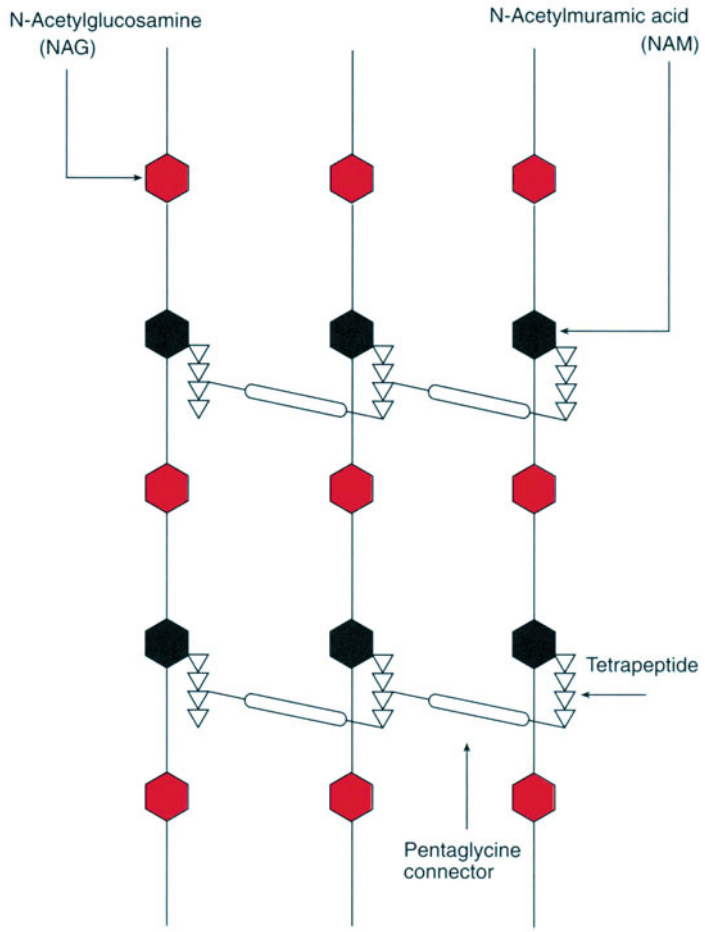


Figure 5.28. Diagram of the peptidoglycan network in the gram-positive bacterium *Staphylococcus aureus*. The network extends up and down to generate a three-dimensional structure. Gram-negative bacteria have a network that differs slightly in the nature and linkage of the peptides.

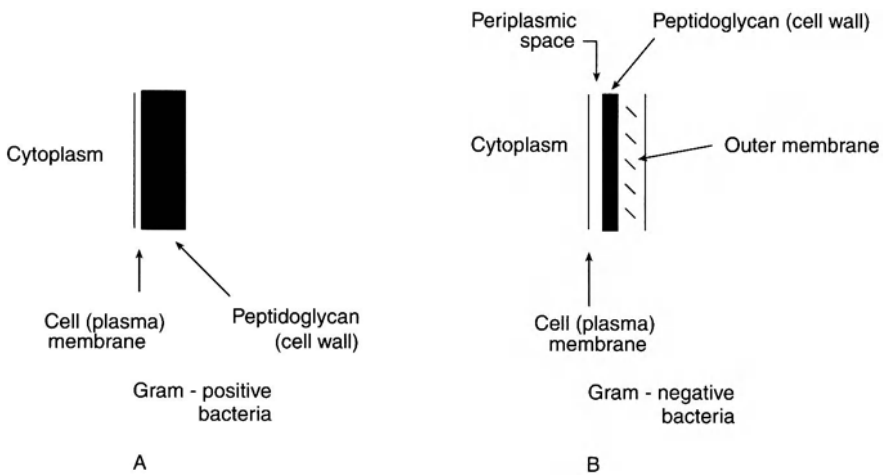


Figure 5.29. Structure of the cell envelope of gram-positive (A) and gram-negative (B) bacteria.

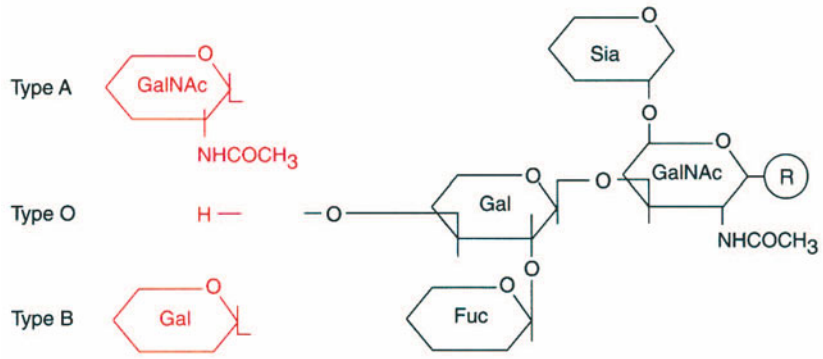


Figure 5.30. Nonreducing ends of the antigens of the ABO blood group system. Gal = galactose, GalNAc = *N*-acetylgalactosamine, Fuc = fucose (6-deoxy-L-galactose), Sia = sialic acid, R = protein or lipid.

its presence often correlates with bacterial pathogenicity. Capsules may function in cell–cell recognition, affect adhesion to surfaces, and serve as permeability barriers.

5.5.5C. Eukaryotes. Eukaryotic cells, other than plants, do not possess a cell wall; however, a *cell coat,* called *glycocalyx,* frequently covers the cell membrane. The cell coat consists of a slimy layer, rich in glycoproteins and glycosaminoglycans, that contributes to the specificity of cells and to the nature of contact inhibition. **Contact inhibition** refers to the inhibition of cell growth that occurs when cells of multicellular organisms come into close contact with one another. It constitutes a cell–cell interaction that prevents disorderly pileup of cells. Loss of contact inhibition is one of the characteristics of cancer cells.

5.5.5D. Red Blood Cells. The surface of *erythrocytes,* human red blood cells, is covered by a network of specific antigens, called *blood group substances.* Some of these are O-linked glycans, joined covalently to membrane proteins. Others are glycolipids—oligosaccharides linked to lipids. The lipid portion aids in attaching the antigen to the cell membrane. The carbohydrate moiety of blood group substances determines the blood group type. Scientists have identified some 100 blood group substances. These belong to 15 blood group systems, of which the best known are the *Rh system* and the *ABO system.*

The antigen of the Rh blood group system is the Rh factor, so called since it was first discovered in the rhesus monkey. Individuals are either Rh positive or Rh negative. An Rh-positive baby of an Rh-negative mother may be born with a hemolytic disease called *erythroblastosis fetalis.*

In the ABO blood group system, there are two antigens, A and B, whose *antigenic determinants* consist of identical oligosaccharides except for their nonreducing ends (Figure 5.30). The antigenic determinant represents that portion of an antigen molecule responsible for the specificity of the antigen; it combines with the antibody. The two antigens give rise to four serum groups denoted A, B, AB, and O. These are characterized, respectively, by carrying antigen A, B, A and B, and neither A nor B.

SUMMARY

Chemically, carbohydrates are aldehyde or ketone derivatives of polyhydroxy alcohols. Carbohydrates serve as a primary source of energy for metabolism, as a source of carbons for synthesis of other biomolecules, as a storage form of energy, as structural elements, and as providers of recognition factors for proteins. Monosaccharides, or simple sugars, typically contain three to seven carbon atoms. Derivatives of simple sugars include sugar acids, sugar alcohols, amino sugars, deoxysugars, and phosphoric acid esters.

Carbohydrates undergo an intramolecular reaction that results in formation of a cyclic structure (hemiacetal or hemiketal) and in conversion of the carbonyl carbon (anomeric carbon) to a chiral center. The configu-

ration of substituents at the anomeric carbon yields two optical isomers (anomers). In solution, many carbohydrates exist predominantly in the form of ring structures, in equilibrium with the open-chain form. Existence of this equilibrium accounts for mutarotation and endows ring structures with potential aldehyde groups. Mono- and oligosaccharides having potential aldehyde groups are reducing sugars.

Carbohydrates undergo various chemical reactions such as oxidation–reduction, fermentation, esterification, and etherification. An analytical technique called exhaustive methylation aids in determining structural details of oligo- and polysaccharides.

In oligo- and polysaccharides, monosaccharide units are linked via bonds involving the anomeric carbon (glycosidic bonds). Two common types are $\alpha(1\rightarrow4)$ and $\beta(1\rightarrow4)$ glycosidic bonds. Amylose and cellulose consist of linear polymers composed entirely of glucose residues. They differ only in that the units are linked by an $\alpha(1\rightarrow4)$ bond in amylose and by a $\beta(1\rightarrow4)$ bond in cellulose. We describe oligo- and polysaccharides by the type, number, and sequence of monosaccharide units and the types and locations of glycosidic bonds between the units, including those at branch points.

Starch and glycogen are the major storage polysaccharides, and cellulose and chitin are the major structural polysaccharides. Important classes of complex polysaccharides include glycosaminoglycans, glycoproteins, proteoglycans, and lipopolysaccharides. Carbohydrates occur as structural components of the cell wall of both gram-positive and gram-negative bacteria and of the cell coat of eukaryotic cells.

SELECTED READINGS

Binkley, R. W., *Modern Carbohydrate Chemistry,* Marcel Dekker, New York (1988).

Chaplin, M. F., and Kennedy, J. F. (eds.), *Carbohydrate Analysis,* 2nd ed., Oxford University Press, Oxford (1994).

Cold Spring Harbor Symposia on Quantitative Biology, *The Cell Surface,* Vol. LVII, Cold Spring Harbor Laboratory Press, New York (1992).

Driouich, A., Faye, L., and Staehelin, L. A., The plant Golgi apparatus: A factory for complex polysaccharides and glycoproteins, *Trends Biochem. Sci.* 18:210–214 (1993).

Dwek, R. A., Edge, C. J., Harvey, D. J., and Wormald, M. R., Analysis of glycoprotein-associated oligosaccharides, *Annu. Rev. Biochem.* 62:65–100 (1993).

Fukuda, M., *Cell Surface Carbohydrates and Cell Development,* CRC Press, Boca Raton, Florida (1991).

Hounsell, E. F., Physicochemical analyses of oligosaccharide determinants of glycoproteins, *Adv. Carbohydr. Chem.* 50:311–350 (1994).

Manzella, S. M., Hooper, L. V., and Baenziger, J. U., Oligosaccharides containing β1,4-linked *N*-acetylgalactosamine, a paradigm for protein-specific glycosylation, *J. Biol. Chem.* 271:12117–12120 (1996).

Shao, M. C., and Wold, F., The effect of protein matrix proximity on glycan reactivity in a glycoprotein model, *Eur. J. Biochem.* 228:79–85 (1995).

Sharon, N., and Lis, H., Carbohydrates in cell recognition, *Sci. Am.* 268:82–89 (1993).

Thiem, J. (ed.), *Carbohydrate Chemistry,* Springer-Verlag, New York (1990).

Zähringer, U., Lindner, B., and Rietschel, E. T., Molecular structure of lipid A, the endotoxic center of bacterial lipopolysaccharides, *Adv. Carbohydr. Chem.* 50:211–276 (1994).

REVIEW QUESTIONS

A. Define each of the following terms:

Mutarotation	Glycosidic bond
Reducing sugar	Nonreducing end
Lactose intolerance	Anomers
Peptidoglycan	Sugar alcohol
Potential aldehyde group	Contact inhibition
Sugar acid	Glycosylation
Haworth projection	Exhaustive methylation

B. Differentiate between the two terms in each of the following pairs:

Aldose/ketose	Cellulose/hemicellulose
α-Amylose/amylopectin	Dextrin/dextran
Proteoglycan/glycoprotein	Starch/glycogen
Anomeric carbon/ epimeric carbon	Homopolysaccharide/ heteropolysaccharide

C. (1) Why do carbohydrates that exist predominantly in the form of ring structures in solution readily undergo reactions characteristic of their open-chain formulations?
(2) Describe the basic structural framework of the bacterial cell wall, and compare and contrast the structure of the cell envelope in gram-positive organisms with that in gram-negative organisms.
(3) Compare and contrast the structures of maltose, lactose, and sucrose.
(4) How are carbohydrates linked covalently to protein?
(5) What are the structures, properties, and functions of glycoproteins and proteoglycans?

PROBLEMS

5.1. What products would be obtained (type and number) if the amylopectin fragment shown in Figure 5.19 were subjected to exhaustive methylation followed by gentle enzymatic hydrolysis?

5.2. Calculate the initial and final values for the specific rotation of a solution containing 20% α-D-glucose and 80% β-D-glucose.

5.3. Based on the linkages between residues, how many different types of glucosyl residues occur in (a) cellulose and (b) amylopectin?

5.4. What weight of silver, in grams, would be deposited when a solution containing 18.0 g of glucose is treated with Tollens' reagent? (MW of glucose = 180; at. wt. of silver = 108).

5.5.* What is the theoretical specific rotation of invert sugar in the absence of mutarotation? The specific rotation of β-D-fructose is −134°.

5.6. Why are gram-negative bacteria not susceptible to lysozyme-catalyzed degradation?

5.7.* The specific rotation of α-D-fructose is −21°, that of β-D-fructose is −134°, and that obtained at equilibrium is −92°. Calculate the percentages of the α- and β-anomers present at equilibrium.

5.8. Write the Fischer projections for: (a) *N*-acetyl-D-2-galactosamine; (b) galacturonic acid; (c) 3-deoxy-D-galactose.

5.9. How does the repeating unit of starch differ from that of cellulose?

5.10. Periodate (IO_4^-) oxidation constitutes an important analytical tool in carbohydrate chemistry. Periodate cleaves the bond between two adjacent carbon atoms that carry carbonyl, hydroxyl, or amino groups. On that basis, predict the number of moles of periodate that will react with one mole of each of the following: (a) glucose 6-phosphate; (b) 2-deoxyglucose; (c) 2-glucosamine.

5.11. Theoretically, how many different disaccharides of glucose can there be?

5.12. What is the minimum molecular weight amylose must have for it to give the characteristic blue color of the starch–iodine complex? (MW of glucose = 180).

5.13. A student treats aliquots of an amylose solution with Tollens' and Benedict's reagents. Both tests are negative. How can you explain this?

5.14. By what factor would monellin be sweeter than sucrose if you compute sweetness on a molar basis? (MW of sucrose = 342; assume that MW of any amino acid residue is 110.)

5.15.* A sample of amylose (3.6 g) is exhaustively methylated and then carefully hydrolyzed by means of *enzymes* so that no methyl groups are lost from anomeric carbons. The experiment yields a total of 2.0 millimoles of 1,2,3,6-tetramethylglucose. Calculate the minimum number of glucose residues per chain of amylose. Use 180 as the molecular weight of any glucose residue.

5.16.* A student subjects a sample of amylopectin (5.4 g) to exhaustive methylation, followed by hydrolysis. The products include 3.0 millimoles of 2,3-dimethylglucose. Calculate the percent of glucose residues located at α(1→6) branch points. Use 180 as the molecular weight of any glucose residue.

5.17. Indicate which of the following pairs of monosaccharides are (A) anomers, (B) epimers, (C) enantiomers, (D) diastereomers, (E) aldehyde–ketone isomers:

1. D-Ribose and L-ribose
2. D-Erythrose and D-erythrulose
3. D-Allose and D-glucose
4. D-Mannose and D-galactose
5. α-D-Xylose and β-D-xylose
6. D-Glucose and D-mannose

5.18. Draw the Haworth projections for:

(a) *O*-α-D-galactopyranosyl-(1→4)-α-D-galactopyranose
(b) *O*-β-D-glucopyranosyl-(1→3)-β-D-mannopyranose
(c) *O*-α-D-ribofuranosyl-(1→2)-α-D-fructofuranoside

5.19. Why is penicillin effective only against *growing* cells?

5.20. How many chiral centers (asymmetric carbons) are there in (a) glucuronic acid, (b) ascorbic acid, and (c) neuraminic acid?

5.21.* Consider two identical amylose fragments, each consisting of 100 glucose residues. You subject the

fragments to hydrolysis, cleaving only three glycosidic bonds in each fragment. You intend to produce minimum degradation with one fragment (a) and maximum degradation with the other (b). What are the final sizes of the fragments obtained and what is the average chain length in cases (a) and (b)?

5.22. D-Galactose shows optical activity, but after reduction with lithium borohydride ($LiBH_4$) you find it to be optically inactive. Why is this so?

5.23.* Hydrolysis of the carbohydrate trehalose yields D-glucose as the only product. Exhaustive methylation of trehalose, followed by *gentle enzymatic hydrolysis,* yields only one product: 2,3,4,6-tetramethyl-α-D-glucose. What is the structure of trehalose? (Give the systematic name and draw the Haworth projection.)

5.24.* A student determines that a sample (0.90 g) of a linear polysaccharide contains 2.0 micromoles of glucose and 3.0 micromoles of ribose as reducing ends. What is the minimum molecular weight of the average chain length?

5.25. Write out the Haworth projection for *N*-acetylglucosamine linked covalently via a β-glycosidic bond to the amide group of asparagine.

5.26. Is there another pentasaccharide of α-D-glucose that, when exhaustively methylated and hydrolyzed, will yield the same experimental results as those shown in Figure 5.13? If so, write out its structure.

5.27. The hydroxyl groups in β-D-glucopyranose are all linked via equatorial bonds. How does this affect the structural stability of cellulose?

Lipids and Membranes

6

Lipids (from the Greek, meaning "fat") comprise a heterogeneous group of organic compounds, insoluble (or sparingly soluble) in water, but soluble in nonpolar solvents such as chloroform, ether, and benzene. Lipids differ in their structure, but all have pronounced nonpolar groups, or both nonpolar and polar groups. Thus, lipids are either *hydrophobic* or *amphipathic* compounds. Unlike proteins and polysaccharides, lipids are small molecules, but they have a strong tendency to associate through noncovalent interactions. As major parts of their structure, many lipids have long hydrocarbon chains that may be saturated or unsaturated. Some lipids have specific roles and high biological activity. Among these, we find several vitamins and hormones. In addition, lipids have a number of general biological functions.

Lipid deposits, particularly in the form of fats, constitute the primary *storage form of energy* in animals. Lipids have the highest energy value of all foodstuffs, making them ideal for energy storage. To store an equivalent amount of energy in the form of either carbohydrate or protein requires deposition of much larger quantities of material. An additional advantage is that lipids are stored with little water, in contrast to storage of carbohydrates. Lipid degradation produces about 38 kJ/g (9 kcal/g).

Lipids serve also as *structural components of biological membranes,* the dynamic permeability barriers that regulate the movement of material into, and discharge of substances from, cells and subcellular structures. Membranes are composed primarily of lipids, particularly phospholipids and glycolipids.

Lastly, lipids have *several protective roles.* In animals, a surrounding layer of fat protects many internal organs against mechanical injury. Lipids are poor conductors and provide thermal insulation, particularly for animals that live in a cold environment and have thick layers of fat under their skin. Lipids also form protective coatings of feathers, skin, and fruits that minimize the chance of infection and prevent excessive water loss.

We can divide lipids into three categories: **simple lipids** are esters of fatty acids and alcohols; **complex lipids** contain a lipid moiety plus some nonlipid component; and **derived lipids** comprise a group of diverse com-

pounds including prostaglandins, isoprenoids, fat-soluble vitamins, and steroids. We will begin our study of lipids with a look at fatty acids.

6.1. FATTY ACIDS

We define **fatty acids** as long-chain carboxylic acids that occur in living organisms (Figure 6.1). Fatty acids constitute parts of many lipids and have the formula R—COOH, where R is a hydrocarbon chain or *tail;* the chain is generally straight and unbranched, and it may be saturated or unsaturated. Small amounts of branched, and even cyclic, fatty acids also occur in nature. **Saturated fatty acids** have a saturated hydrocarbon chain.

We call fatty acids "even-numbered" or "odd-numbered," depending on the *total* number of carbon atoms in the molecule. Most naturally occurring fatty acids are *even-numbered* (Table 6.1), but some fatty acids have an odd number of carbons.

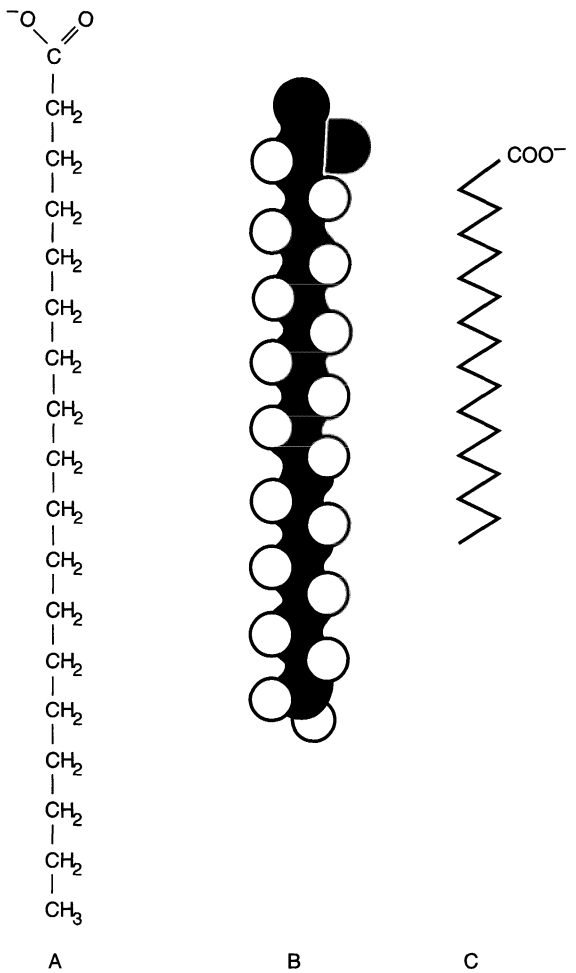


Figure 6.1. A saturated fatty acid. (A) Structural formula of stearic acid; (B) space-filling model; (C) schematic representation in which a carbon atom occurs at each vertex.

Unsaturated fatty acids contain one or more double bonds in their hydrocarbon chain and, hence, can exist in the form of geometric isomers, having either a *cis* or a *trans* configuration. In most naturally occurring fatty acids, double bonds have the *cis* configuration. Occurrence of a *cis* double bond introduces a bend into the hydrocarbon chain of the molecule (Figure 6.2), making close and regular packing of these tails difficult. Consequently, crystals of unsaturated fatty acids form less readily than those of the corresponding saturated compounds; unsaturated fatty acids have lower melting points than saturated ones, and many are liquid at room temperature.

Unsaturated fatty acids have greater chemical reactivity than their saturated counterparts. Addition reactions to double bonds occur readily. Two such reactions, *hydrogenation* and *iodination,* have particular applications.

Hydrogenation converts unsaturated fatty acids to saturated:

$$\text{Hydrogenation:}\quad -CH{=}CH- + H_2 \xrightarrow{\text{Pt or Pd}} -CH_2-CH_2-$$

Hydrogenation represents the fundamental reaction in producing most types of margarine. Margarine consists largely of **fats,** esters of fatty acids and glycerol. Fats containing many unsaturated fatty acids tend to be liquid at room temperature and occur naturally in the form of vegetable oils like olive oil, peanut oil, and corn oil. Hydrogenation of such oils produces fats that contain higher levels of saturated fatty acids and that tend to be solid at room temperature.

A similar addition reaction occurs in the presence of iodine:

$$\text{Iodination:}\quad -CH{=}CH- + I_2 \longrightarrow \underset{\displaystyle I\quad I}{-CH-CH-}$$

Fat iodination has analytical value. When we treat a given amount of fat with a known amount of iodine, we can determine the remaining unreacted iodine by iodometric titration. From the titration data, the amount of iodine taken up by the fat can be calculated. The greater the unsaturation in the fat, the greater is the amount of iodine taken up. We frequently express this in terms of an **iodine**

Table 6.1. Some Naturally Occurring Fatty Acids

Designation[a]	Common name	Structure	Melting point (°C)
Saturated fatty acids			
12:0	Lauric	$CH_3(CH_2)_{10}COOH$	44.2
14:0	Myristic	$CH_3(CH_2)_{12}COOH$	53.9
16:0	Palmitic	$CH_3(CH_2)_{14}COOH$	63.1
18:0	Stearic	$CH_3(CH_2)_{16}COOH$	69.6
20:0	Arachidic	$CH_3(CH_2)_{18}COOH$	76.5
Unsaturated fatty acids			
16:1	Palmitoleic	$CH_3(CH_2)_5CH{=}CH(CH_2)_7COOH$	−0.5
18:1	Oleic	$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$	13.4
18:2	Linoleic	$CH_3(CH_2)_4CH{=}CHCH_2CH{=}CH(CH_2)_7COOH$	−5
18:3	Linolenic	$CH_3CH_2CH{=}CHCH_2CH{=}CHCH_2CH{=}CH(CH_2)_7COOH$	−11
20:4	Arachidonic	$CH_3(CH_2)_4CH{=}CHCH_2CH{=}CHCH_2CH{=}CHCH_2CH{=}CH(CH_2)_3COOH$	−50

[a]Fatty acids are designated by two numbers, separated by a colon; the first indicates the *total* number of carbon atoms, and the second the number of double bonds. All double bonds of the unsaturated fatty acids shown are *cis*.

number, defined as the number of grams of iodine taken up by 100 g of fat.

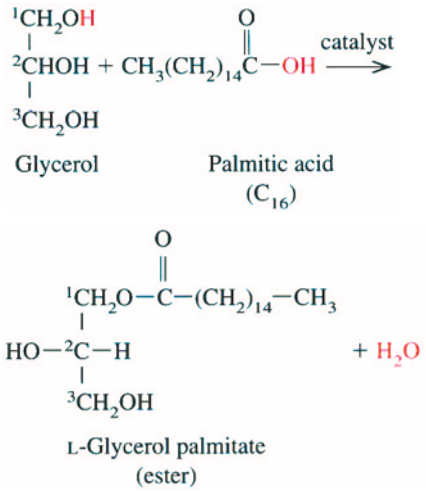
Mammals can synthesize some, but not all, of the unsaturated fatty acids occurring in metabolism. Mammals cannot introduce a double bond beyond C(9)—counting from the carboxyl group. Thus, monounsaturated fatty acids (containing one double bond), such as palmitoleic and oleic acid, can be synthesized, but some polyunsaturated fatty acids (containing more than one double bond), such as linoleic and linolenic acid, cannot be produced and must be obtained through the diet. These polyunsaturated fatty acids constitute **essential fatty acids** in the same sense that some amino acids are essential (Section 2.1). Linoleic acid occurs in plant oils, and linolenic acid in fish oils. Arachidonic acid, another polyunsaturated fatty acid, does not constitute an essential fatty acid; it is synthesized from linoleic acid by elongation and desaturation.

6.2. SIMPLE LIPIDS

Although fatty acids are important components of many lipids, the concentration of free fatty acids in living cells is generally low. Most fatty acids occur as esters that constitute parts of simple and complex lipids. In *simple lipids,* esters form from fatty acids and alcohols. If the alcohol is glycerol, we call the esters *true (neutral) fats,* or simply *fats;* if the alcohol is a long-chain compound, we refer to the esters as **waxes.**

6.2.1. Formation and Hydrolysis of Fats

6.2.1A. Fat Formation. Simple lipids, like all esters, form by elimination of a molecule of water between an acid and an alcohol:



We call esters of glycerol and fatty acids **acylglycerols** but also still use the older term *glycerides.* In the above example, we esterified glycerol once, forming a *monoacylglycerol* or *monoglyceride.* A different monoacylglycerol forms from the same reactants by esterifying palmitic acid to an OH at C(2) or C(3) of glycerol.

Glycerol can also undergo esterification at more than one position. Esterifying at two OH groups produces a *di-*

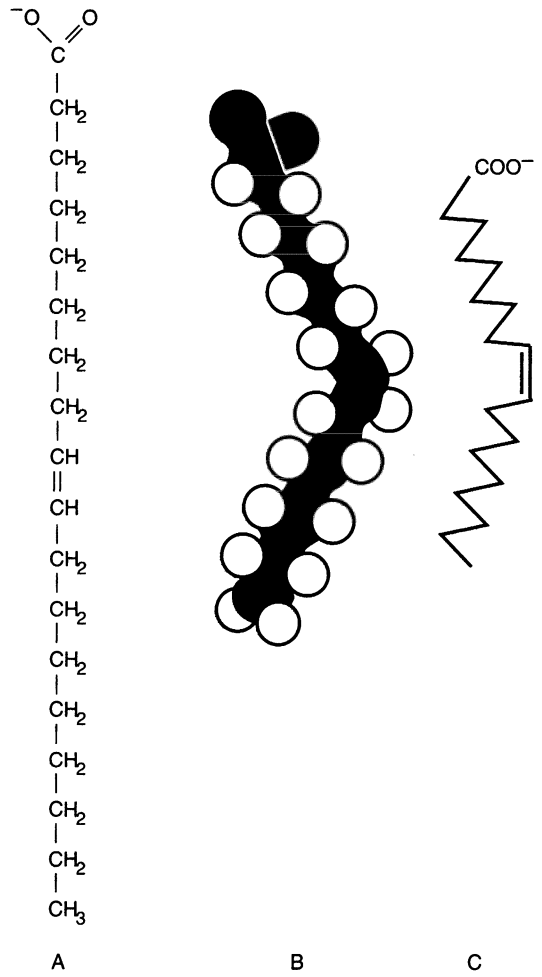


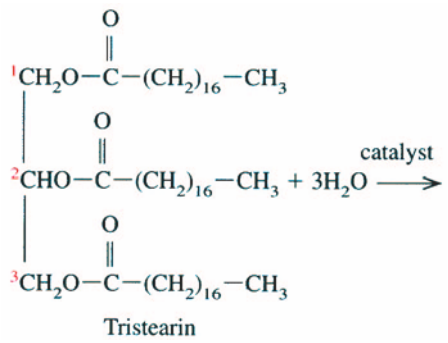
Figure 6.2. An unsaturated fatty acid. (A) Structural formula of oleic acid; (B) space-filling model; (C) schematic representation. Occurrence of a *cis* double bond introduces a rigid bend in the molecule.

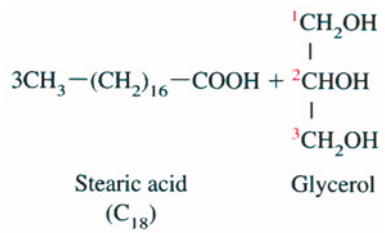
acylglycerol or *diglyceride.* Esterifying at all three OH groups yields a *triacylglycerol* or *triglyceride.* Note that C(2) is a prochiral carbon in glycerol but becomes a chiral center in many acylglycerols, allowing for D- or L-configurations of fats. Of these, the L-isomer is biologically more important.

Esterifying more than one hydroxyl group may involve the same or different fatty acids. We call fats containing nonidentical fatty acids *mixed* di- or triacylglycerols. Fatty acids attached to a single glycerol molecule can differ not only in chain length but also in branching and degree of saturation.

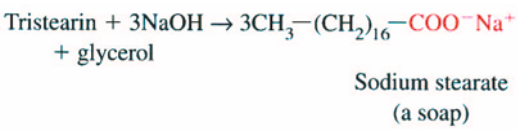
Naturally occurring fats represent mixtures of acylglycerols. Variations in the properties of sheep fat, pig fat, human fat, and the like result from differences in both the *types and relative amounts of acylglycerols.* The greater the percentage of unsaturated fatty acids in a fat, the greater is the likelihood that the fat exists as a liquid at room temperature. We call a liquid fat an **oil.** Well-known vegetable oils, like olive oil, peanut oil, sesame oil, and corn oil, comprise mixtures of acylglycerols that contain a high proportion of unsaturated fatty acids. Coconut oil constitutes an exception in that it actually contains about 90% saturated fatty acids.

6.2.1B. Hydrolysis of Fats. Fats can be hydrolyzed by means of acid, steam, or enzymes. Hydrolysis represents the reverse of fat formation:





We can also hydrolyze fats by using a base like sodium hydroxide. In that case, we obtain as end product not the free fatty acid but rather its sodium salt. For tristearin, the product is sodium stearate:



Scientists refer to base-catalyzed hydrolysis of fats as **saponification** and to salts of fatty acids as **soaps.** They term sodium salts of fatty acids *hard soaps,* and potassium salts *soft soaps.* Hard soaps are lesss soluble in water than soft soaps. You can perform saponification analytically, using known quantities of fat and base. After saponification, you determine the amount of unreacted base by titration with standard acid. Hence you calculate the amount of base required to saponify a given amount of fat. We usually express such data in terms of a **saponification number,** defined as the number of milligrams of KOH re-

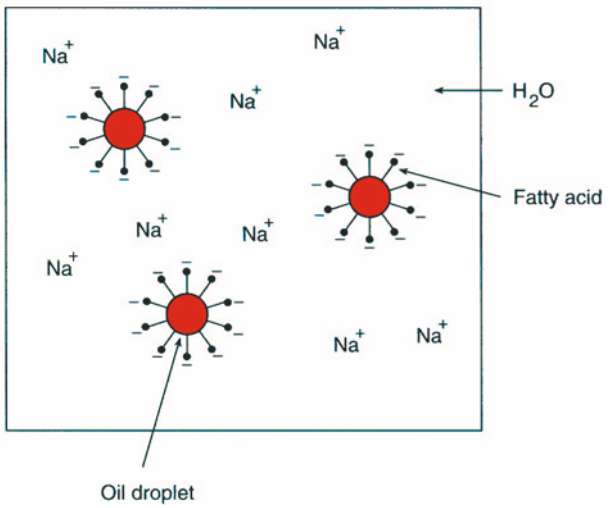


Figure 6.3. Stabilization of emulsions. Soap molecules insert into and coat oil droplets. Positive counterions occur throughout the solution but concentrate around the droplets.

quired to saponify one gram of fat. The shorter the fatty acid chains are, the larger the number of fat molecules in a one-gram sample of fat, and the larger the number of KOH molecules required to saponify this fat. Hence, the greater the saponification number is, the shorter the average chain length of the fatty acids in the fat. We commonly divide lipids into two large groups, *saponifiable* and *nonsaponifiable,* depending on whether or not alkaline hydrolysis produces one or more soap molecules.

6.2.2. Micelles, Emulsions, and Soap Action

As discussed in Section 1.4, soap molecules are *amphipathic,* having both pronounced polar and nonpolar parts. When dissolved in water above a certain concentration (the *critical micelle concentration* or *CMC*), they form supramolecular aggregates called *micelles.*

6.2.2A. Emulsions. The principle that underlies micelle formation—hydrophobic interactions—also applies to stabilization of emulsions by soap molecules. We refer to a colloidal dispersion of one liquid in another, immiscible or partially miscible, as an **emulsion.** Milk, butter, mayonnaise, and cold cream represent different types of emulsions. To illustrate emulsion stabilization, consider the dispersion of oil in water.

When you stir a mixture of oil and water, oil droplets disperse temporarily, but upon standing they coalesce to form a layer floating on the water. If you repeat the experiment with soap present, soap molecules insert themselves into the oil droplets via their long nonpolar hydrocarbon chains (Figure 6.3). Oil droplets thereby acquire a negatively charged coating composed of $-COO^-$ groups. The structure resembles a micelle but with a central core of oil.

Having attained negative surface charges, oil droplets repel each other and do not coalesce. Additionally, water binds to the $-COO^-$ groups at the surface of each droplet. Hydration keeps droplets apart, much as it serves to dissolve simple ions. Both electrical repulsion and hydration combine to keep the oil droplets dispersed in the aqueous phase, resulting in a stabilized emulsion.

6.2.2B. Soap Action and Detergents. Similar hydrophobic interactions to those involved in forming micelles and stabilizing emulsions occur when you use soap for cleaning (Figure 6.4). As you rub soap over a dirty surface, soap molecules insert themselves in the layer of grease and dirt via their nonpolar hydrocarbon chains. The concomitant abrasive rubbing helps break the layer into small micelle-like fragments coated with soap molecules. By passing a stream of water over the surface you rinse off the fragments, leaving a clean surface behind.

Once industry understood the principle of soap action, it proceeded to synthesize amphipathic compounds that could mimic the action of soap molecules. Common household detergents fall into this category. A typical detergent is *sodium lauryl (dodecyl) sulfate (SDS),* a synthetic analog of the soap sodium laurate:

Sodium dodecyl sulfate	$CH_3(CH_2)_{10}CH_2{-}OSO_3^-Na^+$
Sodium laurate	$CH_3(CH_2)_{10}COO^-Na^+$

Detergents function in precisely the same way as soaps. Using detergents allows cleansing in *hard water,* in

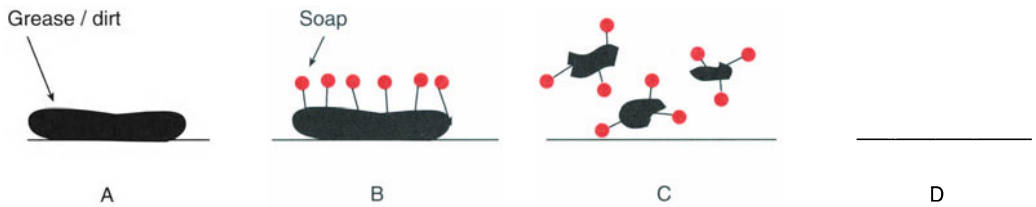


Figure 6.4. The cleansing action of soap. A layer of grease and dirt (A) becomes coated with soap molecules (B), is dispersed as micelle-like fragments (C), and is rinsed off (D).

which ordinary soaps are ineffective because they form insoluble salts. Hard water contains appreciable concentrations of Ca^{2+}, Mg^{2+}, and Fe^{3+} ions, which displace the Na^+ and K^+ ions of ordinary soaps. The insoluble Ca^{2+}, Mg^{2+}, and Fe^{3+} salts of fatty acids precipitate out and float (being less dense than water) to form the scum you get when you use soap in hard water. Detergents, on the other hand, form soluble Ca^{2+}, Mg^{2+}, and Fe^{3+} salts and function effectively under the same conditions.

Following the introduction of detergents, industry synthesized branched compounds that were even more effective than the linear types. However, branched detergents were not *biodegradable;* microorganisms in the environment could not break them down. The use of branched detergents resulted in serious pollution of large bodies of water, and, in 1965, their production was prohibited in the United States.

Phosphate is often mentioned in conjunction with detergents. Phosphate is not a detergent *per se* but rather a "detergent builder" that facilitates detergent action. Phosphate, when added to hard water, chelates Ca^{2+}, Mg^{2+}, and Fe^{3+} ions, thereby increasing the efficiency of soaps and detergents.

6.3. COMPLEX LIPIDS

A *complex lipid* consists of a lipid moiety and a nonlipid component that may be a small molecule or a macromolecule. The combination of lipid and nonlipid components in complex lipids resembles the combination of protein and nonprotein components in conjugated proteins. Major classes of complex lipids include glycerophospholipids, sphingophospholipids, glycolipids, lipoproteins, and proteolipids. Phosphate-containing lipids, or **phospholipids,** play important roles in biological membranes.

6.3.1. Glycerophospholipids

Phosphatidic acid (Figure 6.5) represents the parent compound of **glycerophospholipids** or *phosphoglycerides.*

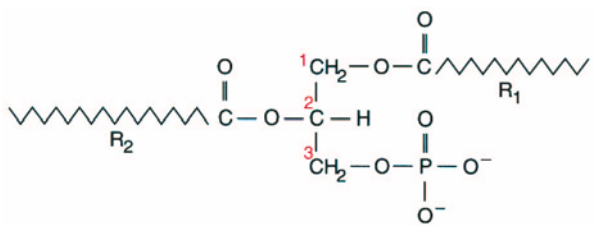
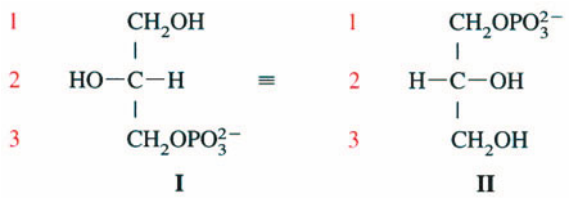


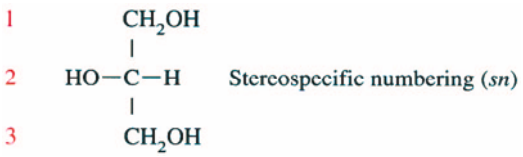
Figure 6.5. Phosphatidic acid.

Phosphatidic acid is derived from glycerol by esterifying two fatty acids at C(1) and C(2) and esterifying phosphoric acid at C(3).

Writing Fischer projections for derivatives of symmetrical compounds such as glycerol can be confusing. For example, L-glycerol 3-phosphate (**I**) and D-glycerol 1-phosphate (**II**) appear to be two different compounds but are, in reality, structurally identical:

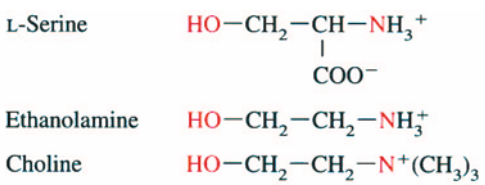


To avoid confusion, the International Union of Biochemistry has adopted the convention to always write the structure of glycerol, by reference to L-glyceraldehyde, with the hydroxyl group to the left of C(2). The carbon above it is C(1); the one below, C(3). The numbers 1 and 3 *cannot be used interchangeably* for the same primary alcohol group. We indicate this convention, called **stereospecific numbering,** by using the prefix *sn:*



On this basis, we call compound **I**, from which phosphatidic acid derives, *sn*-glycerol 3-phosphate. Esterifying compound **I** with two fatty acids, at C(1) and C(2), yields 3-*sn*-phosphatidic acid, the parent compound of glycerophospholipids.

Glycerophospholipids form from phosphatidic acid by esterification of the phosphoric acid residue. In glycerophospholipids, phosphoric acid becomes esterified twice, first to form 3-*sn*-phosphatidic acid, and then by reaction with one of a number of alcohols. Three common ones are:



These three alcohols show some structural similarity in that two carbon atoms separate the OH group from a pos-

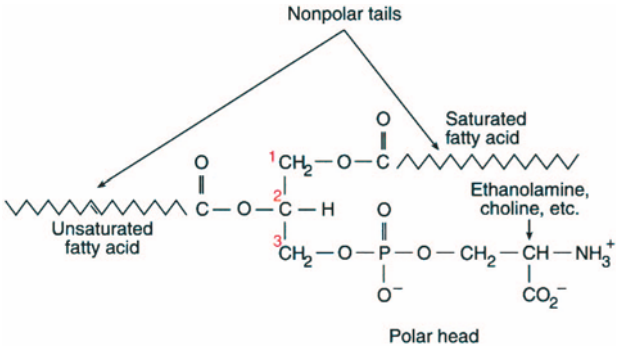


Figure 6.6. Glycerophospholipids. The structural formula of phosphatidyl serine is shown.

itively charged nitrogen. When serine, ethanolamine, and choline become esterified via their hydroxyl groups to phosphatidic acid, they form, respectively, *phosphatidyl serine, phosphatidyl ethanolamine,* and *phosphatidyl choline* (Figure 6.6). We refer to phosphatidyl serine and phosphatidyl ethanolamine, singly or collectively, as *cephalin.* Phosphatidyl choline, or *lecithin,* is sold in health food stores. Still a different glycerophospholipid, *phosphatidyl inositol,* forms by esterification of *myo*-inositol (Figure 5.9) to phosphatidic acid. Note that there exist many different glycerophospholipids of any given kind (e.g., phosphatidyl serine), depending on the fatty acids esterified.

In glycerophospholipids, the second esterification of the phosphoric acid residue makes that portion of the molecule strongly polar. The other part, containing the two esterified fatty acids, is strongly nonpolar. Thus, glycerophospholipids constitute amphipathic molecules, having both pronounced polar and nonpolar sections; we say that they have a polar *head* and two nonpolar *tails.* Glycerophospholipids represent saponifiable lipids, yielding two soap molecules per glycerophospholipid.

6.3.2. Sphingophospholipids

The parent compound of sphingophospholipids is *sphingosine,* a long-chain alcohol (Figure 6.7A). Linking a fatty acid molecule to the amino group of sphingosine yields a *ceramide* (Figure 6.7B). Ceramides occur in only small amounts in plant and animal tissues. **Sphingophospholipids** (Figure 6.7C) form from ceramides by a double esterification of phosphoric acid: to the primary alcohol group of sphingosine and to another alcohol, typically L-serine, ethanolamine, or choline—the same compounds found in glycerophospholipids. Sphingophospholipids are saponifiable lipids, abundant in brain and nervous tissue. They have amphipathic character, containing a polar head and two nonpolar tails, and their overall structure resembles that of glycerophospholipids.

Sphingomyelin (Figure 6.8) is a sphingophospholipid and a key component of the myelin sheath of nerves. This lipid-rich sheath surrounds the axons of nerve cells and promotes rapid transmission of nerve impulses along the nerve. Progressive destruction of the myelin sheath occurs in *multiple sclerosis (MS),* a crippling human disease of unknown cause. Multiple sclerosis is characterized by partial paralysis, weakness, speech and vision problems, and an inability to walk.

6.3.3. Glycolipids

In **glycolipids,** one or more monosaccharide residues are linked via glycosidic bonds to a lipid component such as an acylglycerol or a ceramide. *Cerebrosides,* or glycosyl ceramides (Figure 6.9), are monosaccharide derivatives of ceramides that usually contain glucose or galactose as their sugar. *Gangliosides,* or ceramide oligosaccharides (Figure 6.10), constitute oligosaccharide derivatives of ceramides that contain at least one residue of sialic acid.

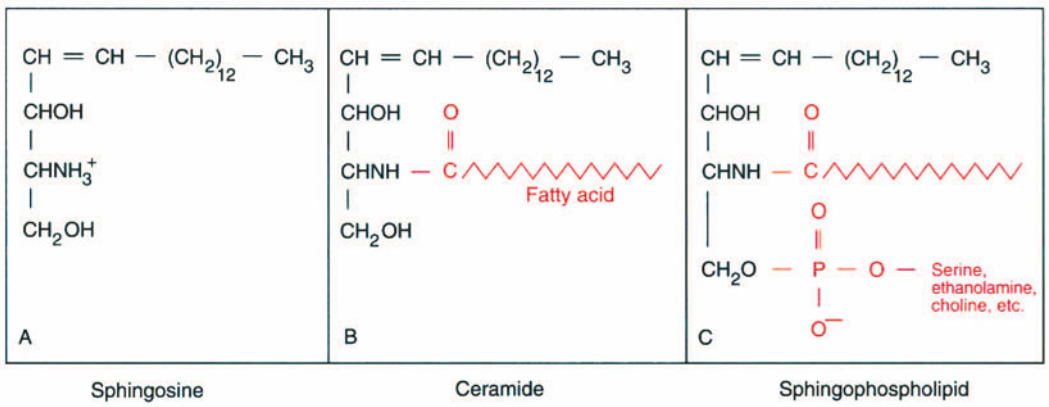


Figure 6.7. Sphingosine and some of its derivatives.

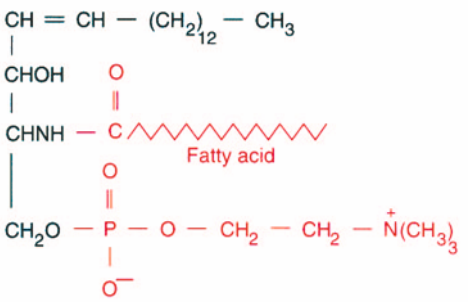


Figure 6.8. Sphingomyelin.

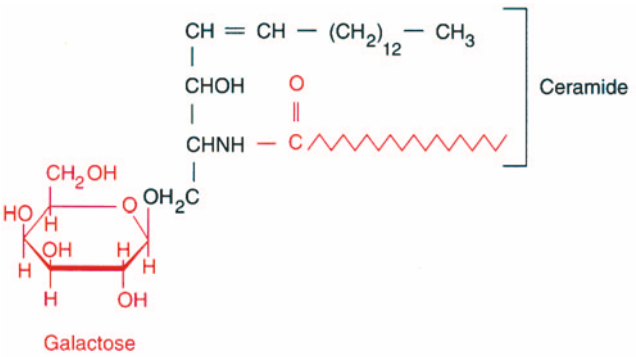


Figure 6.9. Example of a cerebroside.

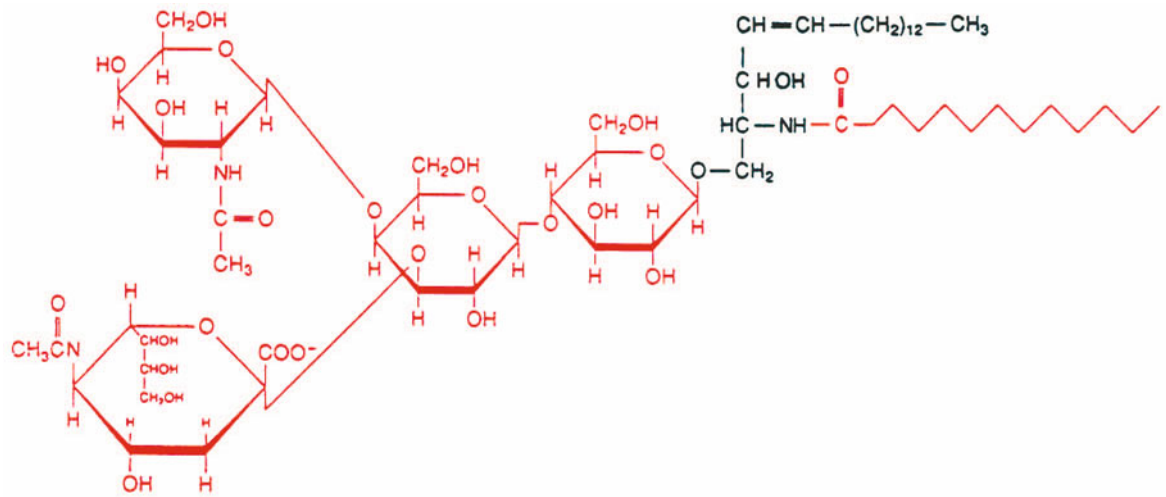


Figure 6.10. *Tay-Sachs ganglioside (GM_2).* GM_2 accumulates in neural tissue of infants suffering from Tay–Sachs disease. Tay–Sachs disease is a genetically inherited disease caused by a deficiency of hexosaminidase A, which catalyzes hydrolysis of a glycosidic bond to release N-acetyl-galactosamine.

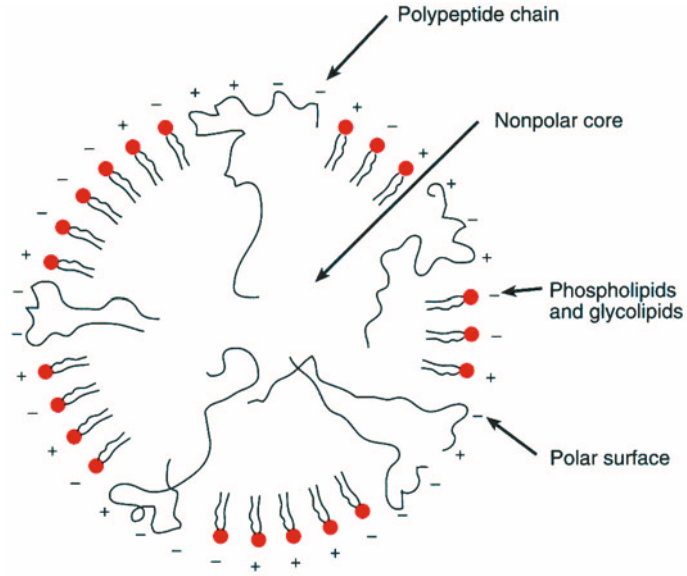


Figure 6.11. Model of a lipoprotein. The globular, micelle-like structure has a hydrophobic core containing triacylglycerols, cholesterol esters, phospholipid tails, and nonpolar segments of polypeptide chains. The core is surrounded by an amphipathic coat containing polar segments of polypeptide chains, phospholipid heads, and unesterified cholesterol.

Cerebrosides and gangliosides occur primarily, but not exclusively, in brain and nervous tissue. Like glycerophospholipids and sphingophospholipids, glycolipids possess a polar head and two nonpolar tails.

6.3.4. Lipoproteins and Proteolipids

Plasma **lipoproteins** comprise water-soluble aggregates of protein and lipid that function as transport vehicles for triacylglycerols and cholesterol. Lipoproteins consist of noncovalently associated proteins and lipids and have a globular, micelle-like structure (Figure 6.11) in which phospholipids and proteins surround a core of cholesterol and triacylglycerols. We classify plasma lipoproteins on the basis of their density, which is a function of their composition (Table 6.2), into five broad categories: *chylomicrons, very-low-density lipoproteins* (*VLDL*), *intermediate-density lipoproteins* (*IDL*), *low-density lipoproteins* (*LDL*), and *high-density lipoproteins* (*HDL*), The LDL particle has a globular structure in which a hydrophobic core of 1500 cholesterol ester molecules is surrounded by an amphipathic coat consisting of 800 phospholipid molecules, 500 cholesterol molecules, and at least one molecule of a large protein, called *apoprotein B-100* (MW = 550,000). We will discuss the role of lipoproteins in atherosclerosis in Section 6.6.

Proteolipids represent a special class of *hydrophobic lipoproteins.* These conjugated proteins have a lipid component and are soluble in some nonpolar solvents but insoluble in aqueous solutions. Biochemists also use the term *proteolipid* to describe a few other proteins that do not contain lipid but are soluble in certain nonpolar solvents because many hydrophobic amino acids cluster at their surfaces. The structure of these proteins differs

Table 6.2. Properties of Lipoproteins[a]

Class	Density (g/ml)	Particle weight $\times 10^{-6}$	Percent (by weight)				
			Protein	Phospholipid	Free cholesterol	Cholesterol esters	Triacylglycerols
Chylomicrons	<0.95	10^3–10^4	1.5–2.5	7–9	1–3	3–5	85–90
VLDL	0.95–1.006	5–10	5–10	15–20	5–10	10–15	50–65
IDL	1.006–1.019	3.5–4.5	15–20	20	7–10	22	20–30
LDL	1.019–1.063	2.0–2.5	20–25	15–20	7–10	35–40	7–10
HDL	1.063–1.210	0.19–0.39	40–55	20–35	3–5	10–12	2–5

[a]Reprinted with permission from N. B. Myant, *Cholesterol Metabolism, LDL, and the LDL Receptor,* Academic Press, San Diego (1990).

greatly from that of globular proteins, in which hydrophobic amino acids tend to be buried inside the molecule. Some membrane proteins are proteolipids and interact strongly with the hydrocarbon core of the membrane bilayer.

6.4. DERIVED LIPIDS

Derived lipids constitute a group of diverse compounds that includes prostaglandins and related compounds, isoprenoids, fat-soluble vitamins, and steroids. Prostaglandins, thromboxanes, and leukotrienes all derive from *arachidonic acid* (Table 6.1). We refer to them collectively as **eicosanoids.** These biologically active lipids, widespread in nature, have a variety of physiological effects.

6.4.1. Prostaglandins, Thromboxanes, and Leukotrienes

Prostaglandins were discovered in Sweden, where Ulf von Euler, in the early 1930s, first identified some compounds in human semen that affected blood pressure and smooth muscle contraction. He believed that the compounds originated in the prostate gland (hence, prostaglandins), but they were later shown to be synthesized in the seminal vesicles. We now know that prostaglandins occur in most, if not all, mammalian tissues. Prostaglandins have various hormonal effects, including lowering blood pressure, stimulating smooth muscle contractions, and regulating inflammatory reactions, blood coagulation, and the immune response.

Biochemists divide prostaglandins into a number of groups depending on the structure of the substituted five-membered ring. The first two prostaglandins to be isolated were called prostaglandin E and F because of their solubility in ether and phosphate buffer ("fosfat" in Swedish), respectively. We now designate all prostaglandins by three letters such as PGE, PGF, and so on (Figure 6.12).

The therapeutic action of **aspirin** (acetylsalicylic acid) is linked to prostaglandin biochemistry. Aspirin has been widely used since the end of the 19th century (1897) to relieve minor pain and treat inflammation and fever, but its mechanism of action was not discovered until some 25 years ago. In part, fever and inflammation arise from an overproduction of prostaglandins. John Vane showed in 1971 that aspirin as well as other "nonsteroidal anti-inflammatory" drugs inhibit *cyclooxygenase,* an enzyme functioning in the biosynthesis of prostaglandins from arachidonic acid. The inhibition by aspirin involves a covalent modification of the enzyme; an acetyl group is transferred from aspirin to a serine residue in the active

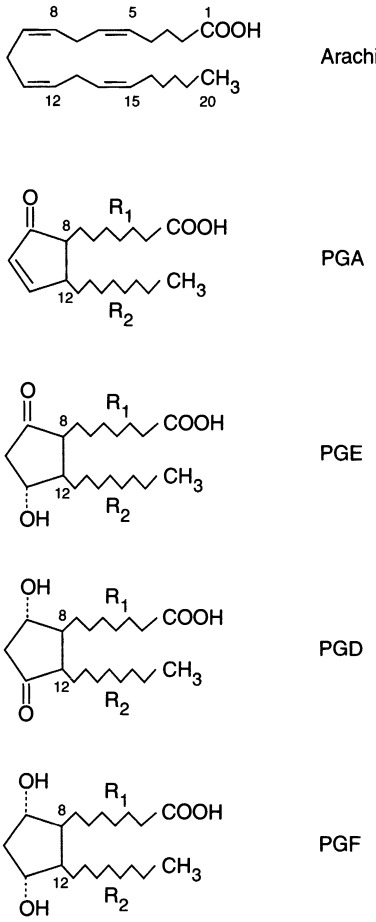


Figure 6.12. Structures of arachidonic acid and some of the prostaglandins derived from it.

site of the enzyme. By blocking cyclooxygenase, aspirin inhibits prostaglandin synthesis and alleviates the disease symptoms.

Thromboxanes contain a six-membered, cyclic ether (*oxane ring*) instead of the five-membered ring characteristic of prostaglandins (Figure 6.13). The compounds were called thromboxanes because they were first isolated from thrombocytes (blood platelets). Thromboxanes

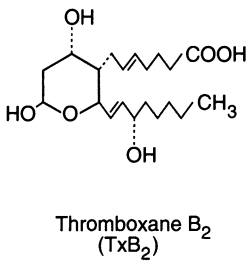
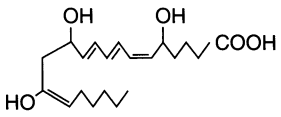


Figure 6.13. Example of a thromboxane.



Leukotriene B_4
(LTB_4)

Figure 6.14. Example of a leukotriene.

stimulate platelet aggregation and smooth muscle contractions.

Leukotrienes are formed by leukocytes (hence the name), macrophages, and other cells in response to immunological (and nonimmunological) stimuli. As distinct from prostaglandins and thromboxanes, leukotrienes have a linear rather than a cyclic structure (Figure 6.14); they contain three conjugated double bonds ("triene"). Leukotrienes lead to contractions of bronchial smooth muscles, stimulate vascular permeability, activate leukocytes, and play a role in allergy and asthma.

6.4.2. Isoprenoids: Carotenoids and Terpenoids

Carotenoids are yellow-orange, water-insoluble pigments that occur widely in plants, animals, and photosynthetic bacteria. They give carrots, pumpkins, cantaloupes, and sweet potatoes their characteristic color. A well-known representative of the group is *β-carotene* (Figure 6.15), a hydrocarbon that occurs in carrots and serves as a precursor of vitamin A.

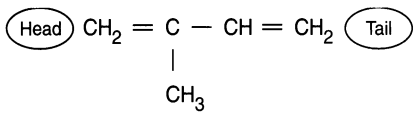
Chemically, carotenoids belong to a large class of compounds termed **isoprenoids.** Isoprenoids include fat-soluble vitamins and steroids (see below); they contain or derive from *isoprene.* The five-carbon molecule of isoprene (Figure 6.16) functions as a building block for many biochemically important molecules. Isoprene does not occur as such in nature; its biologically active equivalent is *isopentenyl pyrophosphate* (Figure 6.16). Isoprene units are usually linked together in a *head-to-tail* fashion, with the double bonds in the *trans* configuration.

Terpenoids are polyisoprenoids that may be linear or cyclic and in which the isoprene units are generally linked in a head-to-tail manner. The smallest terpenoid has a minimum of 10 carbons and is called a monoterpene (C_{10}). Larger terpenes include C_{15}, C_{20}, C_{30}, and C_{40} compounds. Terpenoids comprise a large group of diverse compounds such as natural rubber, plant hormones, sugar carriers, plant pigments, and insect hormones.

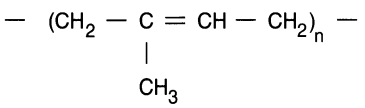
6.4.3. Fat-Soluble Vitamins

Fat-soluble vitamins comprise a group of compounds that are related structurally but have diverse functions. The group includes vitamins A, D, E, and K. By contrast, water-soluble vitamins (discussed in Section 8.4) are not related structurally but have a functional similarity in that they serve as components of coenzymes.

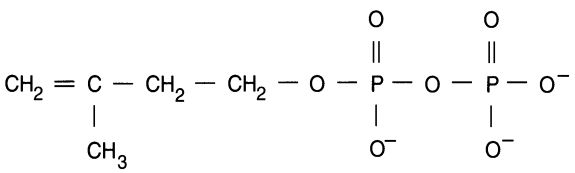
6.4.3A. Vitamin A. Inspecting Figure 6.15, you note that one molecule of β-carotene, when cleaved, should give rise to two molecules of vitamin A. Accordingly, one molecule of β-carotene should be nutritionally equivalent to two molecules of the vitamin. This, however, does not appear to be the case. To account for this discrepancy, researchers have suggested that β-carotene undergoes *asymmetric cleavage* and that the longer fragment



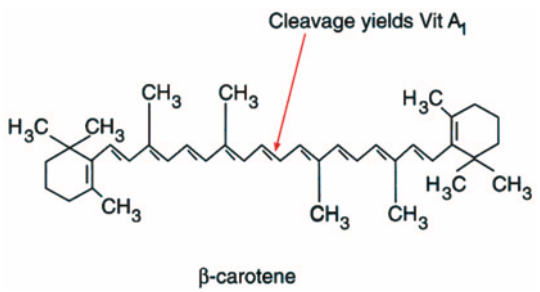
Isoprene



Polyisoprene
(head-to-tail polymer)



Isopentenyl pyrophosphate



β-carotene

Figure 6.15. β-Carotene—a precursor of vitamin A (see Figure 6.17). Cleavage at the arrow yields two molecules of vitamin A_1.

Figure 6.16. Isoprene and its biologically active equivalent, isopentenyl pyrophosphate.

is subsequently degraded further to produce only one vitamin A molecule.

Two forms of vitamin A occcur in nature; they differ only by one double bond (Figure 6.17). Vitamin A_1 has a double bond between C(5) and C(6) in the ring; vitamin A_2 has an additional double bond between C(3) and C(4). Vitamin A_1 predominates in animals and marine fish; vitamin A_2 occurs in freshwater fish. Vitamin A plays a key role in the biochemistry of vision and has an as yet undefined role in animal growth. A deficiency of vitamin A causes *night blindness,* characterized by a poor response of the retinal rods to faint light. Vitamin A deficiency can also lead to *xerophthalmia,* a dryness of the eye's conjunctiva.

While we readily obtain vitamin A through the diet, continuous excessive intake poses danger; too much vitamin A is toxic and can be fatal. A dietary oversupply, as by eating too many carrots, may also produce *carotenosis,* a condition readily diagnosed by yellowing of the skin.

The visual process involves two types of photoreceptor cells, *rods* and *cones.* Rod cells are adapted to respond to low light intensities and function in night vision.

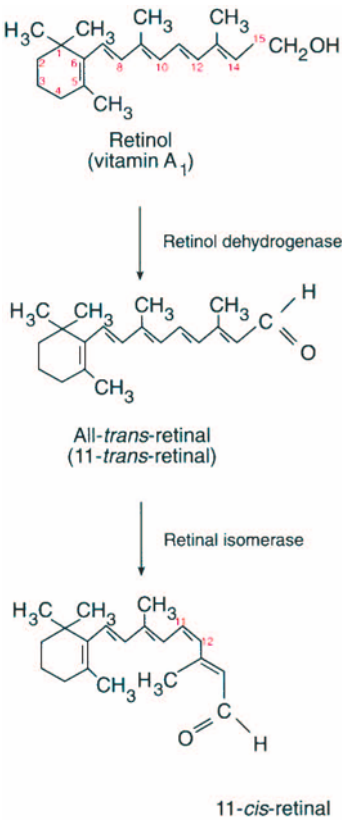


Figure 6.17. Vitamin A_1 (all-*trans*-retinol) and some related compounds. Vitamin A_2 has a second double bond in the ring, between C(3) and C(4).

Cone cells respond to high light intensities and are responsible for color vision. Both types of cells contain photosensitive complexes of protein and visual pigment. The pigment undergoes a number of complex molecular changes.

In rod cells, the photosensitive complex is a conjugated protein called *rhodopsin,* composed of the protein *opsin* and the visual pigment 11-*cis*-retinal (Figure 6.17). This pigment represents the aldehyde form of vitamin A in which the double bond at C(11) is *cis* and those at the remaining carbons of the chain are *trans.* The basic visual process consists of a cyclic set of reactions, termed the **visual cycle** (Figure 6.18). The key reaction involves an excitation of rhodopsin by light. This event comprises several steps, as a result of which 11-*cis*-retinal isomerizes to all-*trans*-retinal. The isomerization represents the primary event in visual excitation and leads to changes in the potential of the rod cell membrane. The changes in potential trigger an electrical impulse, transmitted via the optic nerve to the brain, where it undergoes processing as a visual event. All-*trans*-retinal does not fit into the binding site of the opsin molecule and dissociates from it. An isomerase catalyzes the conversion of all-*trans*-retinal back to 11-*cis*-retinal, which then binds to opsin to reconstitute rhodopsin.

6.4.3B. Vitamin D. Vitamin D is a *steroid* (see below), related to the *sterols* and derived from them by ultraviolet irradiation. Because of this synthetic route, vitamin D is not actually required in the diet provided the organism has adequate access to ultraviolet light, present in sunlight. Hence, vitamin D does not constitute a vitamin in the commonly accepted sense. Because the biologically active form of vitamin D acts as a hormone, we consider vitamin D a *prohormone,* the inactive precursor of a hormone.

Vitamin D_3, or *cholecalciferol,* represents the most abundant form of vitamin D. *1,25-Dihydroxycholecalciferol* is the hormonally active form of vitamin D, and its production from inactive sterol precursor requires the participation of three tissues—skin, liver, and kidney (Figure 6.19).

Cholecalciferol forms in the skin when ultraviolet light irradiates 7-dehydrocholesterol, an intermediate in the biosynthesis of *cholesterol.* Cholecalciferol undergoes hydroxylation by means of two enzymatic reactions, one in the liver, producing 25-hydroxycholecalciferol, followed by one in the kidney, producing 1,25-dihydroxycholecalciferol. 1,25-Dihydroxycholecalciferol acts like a hormone: it stimulates absorption of Ca^{2+} and inorganic phosphate across the intestinal wall, increases their serum concentrations, and regulates their utilization, particular-

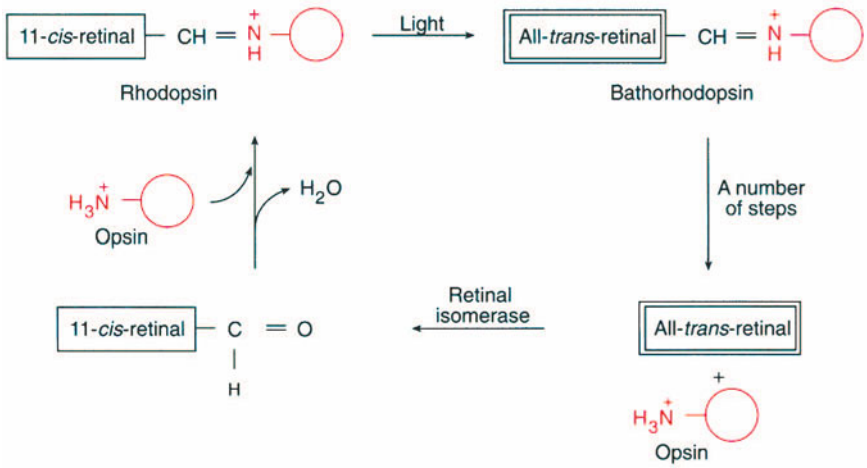


Figure 6.18. The visual cycle. Light excites rhodopsin and causes the isomerization of 11-*cis*-retinal to all-*trans*-retinal and transmission of a nerve impulse to the brain. The all-*trans*-retinal dissociates from the protein opsin and isomerizes to 11-*cis*-retinal. Rhodopsin re-forms by binding of 11-*cis*-retinal to an ϵ-NH_2 group of lysine in opsin.

ly by affecting the synthesis of bone matrix, a substance composed largely of calcium phosphate.

Supplementing milk with vitamin D involves ultraviolet irradiation of added *ergosterol,* a sterol isolated from yeast. Irradiation converts the sterol to *vitamin* D_2, or *ergocalciferol* (Figure 6.20). Vitamin D deficiency in young children leads to **rickets,** a disease characterized by softening and bending of bones. In adults, vitamin D deficiency leads to **adult rickets** or *osteomalacia,* a condition characterized by demineralization of bones so that

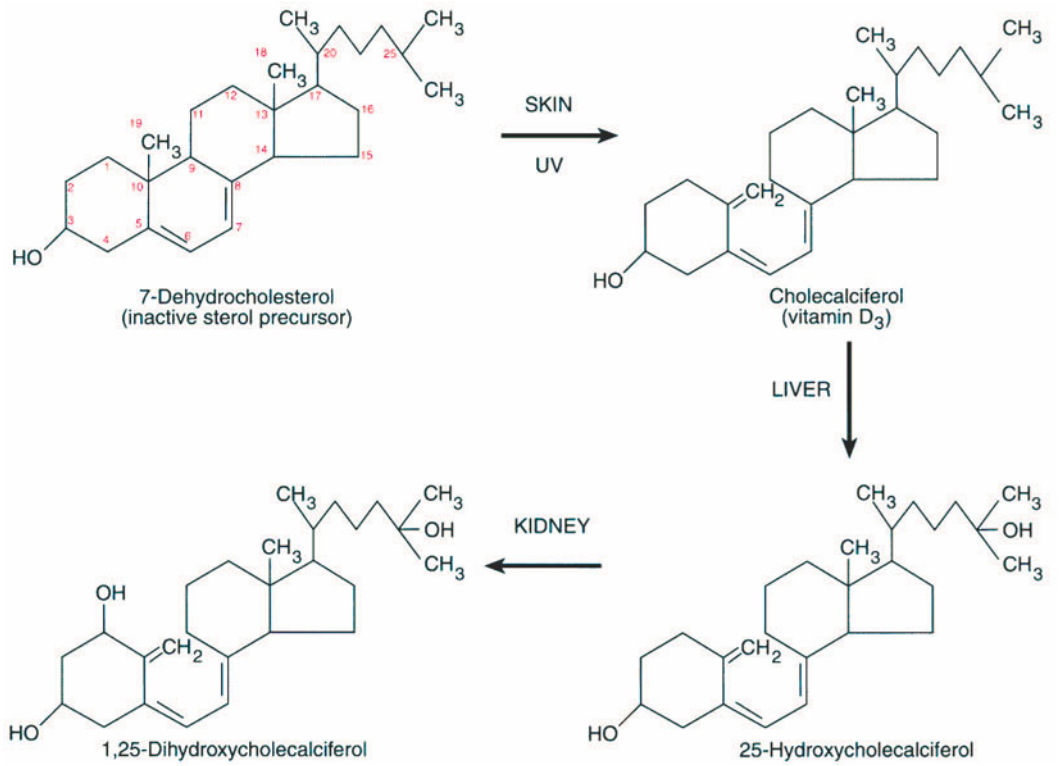


Figure 6.19. Production of the hormonally active form of vitamin D (1,25-dihydroxycholecalciferol) via sequential reactions in skin, liver, and kidney.

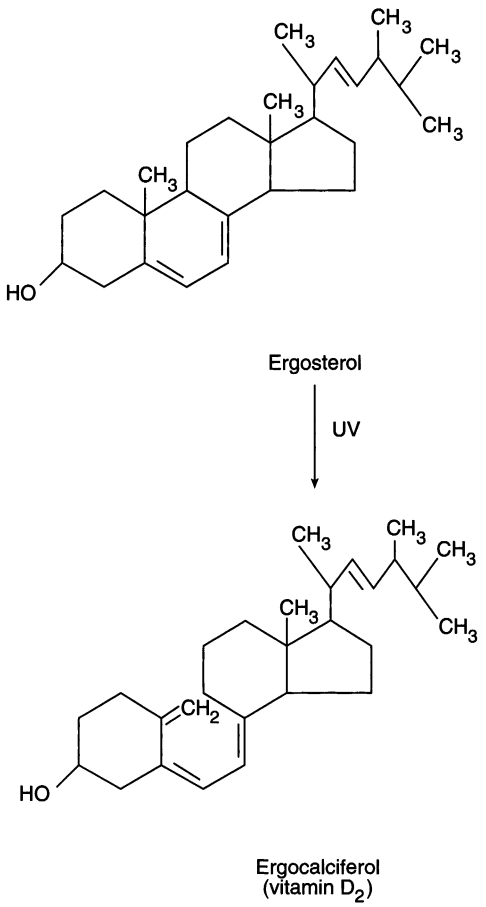


Figure 6.20. Formation of ergocalciferol (vitamin D_2) from ergosterol.

they become softer and more susceptible to fracture. Excessive dietary intake of vitamin D can be harmful, leading to toxic effects in children and to weight loss, irritability, and brittle bones in adults.

6.4.3C. Vitamin E. Vitamin E occurs in many forms, the most active being α-*tocopherol* (Figure 6.21). The vitamin has wide distribution in nature and is required for normal growth and fertility in animals. A deficiency of the vitamin leads to muscular dystrophy in some animals. No conclusive evidence has yet been obtained for the relationship of vitamin E to growth and reproduction in humans. Instead, investigators believe that vitamin E functions mainly as an *antioxidant.* Being a good reducing agent, vitamin E reacts with many strong and potentially very harmful oxidizing agents, such as hydrogen peroxide (H_2O_2), the superoxide anion radical ($O_2 \cdot^-$), and the hydroxy radical (·OH). These oxidizing agents (see also Section 2.5) can attack double bonds of unsaturated fatty acids, producing peroxides, a process termed *peroxidation.* Researchers think that a main function of vitamin E is to prevent the peroxidation of membrane lipids. Peroxidation results in the conversion of an unsaturated fatty acid to a hydroperoxide:

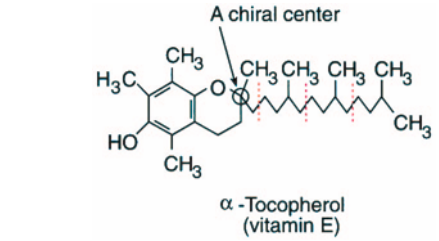
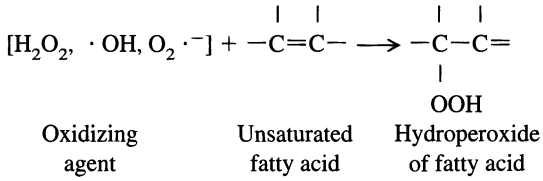


Figure 6.21. Structure of α-tocopherol (vitamin E). The aliphatic side chain contains residues of three isoprene units.



6.4.3D. Vitamin K. Vitamin K (from the German *Koagulation* or coagulation) is required for effective blood clotting. A deficiency of the vitamin results in prolonged clotting times and hemorrhagic disease. A number of forms of vitamin K occur in nature, including *vitamin* K_1 *(phylloquinone),* present in plants, and *vitamin* K_2 *(menaquinone),* found largely in animals and bacteria (Figure 6.22). Vitamin K is required for carboxylation of glutamic acid residues in the protein *prothrombin,* a reaction that produces γ-carboxyglutamate (Figure 6.23). The two negatively charged carboxyl groups of glutamic acid,

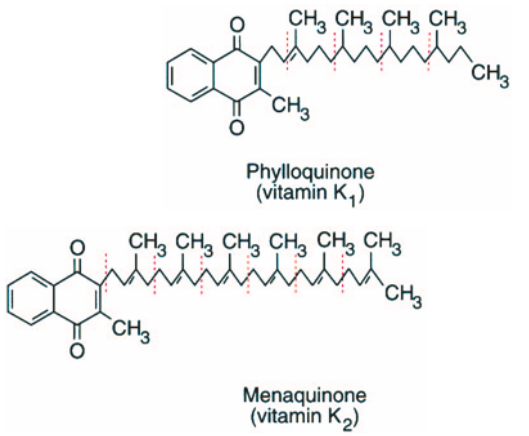


Figure 6.22. Two forms of vitamin K. The aliphatic side chains contain multiple residues of isoprene units.

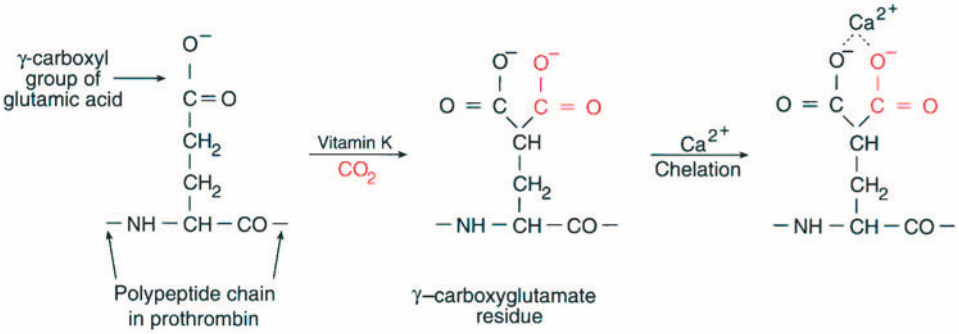


Figure 6.23. Formation of γ-carboxyglutamate, a chelator of Ca^{2+}, by carboxylation of glutamate.

in the modified prothrombin, bind calcium ions by chelation. The presence of Ca^{2+} is essential for blood clot formation.

6.4.4. Steroids

Steroids comprise a very large number of compounds of plant or animal origin that have varied functions. The steroid nucleus, called the *perhydrocyclopentanophenanthrene* nucleus (see Figure 6.24), consists of a rigid system of four fused rings, three six-membered and one five-membered. The cyclohexane rings of the steroids adopt puckered conformations. As for carbohydrates, the chair conformation is usually more stable than the boat conformation. The ring system frequently has a hydroxyl group or an oxygen at C(3), *angular methyl groups* at C(10) and/or C(13), and an *aliphatic side chain* at C(17). In addition to these modifications, a huge number of geometric and optical isomers can occur because of the rigidity of the ring system, the presence of double bonds, and the occurrence of chiral centers. Indeed, the number of known compounds is so large that a separate field of study has emerged, dealing exclusively with steroids.

The length of the aliphatic side chain, attached at C(17), serves as a convenient means for classifying steroids into a few large groups (Table 6.3). **Sterols** have the longest side chain. *Cholesterol* is the best-known sterol (Figure 6.24) and the principal sterol of vertebrates. It serves as precursor for bile acids and steroid hormones. Synthesis of cholesterol occurs in various tissues and involves a long series of steps, outlined in Section 13.6. Cholesterol occurs both as the free sterol and in the form of *cholesterol esters,* produced from cholesterol and fatty acids. Cholesterol esters are very hydrophobic. We will discuss the role of cholesterol in atherosclerosis in Section 6.6.

Bile acids are major degradation products of cholesterol and occur in the bile; they function in fat digestion. The two major bile acids are *cholic acid* and *deoxycholic acid* (Figure 6.25).

Bile acids become conjugated to either glycine or taurine, an aminosulfonic acid derived from cysteine. The reaction requires formation of an amide bond and results in conjugates called **bile salts** (Figure 6.25). Because bile salts are amphipathic, they function like soaps and detergents. When bile empties into the intestine, bile salts emulsify lipid components of the diet, primarily fats. They stabilize the emulsion of fats in the aqueous medium of the intestine, so water-soluble enzymes that catalyze the hydrolysis of fats (*lipases*) can bind to them.

Table 6.3. Classification of Steroids Based on the Side Chain at C(17)

	Number of carbons	
Class of compounds	Side chain	Molecule[a]
Sterols	8, 9, or 10	27, 28, or 29
Bile acids	5	24
Adrenal cortical steroids	2	21
Sex hormones	—	19

[a]Including two angular methyl groups.

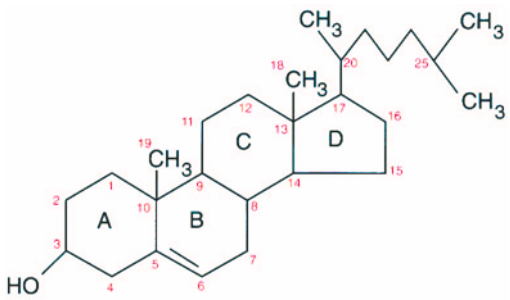


Figure 6.24. Cholesterol. The four fused rings (without substituents) constitute the perhydrocyclopentanophenanthrene nucleus.

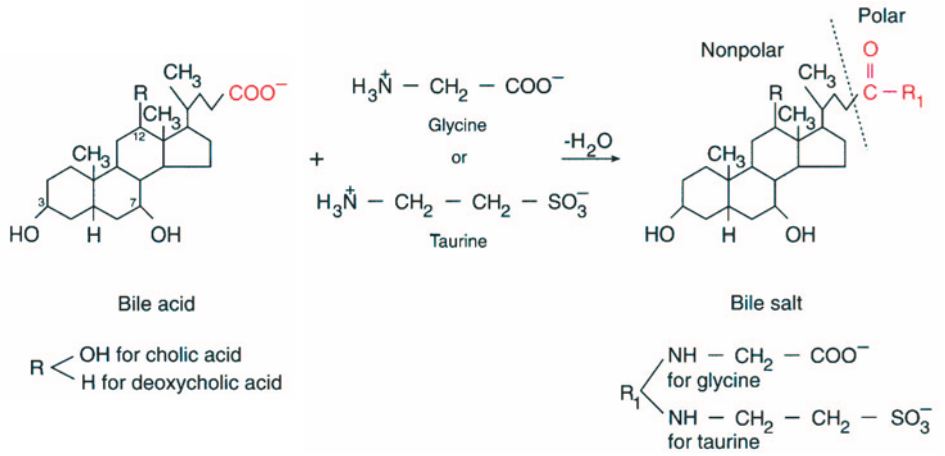


Figure 6.25. Conjugation of glycine and taurine with bile acids to form bile salts.

Adrenal cortical steroids form in the adrenal cortex of the *adrenal gland.* Many, but not all, adrenal cortical steroids function as hormones. Biochemists recognize two major classes of adrenal cortical steroids—glucocorticoids and mineralocorticoids. *Glucocorticoids* lead to gluconeogenesis, lipid mobilization, and protein catabolism. They also have antiallergic and anti-inflammatory effects. *Mineralocorticoids* act primarily on water and electrolyte balance by stimulating retention of sodium and excretion of potassium by the kidney.

Sex hormones (Figure 6.26) affect sexual development and function. They are formed primarily by the gonads (testes in males, ovaries in females) and secondarily by the adrenal cortex. Sex hormones include *androgens,* male sex hormones produced by the testes, and *estrogens,* female sex hormones produced by the ovaries. Testes also produce estrogens, and ovaries also produce androgens. However, the balance is such that testes produce mostly androgens, and ovaries mostly estrogens. Androgens and certain androgen derivatives have the additional effect of stimulating muscle growth and muscle strength and are known as *anabolic steroids.* Debate surrounds use of these compounds by athletes.

Figure 6.26. Testosterone and estradiol, the principal male and female sex hormones, respectively.

6.5. BIOMEMBRANES—STRUCTURAL ASPECTS

6.5.1. Composition and Structure

Membranes constitute essential components of all living cells. They represent water-insoluble permeability barriers between the interior of the cell and the external environment, as well as between the contents of subcellular organelles and the cytosol. Membranes permit the existence of unique cellular compartments. Biological membranes, or **biomembranes,** consist of self-assembling sheetlike structures, held together by noncovalent interactions.

6.5.1A. Membrane Composition. All biomembranes consist of lipid (20–80%) and protein (80–20%), and most contain carbohydrate as well (Table 6.4).

Table 6.4. Chemical Composition of Some Biomembranes[a]

Membrane[b]	Percentage (by weight)		
	Protein	Lipid	Carbohydrate
Myelin	18	79	3
Liver cell (mouse)	46	54	3
Red blood cell (human)	49	43	8
Mitochondrion (outer)	52	48	2–4
Chloroplast (spinach)	70	30	6
Halobacterium	75	25	0
Gram-positive bacterium	75	25	10
Mitochondrion (inner)	76	24	1–2

[a]Reprinted with permission from G. Guidotti, *Annu. Rev. Biochem.* 41:731–750 (1972). © 1972 by Annual Reviews, Inc.
[b]Cell membranes, except for mitochondria and chloroplasts.

Biomembranes are only a few molecules thick, but their lipid interior constitutes an effective barrier to polar molecules and ions. In addition, their structural proteins and many of their lipids function in three key cellular processes—catalysis, binding at receptor sites, and transport. *Catalysis* occurs when an enzyme is located in or bound to the membrane. Binding of substances to the membrane may involve specific *receptor sites,* such as those for LDL (discussed in Section 6.6). Lastly, membranes provide specific *transport mechanisms* that regulate the flow of substances into and out of the cell or the organelle. As we shall see, both with respect to their structure and their function, membranes represent dynamic rather than static biochemical systems.

6.5.1B. Membrane Structure. Although each type of membrane has a unique design and function, all membranes share general structural features, the main one being a central core composed of a double layer of amphipathic lipid molecules. The two *monolayers* make up a **lipid bilayer** that consists largely of phospholipid molecules. The *nonpolar (hydrophobic) tails* form the interior of the double layer, held together by hydrophobic interactions, and the *polar (hydrophilic) heads* are at the two membrane surfaces (Figure 6.27).

Two Dutch investigators, E. Gorter and F. Grendel, first proposed the concept of a lipid bilayer in 1925. Hugh Davson and James Danielli elaborated on this proposal in 1935 by suggesting that protein molecules attach to the exterior membrane surface via noncovalent electrostatic interactions. The Davson–Danielli model was subsequently refined by J. D. Robertson (1959) and became known as the **unit membrane hypothesis.**

Support for the bilayer nature of biological membranes comes from a variety of sources. First, we can easily see double layers in electron micrographs of membrane preparations (Figure 6.28). Using a special technique called *freeze fracture,* investigators can cleave the membrane between its two monolayers and examine each layer separately (Figure 6.29). Second, we know that the preferred structure of phospholipids in an aqueous environment is a

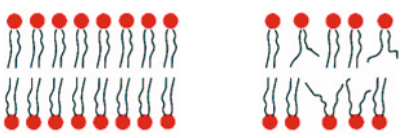


Figure 6.27. (A) A phospholipid bilayer containing only saturated fatty acids; (B) a bilayer whose regular structure is disrupted by unsaturated fatty acids.

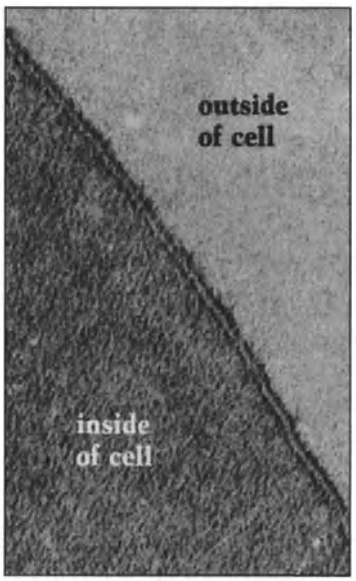


Figure 6.28. Electron micrograph showing the bilayer of a red blood cell plasma membrane. (Courtesy of Dr. J. D. Robertson.)

lipid bilayer rather than a micellar arrangement. The preference results from the nonpolar tails of these lipids being too bulky to fit in the interior of a micelle. A micelle accommodates the hydrocarbon chains of soap molecules but cannot accommodate the larger phospholipid structures. Lastly, we are able to form synthetic bilayer assemblies. *Sonication* of suspensions of phospholipids in aqueous salt solutions can produce cell-like structures in which a lipid bilayer surrounds an aqueous compartment. Such aggregates, called *liposomes* (Figure 6.30), constitute excellent models for studying membrane properties.

Subsequent investigations have necessitated a revision of the unit membrane hypothesis. Two major changes deal with *location of proteins* and *fluidity of the membrane.*

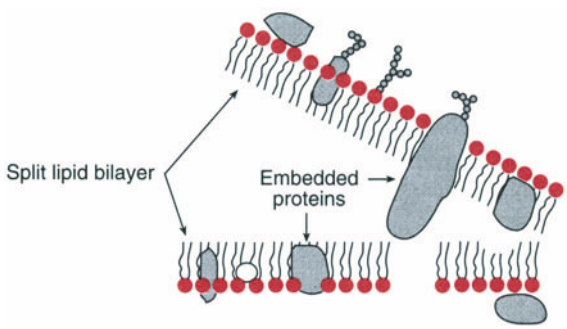


Figure 6.29. The freeze fracture technique. A frozen specimen is fractured with a knife, yielding cleavage surfaces that are examined by electron microscopy. Freeze fracture of membranes sometimes results in a cleavage between the two monolayers.

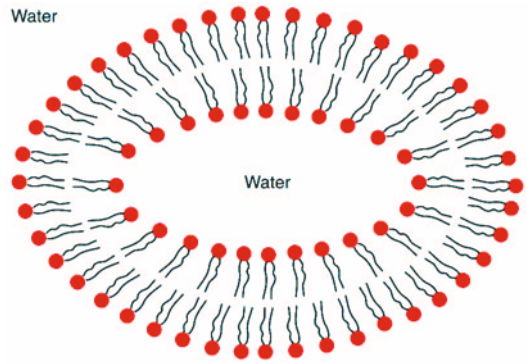


Figure 6.30. Structure of a liposome.

According to present views, lipids, primarily phospholipids, form a bilayer as described above. Proteins, however, do not attach only at the external surface of this bilayer. Instead, they may also attach at the inner surface and, most importantly, may occur inside the membrane as well. Some proteins are "dissolved" in the lipid bilayer. Thus, lipid and protein molecules form a mosaic pattern throughout the membrane. The lipid bilayer has a thickness of about 35–40 Å, and the total thickness of the membrane, including associated proteins, typically amounts to 70–90 Å.

The second change made in the original hypothesis deals with membrane rigidity. We now view the entire membrane structure as a dynamic rather than a static assembly. Both the lipids and the proteins have some flexibility, permitting a certain amount of lateral movement within each monolayer. Lipids, but not proteins, are also capable of some transverse movement from one monolayer to the other. These various molecular movements endow the membrane with an important degree of fluidity.

6.5.1C. Membrane Proteins. Membrane proteins fall into two categories. **Integral** or **intrinsic proteins** are embedded in the hydrophobic matrix of the membrane. They may lie within one layer, extend partially through it, or even extend completely through the bilayer, spanning the membrane from one side to the other (*transmembrane proteins*). We usually cannot remove integral proteins without disrupting the membrane by means of organic solvents or detergents. Some integral proteins represent genuine structural components of the membrane, often specially suited to fit into the lipid matrix. Such proteins may be amphipathic and contain high proportions of hydrophobic amino acids, especially where embedded in the membrane. These structural proteins may also be elongated or fibrous in shape. Protein parts that span membranes frequently have an α-helical configuration.

Other integral proteins constitute functional components of the membrane. Some serve as transport proteins involved in the regulation of solute movement across the membrane. Others have enzymatic properties and catalyze specific chemical reactions taking place in the membrane. Some function to define receptor sites on the surface of the membrane and bind specific substances to the membrane.

Peripheral or **extrinsic** proteins, by contrast, are located at one or the other surface of the lipid bilayer. They bind electrostatically to the polar region of a lipid monolayer on either the external or the internal side of the membrane. We can extract peripheral proteins under proper conditions of ionic strength or in the presence of chelating agents, without disrupting the hydrophobic structure of the membrane.

6.5.2. Fluidity and Asymmetry

6.5.2A. Membrane Fluidity. The currently accepted model of biomembranes postulates a significant degree of fluidity. Both the lipids and some proteins are capable of lateral movement within the lipid matrix of the membrane, and lipids can undergo transverse movement from one monolayer to the other as well. Accordingly, we refer to this model of membrane structure, proposed by Jonathan Singer and Garth Nicolson in 1972, as the **fluid mosaic model** (Figure 6.31).

Fluidity of the membrane results from the occurrence of unsaturated fatty acids in the phospholipid bilayer. Recall that every *cis* double bond in an unsaturated fatty acid results in a bend of the hydrocarbon chain (Figure 6.2). The presence of unsaturated fatty acids in the lipid bilayer prevents close packing of the nonpolar components of the lipid matrix (see Figure 6.27). Membrane stiffness correlates with an increased ratio of saturated/unsaturated fatty acid chains. Increased concentrations of saturated fatty acids result in closer packing in the lipid bilayer and less opportunity for lateral movement of lipids. High concentrations of the rigid ring system of cholesterol also lead to increased membrane stiffness.

We now know that lipid molecules can rotate and diffuse laterally within one monolayer, or *leaflet,* of the lipid bilayer. Lipids can also, though less frequently, move from one monolayer into the other (Figure 6.32). Such *transverse diffusion,* or *flip-flop,* is much more difficult to achieve than *lateral diffusion* or *rotational diffusion.* In a flip-flop transition, the polar head of a lipid must move through a nonpolar environment, an energetically unfavorable transfer. Lateral movement of proteins appears to be more limited than that of lipids, and these proteins almost never rotate.

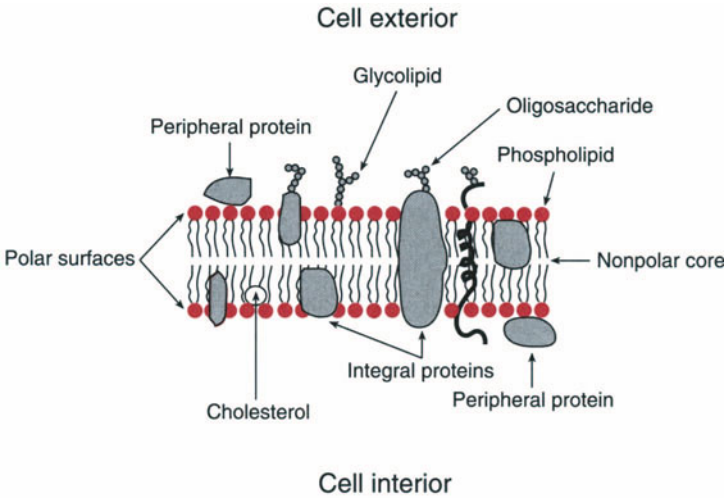


Figure 6.31. Diagram of the fluid mosaic model of biological membranes.

6.5.2B. Membrane Asymmetry. The original unit membrane hypothesis alluded to membrane asymmetry by postulating that proteins were attached only to the membrane exterior. While we now recognize that proteins can be attached to *either* side of the membrane, as well as being located internally, the concept of membrane asymmetry remains firmly established. The nature of the proteins and their attachment at the internal and external sides of the membrane are not identical. Integral proteins have asymmetrical locations within the membrane; some proteins occur in one monolayer, other proteins occur in the second monolayer (see Figures 6.29 and 6.31), and transmembrane proteins have different sections of their molecules in each monolayer. Additionally, the phospholipid compositions of the bilayers differ. Lastly, in cell membranes, carbohydrates are attached only to the exterior membrane surface. Consequently, biological membranes show a definite sidedness. The internal membrane surface differs from the external one in both composition and function, and the nature and distribution of proteins in the two monolayers is not identical. This asymmetry ensures that membrane processes possess directionality; a solute or signal may move across the membrane in one direction but not in the other.

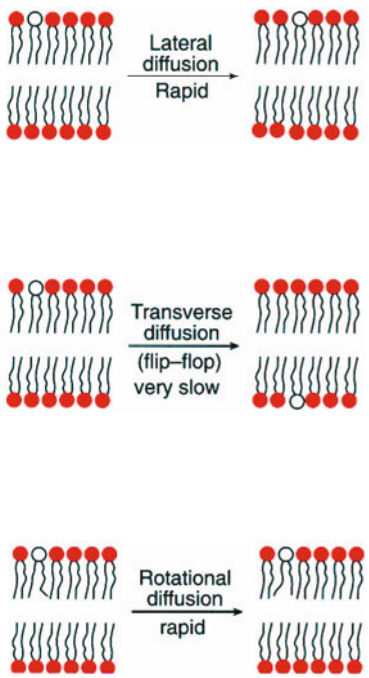


Figure 6.32. Movement of phospholipids in biological membranes.

6.6. BIOMEMBRANES—FUNCTIONAL ASPECTS

6.6.1. Transport across Membranes

Three major types of transport mechanisms operate in biological membranes: *passive transport, facilitated transport,* and *active transport.* We will discuss each of these in turn.

6.6.1A. Passive Transport. **Passive transport** refers to the movement of substances across a semipermeable membrane that results from **diffusion**—the random Brownian motion of ions and molecules. Diffusion is a function of the **concentration gradient,** $\Delta c/\Delta x$, across the membrane; Δc represents the change in con-

centration over the distance Δx. Diffusion proceeds *along the concentration gradient;* a substance diffuses *downward* in the gradient (*net* transfer is downward), from higher to lower concentration. Diffusion also depends on the *diffusion coefficient, D,* of the moving particle, which is a function of the particle's size and shape. Lastly, in the case of biological membranes, the relative polarity or nonpolar character of a diffusing substance affects its ability to penetrate into and move through the lipid bilayer.

Diffusion proceeds until an equilibrium is established at which the concentration of the diffusing substance is equal on both sides of the membrane. At that point the driving force of diffusion, the concentration gradient (more precisely, the entropy change), equals zero, and no further *net* movement of material across the membrane takes place. We also refer to the passive transport produced by diffusion as *nonmediated transport.*

6.6.1B. Facilitated Transport. Another type of transport is also driven by a concentration gradient but requires the assistance of some factor over and above the movement due to simple diffusion. **Facilitated transport,** also termed *mediated transport,* can proceed by means of three types of mechanisms (Figure 6.33), involving (i) a mobile carrier, (ii) a passage through the membrane (pore or channel), or (iii) an integral transmembrane protein.

Facilitated transport by means of a **mobile carrier** requires that a substance combine with some carrier that moves the substance across the membrane. The mobile carrier may be a protein, called variously *transport protein, translocase,* or *permease,* but it may also be some other component, such as an *ionophorous antibiotic.* **Ionophores** are lipid-soluble compounds that mediate the transport of ions across membranes. Because of their hydrophobic exteriors, ionophores are soluble in the membrane, and their hydrophilic interiors provide polar environments through which ions can move. Some ionophores constitute *channel formers* (see below); others function as mobile carriers. In some cases, the ionophore is an antibiotic. Valinomycin is an ionophorous antibiotic that functions as a mobile carrier (Figure 6.34). Valinomycin is a *depsipeptide,* a compound consisting of alternating amino and hydroxy acids, linked by alternating peptide and ester bonds. Valinomycin permits the movement of up to 10^4 K^+ ions per second across a membrane.

Facilitated transport can also occur through a **passage** in the membrane. A passage consists of either a pore or a channel. A spacing between adjacent phospholipids

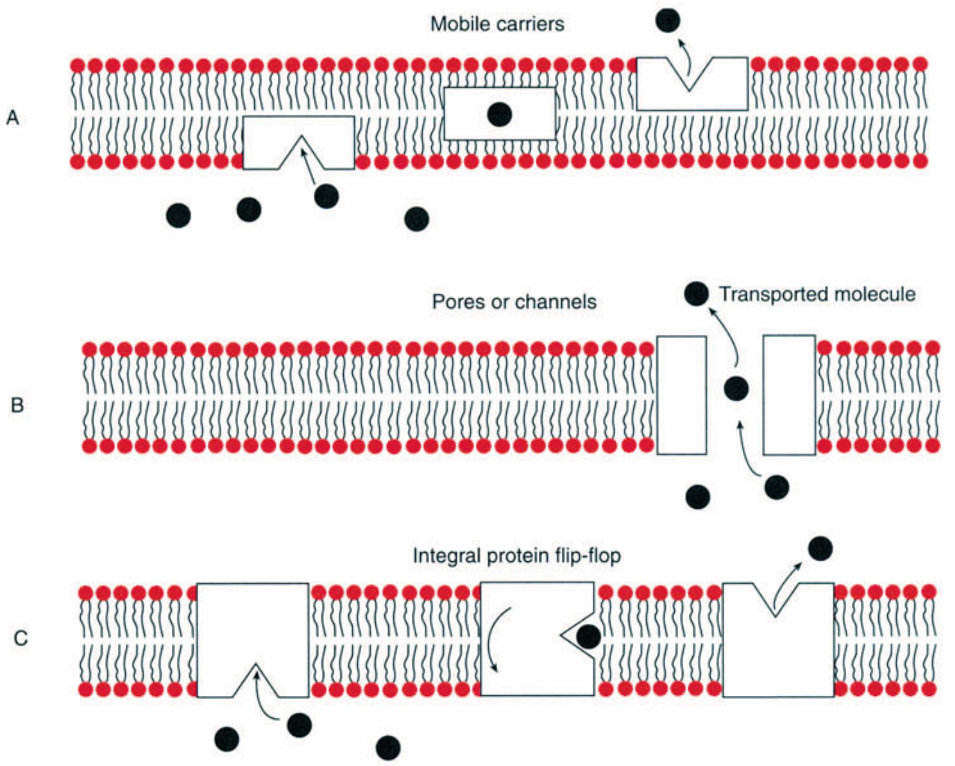


Figure 6.33. Three possible mechanisms of facilitated transport.

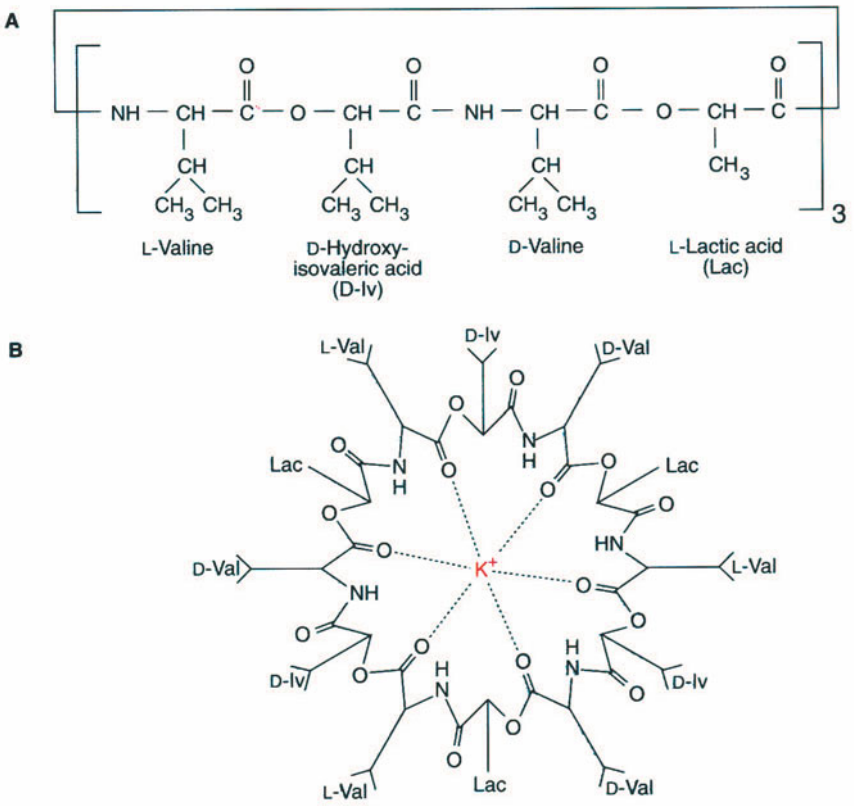


Figure 6.34. Valinomycin, an ionophorous antibiotic. (A) Structure of the repeating unit. (B) Structure of the valinomycin–potassium complex.

or a channel. A spacing between adjacent phospholipids constitutes a *pore*. Compounds, especially nonpolar ones, can readily diffuse through such pores. A *channel* consists of an opening through the bilayer, lined by hydrophilic groups that provide a polar environment for movement of polar solutes. Most channels are lined by integral proteins, but some are lined by ionophores. Gramicidin A represents an ionophorous antibiotic and a channel former (Figure 6.35). The gramicidin A channel provides highly efficient transport, allowing passage of over 10^7 K^+ ions per second.

Most channels show high selectivity and allow passage of only certain compounds. Some channels are *gated,* meaning they open and close in response to specific control mechanisms that can involve binding of a ligand to a receptor (*ligand-gated channels*) or changes in the potential of the membrane (*voltage-gated channels*). At times, a channel connects two adjacent cells separated by a tiny gap of about 3 nm. Such a communicating junction between two cells, or *gap junction* (Figure 6.36), forms by juxtaposition of clusters of a transmembrane protein called *connexin.* A cluster of six connexins in one cell membrane faces a similar cluster in a second membrane.

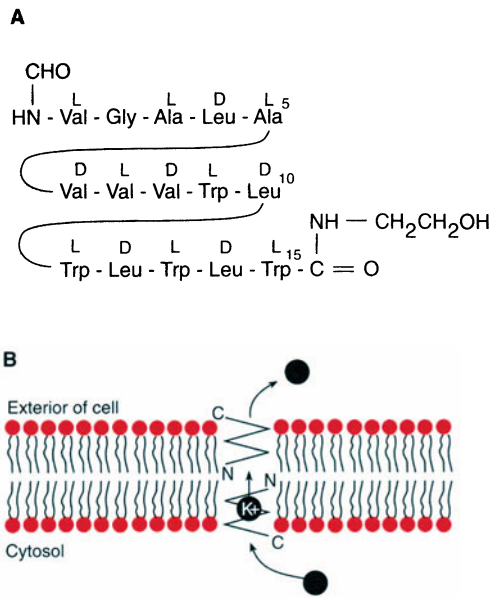


Figure 6.35. Gramicidin A, an ionophorous antibiotic. (A) Structure of gramicidin A. The molecule consists of an alternating sequence of L- and D-amino acids. (B) Proposed transmembrane channel formed by two helical molecules of gramicidin A, joined head-to-head. The channel is large enough to permit passage of alkali-metal cations.

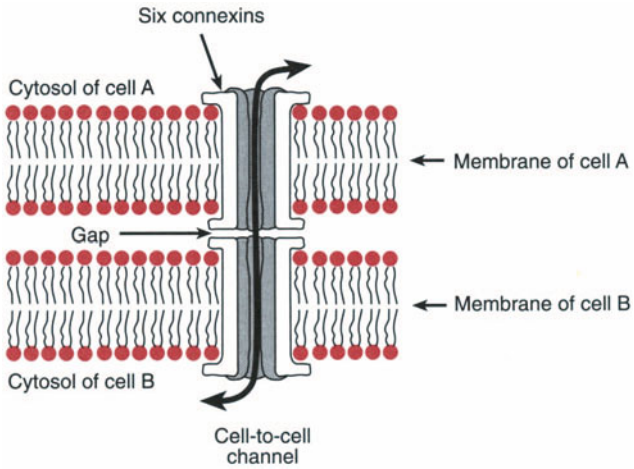


Figure 6.36. A gap junction between cells, made by juxtaposition of a transmembrane protein (connexin). Clusters of six connexins in one membrane face a similar cluster in a second membrane.

The connexins surround a central aqueous channel through which solutes move directly from one cell to another. Gap junctions are relatively nonspecific and function as gated channels that respond to specific stimuli.

In theory, membrane transport by means of a **transmembrane protein** could serve as a third possible mechanism of facilitated transport. An integral membrane protein having a binding site for a specific ligand on one side of the membrane could transfer that ligand to the other side by a flip-flop (Figure 6.33C). However, since it currently seems that integral membrane proteins are incapable of transverse movement from one monolayer to the other, a flip-flop mechanism appears unlikely.

Facilitated transport differs from passive transport in that the system can be *saturated.* Much as the reaction between an enzyme and its substrate reaches a saturation stage when all the enzyme molecules are combined with substrate per unit time, so the combination of a compound with its carrier will reach a point at which all of the carriers are combined with compound per unit time. At that point, the transport system is saturated. Because of this property, facilitated transport systems can be described by hyperbolic curves similar to the v versus [S] curve of a typical enzyme reaction. Transport via channels and pores likewise shows saturation, because the number of pores or channels is limited and the capacity for transfer, no matter how effective, does have a physical limit.

Facilitated transport constitutes a more complex process than passive transport and is often very rapid and highly selective. However, it still represents only a special case of diffusion. Ultimately, given enough time and in the absence of other intervening factors, facilitated transport will attain an ordinary equilibrium. At that point, the concentrations of the free substance being transported will be identical on both sides of the membrane.

6.6.1C. Active Transport.

Active transport involves the movement of a substance across a membrane *against its concentration gradient.* The substance moves *upward* in the gradient, from lower to higher concentration. It accumulates on one side of the membrane, despite the force of diffusion driving it in the opposite direction. Active transport requires the expenditure of energy, and an equilibrium in concentration across the membrane does not occur as long as a source of energy exists to drive the transport. Active transport must, therefore, be coupled in some fashion to a source of energy. Energy coupling may involve hydrolysis of an energy-rich compound, simultaneous transport of a second substance, or a molecular modification of the substance transported.

An important mechanism of using a source of energy for active transport consists of *coupling the transport directly to hydrolysis of ATP (adenosine triphosphate) or some other energy-rich compound.* (More will be said about this coupling in Section 9.3.) We call active transport that proceeds in this fashion a **primary active transport** or a **pump.**

The *sodium–potassium pump,* or *Na^+-K^+ ATPase,* represents one of the best-studied primary active transport systems. This tetrameric protein functions as both an enzyme, catalyzing ATP hydrolysis, and as a transport protein. Two large (α) subunits (MW = 110,000) are nonglycosylated transmembrane proteins, have the enzymatic activity, and bind ATP on the internal side of the membrane. Two small (β) subunits (MW = 55,000) are glycosylated proteins located on the external side of the mem-

brane. The overall stoichiometry involves transporting three Na^+ ions out and two K^+ ions in for every molecule of ATP hydrolyzed. Both ions are moved *against* their concentration gradients, since typical concentrations in animal cells are approximately:

	Intracellular	Extracellular
Na^+	10m*M*	140m*M*
K^+	100m*M*	5m*M*

The Na^+-K^+ ATPase constitutes an *electrogenic pump;* it generates a gradient of electrochemical potential as a result of the unequal movement of the two cations. A postulated mechanism for the pump includes two conformational changes of the enzyme resulting from phosphorylation and dephosphorylation (Figure 6.37). These conformational changes open the channel to the outside and inside, respectively.

A second mechanism of using a source of energy for active transport involves *coupling the transport to the simultaneous movement of a second substance down its concentration gradient.* We term this mode of active transport **secondary active transport** or **cotransport.** The second substance may move in the same direction as the first (*symport*) or in the opposite direction (*antiport*). The energetically favored transport of the second substance provides the energy to drive the transport of the first against its concentration gradient. Absorption of glucose from the intestine illustrates the principle of cotransport. Movement of glucose into the cells of the intestine is coupled to, and driven by, the simultaneous movement of sodium into the cells. Glucose moves upward and sodium moves downward in their respective concentration gradients.

Lastly, energy for active transport may be derived from a *molecular modification of the substance transported.* We refer to this type of active transport as **group translocation.** The transport of sugars in bacteria illustrates this mechanism. In bacteria, sugars are phosphorylated, either as they move through the membrane or as soon as they emerge from it on the interior side. The charged phosphorylated sugars cannot move back through the membrane and are trapped inside the cell.

All three forms of active transport, like those of facilitated transport, show *saturation* of the system, involve *substrate specificity,* and are subject to *inhibition* by particular inhibitors.

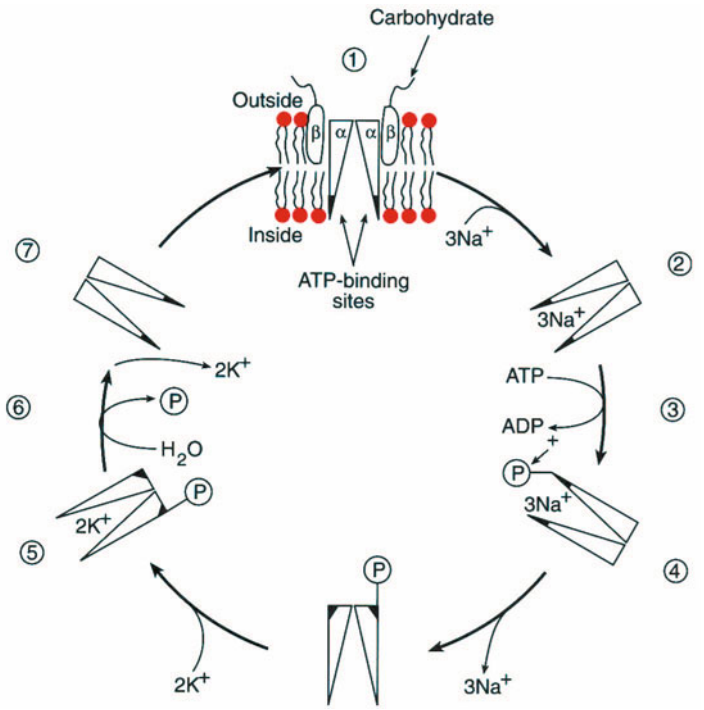
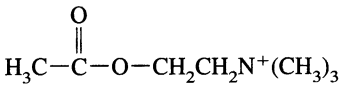


Figure 6.37. Operation of the Na^+-K^+ pump. (1) Pump is open to inside; (2) three Na^+ ions are taken up from inside; (3) ATP phosphorylates one subunit inside; (4) phosphorylation induces a conformational change in the pump, resulting in expulsion of three Na^+ ions to the outside; (5) two K^+ ions are taken up from outside; (6) pump is dephosphorylated inside; (7) dephosphorylation induces a conformational change in the pump, resulting in expulsion of two K^+ ions to the inside. For simplicity, the membrane bilayer has been omitted in steps 2–7.

6.6.2. The Cell Membrane and Nerve-Impulse Transmission

We will conclude our discussion of biomembranes by describing the cell membrane's role in two important physiological systems. We will first consider transmission of a nerve impulse and then, in the next section, atherosclerosis.

Transmission of a nerve impulse from one *neuron* (nerve cell) to another involves a special communicating junction called a *synapse.* A synapse comprises the functional area of contact between two nerve cells. The nerve impulse crosses a gap, called the *synaptic cleft,* that separates the *presynaptic* and *postsynaptic* cells (Figure 6.38). When the nerve impulse arrives at the presynaptic cell in the form of an *action potential* (a *depolarization* of the membrane), it triggers the opening of a voltage-gated channel that allows Ca^{2+} to flow into the cell. The influx of Ca^{2+} causes fusion of *synaptic vesicles,* containing the neurotransmitter acetylcholine, with the cell membrane.



Acetylcholine

The acetylcholine is released by *exocytosis* (discharge of fluids and particles from a cell) into the synaptic cleft, diffuses across it, and becomes bound to

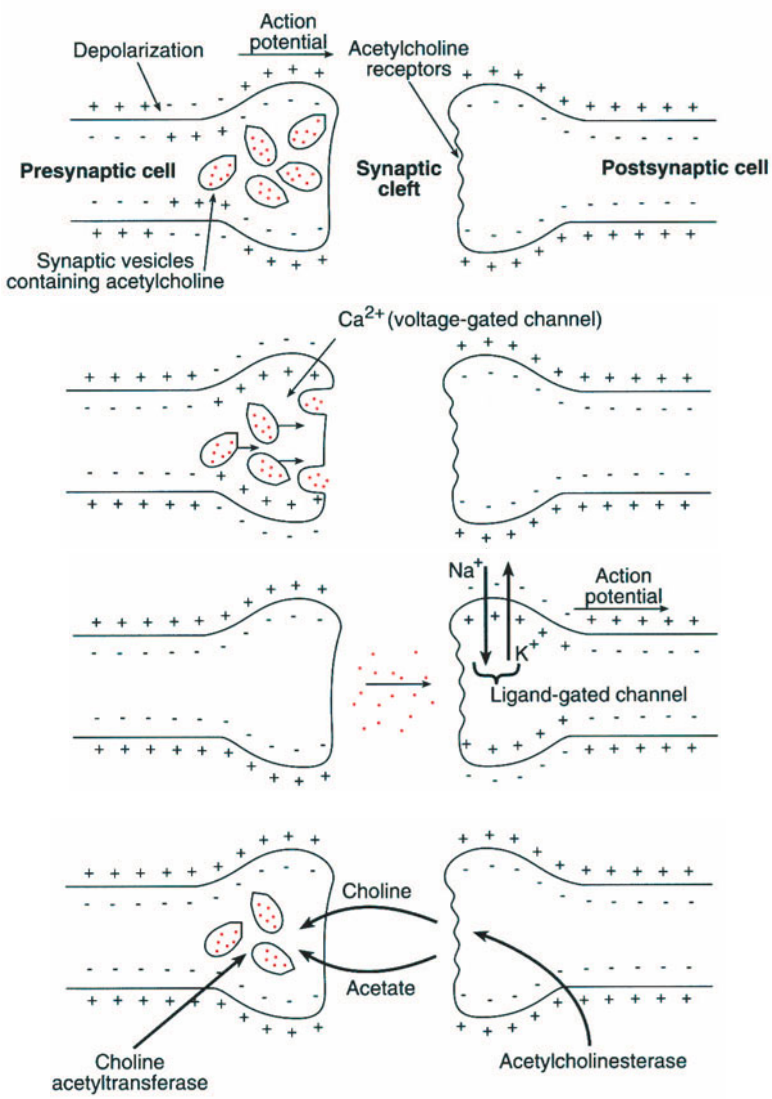


Figure 6.38. The sequence of events in transmission of a nerve impulse from one neuron to another at a synapse.

specific acetylcholine receptors on the membrane of the postsynaptic cell. These receptors represent ligand-gated channels that, upon binding acetylcholine, open and allow Na^+ to diffuse in and K^+ to diffuse out. Both ions move along their respective concentration gradients. The acetylcholine receptor is a pentameric protein (MW ≈ 250,000) that consists of four different types of rod-shaped subunits ($\alpha_2\beta\gamma\delta$) arranged in such a fashion that they describe a central, water-filled channel through which both sodium and potassium ions move. The channel permits very rapid diffusion, so that some 20,000 ions of each type can move through per millisecond.

Movement of Na^+ and K^+ across the postsynaptic membrane depolarizes it. The resulting action potential is propagated along the cell. In this way the nerve impulse has been transmitted from the presynaptic cell to the postsynaptic one. The entire process is extremely rapid, occurring in about one millisecond.

It is essential that acetylcholine be degraded immediately, before another nerve impulse arrives. This occurs by means of *acetylcholinesterase,* a serine protease that catalyzes the hydrolysis of acetylcholine to choline and acetate:

$$\text{Acetylcholine}^+ + H_2O \longrightarrow \text{acetate}^- + \text{choline}^+ + H^+$$

Acetylcholinesterase is inhibited by diisopropyl fluorophosphate (Figure 4.16), a component of nerve gas. Inhibition of the enzyme blocks transmission of nerve impulses, which rapidly paralyzes vital functions. Following the action of acetylcholinesterase, choline and acetate diffuse back into the presynaptic cell, where they are recombined by action of a second enzyme, *choline acetyltransferase:*

$$\text{Acetyl-S-CoA}^{4-} + \text{choline}^+ \longrightarrow \text{acetylcholine}^+ + \text{HS-CoA}^{4-}$$

In this reaction, the acetyl group is transferred to choline via coenzyme A ($HS\text{-}CoA^{4-}$). The structure of this coenzyme is given in Figure 11.7, and its properties are discussed in Section 11.1.

6.6.3. The Cell Membrane and Atherosclerosis

Atherosclerosis is a disease of the arteries characterized by gradual accumulation of deposits in the arterial wall. These deposits, termed *plaques,* consist of cholesterol, cholesterol esters, collagen, elastic fibers, and proteoglycans. Researchers believe that an injury to the arterial lining, followed by leakage of plasma into the arterial wall, triggers plaque formation. As the plaque grows, the diameter of the artery gradually diminishes and impedes blood flow. When the flow of blood is blocked completely, a heart attack, a stroke, or death can ensue.

The role of cholesterol (Figure 6.24) in atherosclerosis has been investigated and debated for many years. High levels of plasma cholesterol (normal levels range from about 130 to 200 mg/100 ml) have long been associated with increased risk of cardiovascular disease, but there exists a complex interplay of this and other factors.

We do not question that the blood flowing through the arteries carries cholesterol and that cholesterol constitutes a major part of atherosclerotic plaques. We can understand, therefore, that early investigators thought that blood cholesterol deposited in the plaques *caused* the disease. We now know, however, that a variety of tissues in the body, including arterial walls, synthesize cholesterol. We do not know how much plaque cholesterol is actually synthesized in the arteries and how much comes directly from the bloodstream. Moreover, we now consider deposition of cholesterol to be a *symptom* rather than a cause of the disease. In recent years, investigators have established that the incidence of atherosclerosis correlates with increased plasma levels of LDL and that the nature of the disease is related to the metabolism of LDL. Because of its role, LDL has been dubbed "bad" cholesterol.

LDL is taken up by cells in a process termed *receptor-mediated endocytosis* and elucidated by M. S. Brown and J. L. Goldstein in 1977. The receptors are specific binding sites for LDL on the cell membrane. Endocytosis describes the process whereby cells pinch off a membrane section to engulf and take up fluids or particles.

The LDL particle binds via its *apoprotein B-100* to a specific receptor in the cell membrane. The receptor is a large glycoprotein that spans the membrane. Receptors are clustered in membrane indentations, lined with a protein (*clathrin*) that can form a lattice structure. The membrane indentations are called *coated pits* (Figure 6.39).

When LDL binds to receptors in a coated pit, the entire complex buds off into the cell, forming first a clathrin-coated vesicle and then an uncoated vesicle, or *endosome* (Figure 6.40). Subsequently, the LDL particle degrades, and the receptors recycle to the plasma membrane. Accumulation of intracellular cholesterol inhibits the synthesis of LDL receptors and their replenishment at the cell membrane, thereby blocking further uptake of cholesterol. The decrease in activity of cell membrane receptors produced by reaction with their specific ligands is called *down regulation.*

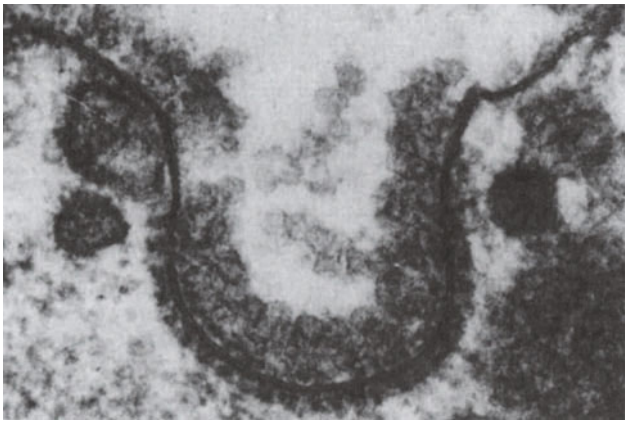


Figure 6.39. Electron micrograph of a coated pit in the plasma membrane of a chicken oocyte. Many LDL particles are bound in the pit, and a coat of clathrin molecules is located below the pit. [Reprinted, with permission, from R. G. W. Anderson, M. S. Brown, and J. L. Goldstein, *Cell* 10: 351–364 (1977).]

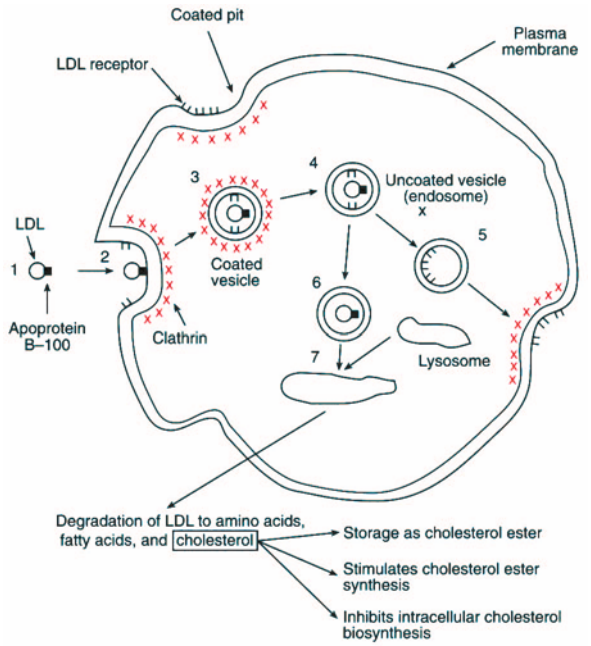


Figure 6.40. The sequence of events in receptor-mediated endocytosis of LDL. LDL (1) binds to a clathrin-lined coated pit (2) that pinches off as a coated vesicle (3) and is converted to an uncoated one (4) by removal of clathrin. Uncoated vesicles can give rise to receptor-containing (5) and LDL-containing (6) vesicles. The former are recycled to the cell membrane; the latter fuse with lysosomes (7) and lead to LDL degradation.

Cholesterol has two other regulatory functions. It activates an enzyme that catalyzes the conversion of cholesterol to cholesterol esters, which are then stored in the form of lipid droplets. At high cholesterol concentration, cholesterol derivatives inhibit an enzyme involved in intracellular cholesterol biosynthesis. This avoids cellular overproduction of cholesterol and ensures preferential use of LDL-derived cholesterol.

Investigators currently believe that a deficiency of LDL receptors constitutes the primary cause of atherosclerosis. Such a deficiency leads to increased plasma levels of LDL. Circulating LDL invades arterial linings and participates in plaque formation.

Strong evidence in support of this theory is provided by patients afflicted with a hereditary disease called **familial hypercholesterolemia.** Individuals in the homozygous state have extremely high levels of plasma LDL-cholesterol (three to five times the normal value) and such an accelerated degree of atherosclerosis that the disease usually becomes fatal by age 20. We now know that these individuals suffer from a complete lack of functional LDL receptors.

The effect of HDL ("good" cholesterol) is opposite to that of LDL. Increased plasma levels of HDL are associated with decreased incidence of atherosclerosis. Researchers believe that HDL functions as cholesterol scavenger, removing it from peripheral tissues and transporting it back to the liver. Many factors known to decrease the incidence of cardiovascular disease—such as weight loss and exercise—correlate with increased levels of HDL. Smoking, a factor that enhances the incidence of the disease, correlates with lowering the level of HDL. The reasons for these various effects are unknown.

Receptor-mediated endocytosis, or *internalization of receptors,* applies not only to endocytosis of LDL but also to the uptake by cells of, and down regulation by, a number of other substances like hormones, growth factors, toxins, and viruses.

SUMMARY

Lipids comprise a diverse group of water-insoluble (or sparingly soluble) substances that constitute a storage form of energy, serve as structural components of biological membranes, and have a number of protective roles. Simple lipids are esters of fatty acids and alcohols. Fats form from fatty acids

and glycerol; waxes from fatty acids and other alcohols. Saponification of fats yields glycerol and salts of fatty acids, or soaps. Soap molecules form micelles, stabilize emulsions, and have a cleansing action. Synthetic detergents have structural similarity to soaps and comparable properties.

Complex lipids have a nonlipid component in addition to their lipid moiety. Typical ones include phospholipids, glycolipids, and lipoproteins. Glycerophospholipids and sphingophospholipids constitute phospholipids derived from, respectively, phosphatidic acid and sphingosine. These phospholipid molecules are amphipathic and have a polar head and two nonpolar tails. Phospholipids have properties like those of soaps and detergents and occur as major constituents of the lipid bilayer of biomembranes.

Derived lipids constitute a heterogeneous group of substances having many different functions. Major derived lipids include prostaglandins and related compounds that have various hormonal effects; isoprenoids such as carotenoids and terpenoids; fat-soluble vitamins (A, D, E, and K); and steroids, a very large group of compounds that includes cholesterol, adrenal cortical steroids, and sex hormones.

We currently describe biomembranes by the fluid mosaic model, according to which a membrane consists of a lipid bilayer, composed primarily of phospholipids oriented so that their polar heads are at the two surfaces of the membrane. Proteins occur at the surfaces of the membrane and distributed within the membrane, extending to varying degrees into the lipid matrix. Biomembranes are fluid, allowing for some lateral movement of lipids and proteins and some transverse movement of lipids. Biomembranes are asymmetric assemblies, so that membrane processes have directionality.

There exist three modes of membrane transport: passive transport consists of simple diffusion; facilitated transport requires a mobile carrier or a passage (pore or channel) through the membrane; and active transport entails moving a substance against its concentration gradient and requires the expenditure of energy. Among other functions, the cell membrane plays an important role in transmitting nerve impulses and in atherosclerosis.

SELECTED READINGS

Bach, T. J., Some new aspects of isoprenoid biosynthesis in plants—a review, *Lipids* 30:191–202 (1995).

Bell, G. I., Burant, C. F., Takeda, J., and Gould, G. W., Structure and function of mammalian facilitative sugar transporters, *J. Biol. Chem.* 268:19161–19164 (1993).

Catteral, W. A., Structure and function of voltage-gated channels, *Annu. Rev. Biochem.* 64:493–531 (1995).

Darwish, H. M., and DeLuca, H. F., Recent advances in the molecular biology of vitamin D action, *Prog. Nucleic Acid Res. Mol. Biol.* 53:321–344 (1996).

Goodenough, A., Goliger, J. A., and Paul, D. L., Connexins, connexons, and intercellular communication, *Annu. Rev. Biochem.* 65:475–502 (1996).

Hoberman, J. M., and Yesalis, C. E., The history of synthetic testosterone, *Sci. Am.* 272:76–81 (1995).

Khorana, H. G., Rhodopsin, photoreceptor of the rod cell. An emerging pattern for structure and function, *J. Biol. Chem.* 267:1–4 (1992).

Lawn, R. M., Lipoproteins in heart disease, *Sci. Am.* 266:54–60 (1992).

Liscum, L., and Underwood, K. W., Intracellular cholesterol transport and compartmentation, *J. Biol. Chem.* 270:15443–15446 (1995).

Moolenaar, W. H., Lysophosphatidic acid, a multifunctional phospholipid messenger, *J. Biol. Chem.* 270:12949–12952 (1995).

Petty, H. R., *Molecular Biology of Membranes: Structure and Function,* Plenum, New York (1993).

Singer, S. J., and Nicolson, G. L., The fluid mosaic model of cell membranes, *Science* 175:720–731 (1972).

Smith, W. L., Garavito, R. M., and DeWitt, D. L., Prostaglandin endoperoxidase H synthases (cyclooxygenases)-1 and -2, *J. Biol. Chem.* 271:33157–33160 (1996).

Stein, W. D., *Channels, Carriers, and Pumps: An Introduction to Membrane Transport,* Academic Press, San Diego (1990).

Vance, D. E., and Vance, J. E. (eds.), *Biochemistry of Lipids, Lipoproteins, and Membranes,* Elsevier, Amsterdam (1991).

REVIEW QUESTIONS

A. Define each of the following terms:

Acylglycerol	Glycerophospholipid
Sphingophospholipid	Soap
Saponification number	Iodine number
Ionophore	Transmembrane protein
Lipid bilayer	Sex hormones
Primary active transport	Secondary active transport
Group translocation	Visual cycle

B. Differentiate between the two terms in each of the following pairs:

Active transport/cotransport	Fat/wax
Simple lipid/complex lipid	Lipoprotein/proteolipid
Diffusion/facilitated transport	Integral protein/peripheral protein
Bile acids/bile salts	Sterols/steroids

C. (1) What are the major mechanisms by which substances are transported across biological membranes?
(2) Describe the operation of the sodium–potassium pump.
(3) What are the main steps in receptor-mediated endocytosis?
(4) Discuss the structure and properties of biomembranes according to the fluid mosaic model.
(5) Give the names and write out the structures of representative compounds from the major types of simple, complex, and derived lipids.
(6) List the ways in which active transport is comparable to the reaction of an enzyme with its substrate.

PROBLEMS

6.1. Calculate the average density of the human red blood cell membrane (Table 6.4) on the assumption that the densities of protein, lipid, and carbohydrate are 1.33, 0.92, and 1.48 g/ml, respectively.

6.2. Write balanced equations, using structural formulas, for: (a) the saponification of tripalmitin with potassium hydroxide; (b) the formation of an optically inactive monoacylglycerol from glycerol and oleic acid; (c) the formation of phosphatidyl tyrosine from phosphatidic acid and tyrosine.

6.3. How many different triacylglycerols can be made with *sn*-glycerol and three different fatty acids? How many of these triacylglycerols have optical activity?

6.4. Which of the following are saponifiable lipids?

Vitamin A	Vitamin E	Vitamin K
Lecithin	Cephalin	Sphingomyelin
Cholesterol	Cholic acid	Phosphatidic acid
Ceramide	Bile salts	Ganglioside GM_2

6.5. Calculate the theoretical saponification number for the triacylglycerol of arachidic acid (MW of KOH = 56.0).

6.6. Calculate the theoretical iodine number for the optically inactive monoacylglycerol of arachidonic acid (atomic weight of I = 127).

6.7. Which compound in each of the following pairs would you expect to be less soluble in polar solvents?

Cholesterol/deoxycholic acid
Phosphatidic acid/phosphatidyl serine
Sphingosine/sphingomyelin

6.8. What might happen if you dissolved sodium palmitate, at a high concentration, in chloroform?

6.9. How many milliliters of 1.0*M* KOH do you need to hydrolyze off the fatty acid contained in 2.4 g of sphingomyelin (MW = 730)?

6.10.* At what pH would you be able to separate phosphatidyl serine, phosphatidyl ethanolamine, and phosphatidyl choline by electrophoresis (see Appendix C): pH 1.0, pH 7.0, or pH 12.0? Give the net charge of each compound at the selected pH, and indicate whether it moves toward the anode or the cathode. See Table 2.1 for the pK_a' values of serine. The pK_a' of the amino group of ethanolamine is 9.5, and that of the phosphate group in glycerophospholipids may be taken as 2.1.

6.11.* It takes exactly 20 ml of 0.2*M* KOH to fully saponify 1.2 g of an acylglycerol with a molecular weight of 600. Is the compound a mono-, a di-, or a triacylglycerol?

6.12.* Calculate the number of double bonds in an acylglycerol that has a molecular weight of 878 and an iodine number of 174 (atomic weight of I = 127).

6.13. One membrane, Mb-I, contains only saturated fatty acids, and a second one, Mb-II, contains 30% unsaturated fatty acids. Which membrane do you expect to be more sensitive to disruption by an increase in temperature? To disrupt Mb-I and Mb-II by chemical means, would you choose (a) sodium dodecyl sulfate (SDS), (b) urea, or (c) glucose? Explain your answers.

6.14. The ATP-binding site of the sodium pump is located inside the cell. What types of amino acid residues might constitute that site, and why?

6.15.* The efficiency of a gramicidin A channel across a cell membrane is such that 10^7 K^+ ions are transported per second from the inside to the outside. The intracellular volume is 2.0×10^{-12} ml, and the initial intracellular concentration of K^+ is 0.10*M*. What will be the intracellular concentration of K^+ after the channel has been in operation for five seconds?

6.16.* A red blood cell's surface area is approximately 100 μm^2. The membrane lipids isolated from 4.74×10^9 red blood cells, when spread over water, form a *monolayer* that covers an area of 0.890 m^2. Show that the membrane covering the red blood cell must be a bilayer; that is, the membrane must be two lipid molecules thick.

6.17.* Given that one phospholipid molecule occupies 0.70 nm^2, calculate the number of phospholipid molecules in a section of a phospholipid bilayer that has a surface area (on one side of the bilayer) of 1.0 μm^2?

6.18. Assume that you find two types of margarine in the supermarket, both made essentially from 100% corn oil. One margarine is made by hydrogenating the corn oil, and the other by emulsifying it. Which of the two margarines would be the better source of unsaturated fatty acids?

6.19. Why might some fat-soluble vitamins be toxic at high doses whereas water-soluble vitamins (like vitamin C) can be taken in megadoses without any apparent complications?

6.20. Calculate the molar ratio of lipid to protein for the membrane of *Halobacterium* (Table 6.4), assuming average molecular weights of 800 for the lipid and 50,000 for the protein.

6.21. What is the N/P ratio in terms of numbers of atoms for each of the following? (a) lecithin; (b) cephalin; (c) sphingomyelin

6.22. Which of the following do you expect bile salts to form? (a) micelles; (b) monolayers; (c) bilayers. Explain your answers.

Nucleic Acids

7

The study of nucleic acids began in 1869 with the isolation of a hitherto unknown phosphate-rich substance from the nuclei of white blood cells. The discovery was made by Friedrich Miescher, a young Swiss studying with the great German biochemist Ernst Hoppe-Seyler. Miescher called the isolated substance *nuclein*. Subsequently, nuclein was shown to be acidic in nature, and its name was changed to **nucleic acid.** Soon it became apparent that there were two types of nucleic acids, now called **ribonucleic acid (RNA)** and **deoxyribonucleic acid (DNA)**. In the early days of nucleic acid research, investigators frequently isolated RNA from yeast and DNA from the thymus gland. Hence, RNA and DNA are referred to in the older literature as *yeast nucleic acid* and *thymus nucleic acid,* respectively.

Although DNA and RNA were studied extensively, it was not until the middle of the 20th century that specific biological functions were assigned to them. In 1953, James Watson and Francis Crick proposed a double-helical model for the structure of DNA. The model, for which both received the Nobel Prize in 1962, immediately suggested a molecular basis for the transfer of genetic information. Acceptance of the double-helix structure for DNA has revolutionized research in biochemistry and related sciences. It is now firmly established that DNA is the genetic material in all bacterial, plant, and animal organisms. (Viruses contain either DNA or RNA as their genetic material.) DNA stores hereditary information in the form of genes and transmits that information from generation to generation.

The study of DNA, RNA, and protein synthesis, called **molecular biology,** has progressed at an explosive rate over the last 50 years. The discoveries and applications of molecular biology undoubtedly constitute one of the most astounding chapters in the entire history of science.

7.1. STRUCTURAL COMPONENTS

7.1.1. Purines and Pyrimidines

Both DNA and RNA are polymers containing only a few types of building blocks called **nucleotides.** Each nucleotide consists of three chemical parts—a sugar, a phosphate group, and a base. The sugar of RNA is *D-ribose* and that of DNA is *D-2-deoxyribose* (see Figures 5.5 and 5.11). The phosphate group, in both RNA and DNA nucleotides, is esterified to the sugar. The **base,** or **nitrogen(ous) base,** occurs in two forms—as a **purine** (**Pu**) or a **pyrimidine**

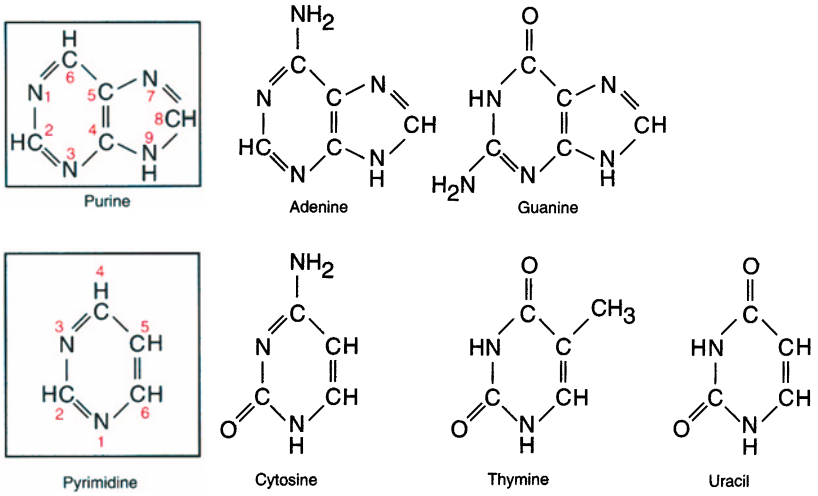


Figure 7.1. Major purines and pyrimidines in DNA and RNA.

(**Py**)—so designated by reference to their respective parent compounds (Figure 7.1).

There are two major purines, **adenine (A)** and **guanine (G),** and three major pyrimidines, **cytosine (C), thymine (T),** and **uracil (U)** (Figure 7.1). Adenine, guanine, and cytosine occur in both RNA and DNA. Uracil, however, is found only in RNA, and thymine occurs mainly in DNA. Thus, on the basis of overall composition, RNA and DNA differ in two respects—the type of sugar and the presence of either uracil or thymine. A number of **modified bases** occur in some nucleic acid structures (Figure 7.2). Many of these represent methylated derivatives of common bases.

The bases have a number of properties that relate to their roles in nucleic acid structure and function. First of all, the bases are *planar.* Both purine and pyrimidine rings are essentially flat because of the presence of conjugated double bonds. In addition, purines and pyrimidines are *weak bases* in the acid–base sense and can participate in various proton equilibria (Figure 7.3).

All of the bases undergo tautomerism (an equilibrium of rapidly interconvertible structural isomers). In purines and pyrimidines, this takes the form of an *enol–keto tautomerism,* involving the shift of a hydrogen. Figure 7.4 shows the corresponding two isomers of uracil. Other bases undergo similar tautomeric shifts. Keto configurations of the bases constitute their biologically important structures; under physiological conditions (pH 7.0), the bases exist predominantly in their keto forms.

Lastly, purines and pyrimidines *absorb in the ultraviolet region* because of their aromatic structures (Figure 7.5). The absorption spectra vary as a function of pH. Nucleic acids and their components absorb maximally at or near 260 nm (in contrast to proteins, which have maximum absorption at about 280 nm), and nucleic acids absorb more strongly at 260 nm than proteins do at 280 nm. Researchers frequently use ultraviolet absorption to analyze nucleic acid components.

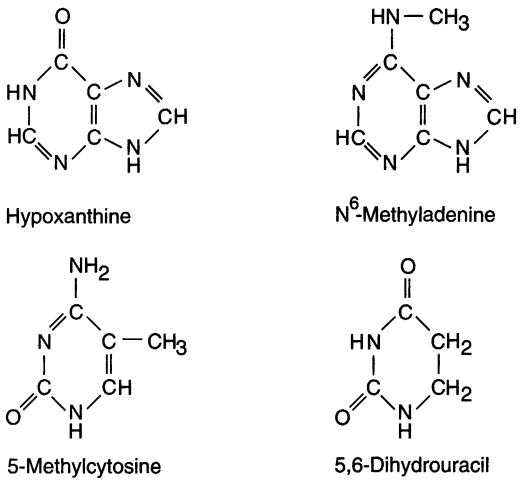


Figure 7.2. Some modifed bases of nucleic acids.

7.1.2. Nucleosides

A sugar and a base combine to form a **nucleoside.** In this structure, the sugar is linked to the base via a *β-glycosidic bond* involving the anomeric carbon of the sugar and N(9) of a purine or N(1) of a pyrimidine (Figure 7.6). By convention, we indicate positions in the sugar moiety of nucleosides and nucleotides by a prime to differentiate them from positions in the attached base. The glycosidic bond of a nucleoside is the same whether formed with ribose or deoxyribose. The configuration at C(2′) of the sugar has no effect on formation of the glycosidic bond at the anomeric carbon, C(1′). However, the nature of the base

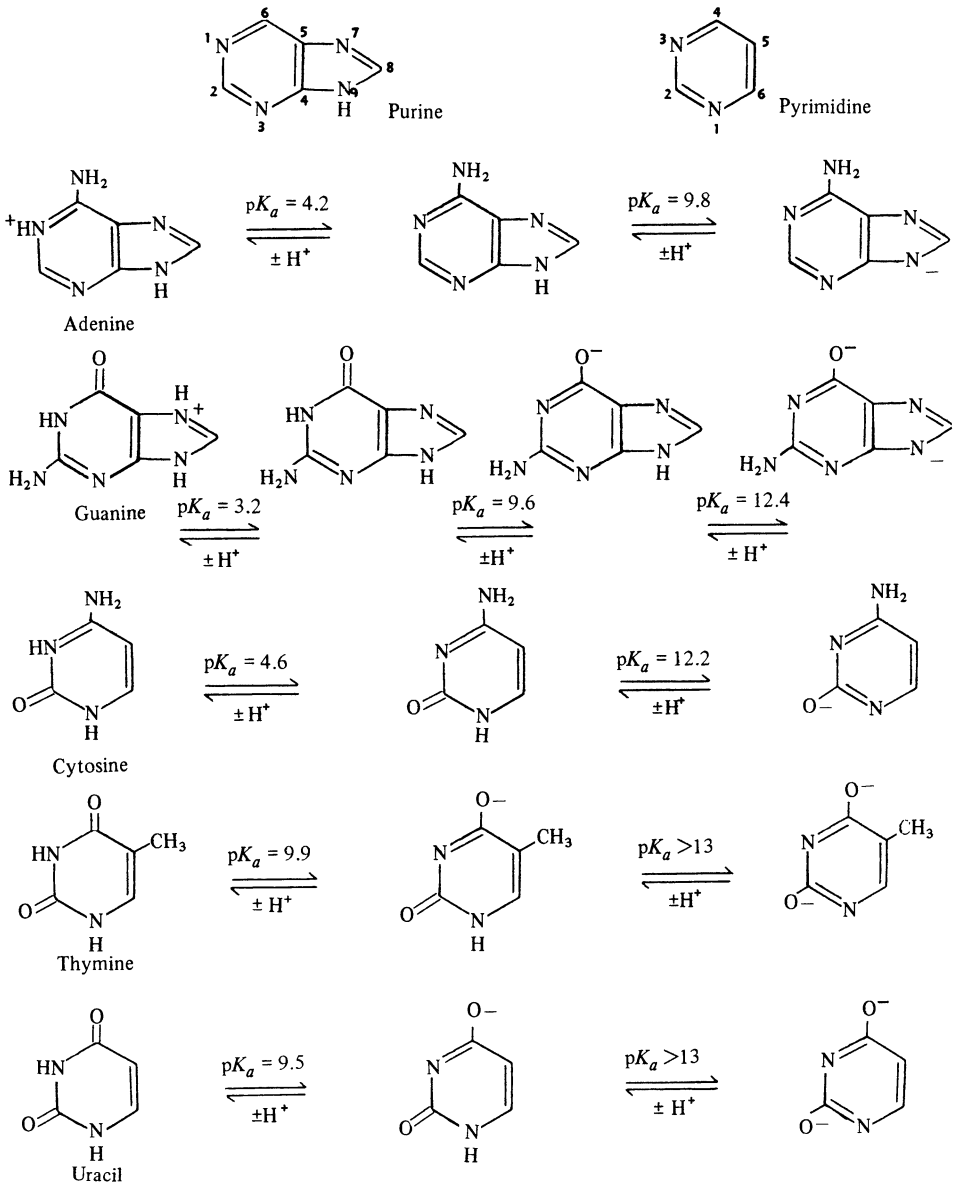


Figure 7.3. Proton equilibria of the common purines and pyrimidines. [Reprinted, with permission, from J. Stenesh, *Experimental Biochemistry*, Allyn and Bacon, Boston (1984).]

does have steric implications for the combination of base and sugar.

For purines, there exist two permissible orientations, *anti* and *syn* (Figure 7.7). Most nucleic acid structures involve the less hindered *anti* configuration of purine nucleosides. For pyrimidines, only the *anti* conformation can readily form; the *syn* conformation is difficult to achieve because of the steric hindrance due to the substituent at C(2) of the pyrimidine nucleus.

We name nucleosides as outlined in Table 7.1. Purine nucleosides have the ending "osine" (from pur*ine*), and

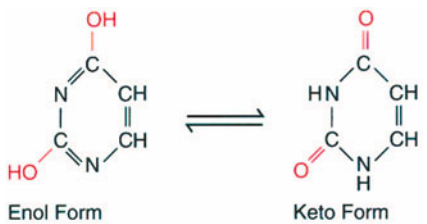


Figure 7.4. Enol–keto tautomerism of uracil.

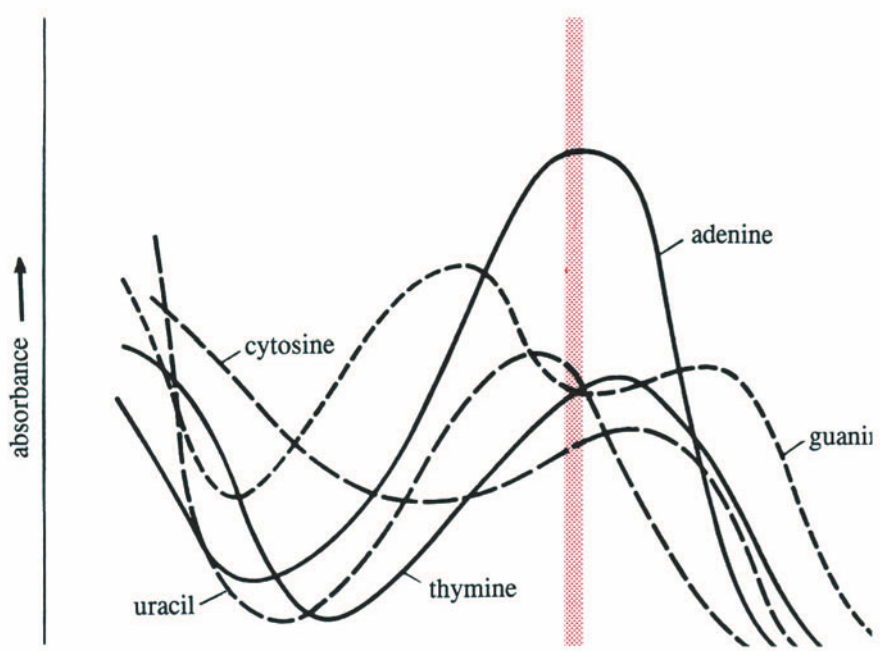


Figure 7.5. Ultraviolet absorption spectra of equimolar solutions of the major purines and pyrimidines at pH 7.0. [Reprinted, with permission, from J. N. Davidson, *The Biochemistry of the Nucleic Acids,* 8th ed., Academic Press, New York (1976).]

pyrimidine nucleosides have the ending "idine" (from pyrim*idine*). Note that investigators originally believed that thymine occurred only in DNA. Therefore, they dropped the prefix "deoxy," and thymidine came to mean deoxythymidine. We now know that thymine does occur occasionally in RNA. Accordingly, we explicitly designate the ribonucleoside of thymine by the prefix "ribo" or "r" as either *ribothymidine* or *rT.*

A number of antibiotics produced by species of *Streptomyces* contain a nucleoside or nucleoside derivative. These antibiotics, termed *nucleoside antibiotics,* inhibit the growth of microorganisms by blocking purine and pyrimidine metabolism. A synthetically produced nu-

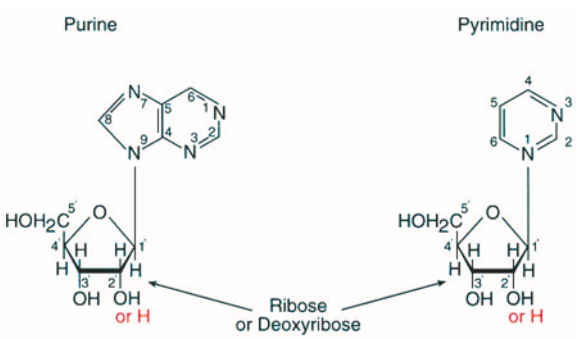


Figure 7.6. Structure of a nucleoside.

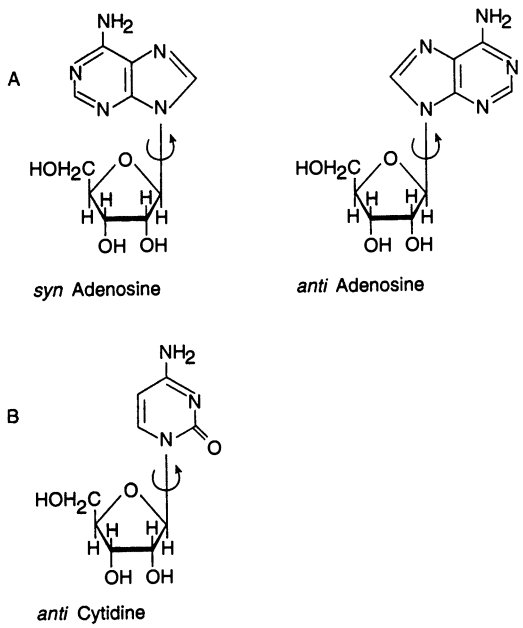


Figure 7.7. Stereochemical orientations of bases about the glycosidic bond in nucleosides and nucleotides. (A) Purines; (B) pyrimidines.

Table 7.1. Names and Abbreviations of Nucleosides

Base	Nucleoside[a]	Abbreviation[b]
Adenine	Adenosine	A
Guanine	Guanosine	G
Cytosine	Cytidine	C
Thymine	Thymidine	T
Uracil	Uridine	U

[a] Deoxyribonucleosides are indicated by the prefix "deoxy" or "d."
[b] Single-letter abbreviations are generally used for bases but occasionally also for nucleosides.

cleoside that has received a great deal of recent attention is *3′-azido-2′,3′-dideoxythymidine* (*AZT*) (Figure 7.8). AZT is used to treat individuals afflicted with AIDS (acquired immune deficiency syndrome).

7.1.3. Nucleotides

A nucleotide forms when a molecule of phosphoric acid becomes esterified to the sugar of a nucleoside. Nucleotides differ depending on the sugar position at which esterification occurs. RNA and DNA nucleotides have the phosphate group attached at C(5′) of ribose or deoxyribose; they are **5′-nucleotides** (Figure 7.9A). We can summarize the different possibilities for esterifying phosphoric acid to a nucleoside as follows:

1. Ester formation for ribonucleosides can occur at the 2′, 3′, or 5′ position of the sugar, but for deoxyribonucleosides it can only occur at positions 3′ and 5′.
2. Ester formation can occur at more than one position on the same ribo- or deoxyribonucleoside (Figure 7.9B).
3. Any ribo- or deoxyribonucleoside can have a single phosphoric acid group esterified at two positions on the sugar, forming a cyclic phosphodiester (Figure 7.9C).

Note that only the 5′-nucleotides (*nucleoside 5′-monophosphates*) serve as building blocks of nucleic

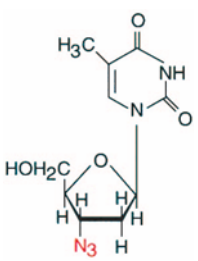


Figure 7.8. The drug 3′-azido-2′,3′-dideoxythymidine (AZT).

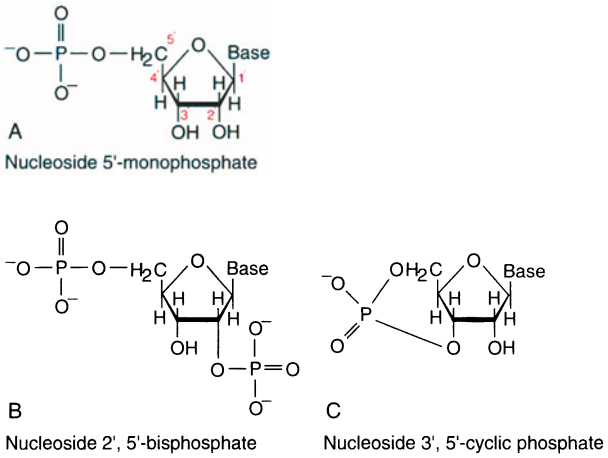


Figure 7.9. Examples of phosphoric acid esters of nucleosides.

acids. 5′-Ribonucleotides constitute the structural components of RNA, and 5′-deoxyribonucleotides those of DNA. Other types of nucleotides function in various capacities in metabolism but not as nucleic acid building blocks.

Some nucleotide derivatives have the phosphate group extended by anhydride linkages so that two or even three phosphate groups are linked in series at one position on the sugar (Figure 7.10). Nucleoside di- and triphos-

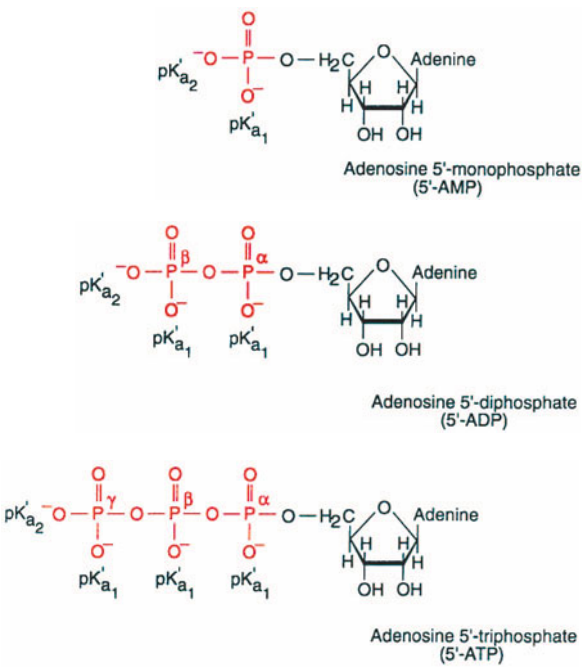


Figure 7.10. A 5′-nucleotide and two derivatives formed by joining additional phosphate groups through anhydride bonds. Multiple phosphate groups (designated α, β, and γ) have pK'_a values that approximate pK'_{a_1} and pK'_{a_2} of H_3PO_4.

Table 7.2. Names and Abbreviations of Nucleotides[a]

System I, for mono,di-,tri-, and cyclic phosphates[b]	System II, only for monophosphates[b]	Abbreviations
Adenosine (′) . . . . phosphate	(′) Adenylic acid	AMP, ADP, ATP
Guanosine (′) . . . . phosphate	(′) Guanylic acid	GMP, GDP, GTP
Cytidine (′) . . . . . phosphate	(′) Cytidylic acid	CMP, CDP, CTP
Thymidine (′) . . . . phosphate	(′) Thymidylic acid	TMP, TDP, TTP
Uridine (′) . . . . . phosphate	(′) Uridylic acid	UMP, UDP, UTP

phates play important roles in energy metabolism. Nucleoside triphosphates also serve as reactants for nucleic acid biosynthesis: each provides a 5′-nucleoside monophosphate building block while the remaining phosphate is cleaved out as *pyrophosphate* ($P_2O_7^{4-}$ or PP_i).

Table 7.2 outlines two systems of nomenclature for various types of phosphorylated nucleosides. We can use system I for any phosphorylated derivative; system II is strictly for naming nucleoside monophosphates. Some examples follow:

- Adenosine 5′-triphosphate (5′-ATP); three phosphate groups at C(5′) of adenosine
- Adenosine 3′,5′-cyclic monophosphate, or 3′,5′-cyclic adenylic acid (3′,5′-cyclic AMP); one phosphate group esterified twice, at C(3′) and C(5′) of adenosine
- Guanosine 3′-diphosphate (3′-GDP); two phosphate groups at C(3′) of guanosine
- Cytidine 5′-monophosphate (5′-CMP), or 5′-cytidylic acid; one phosphate group at C(5′) of cytidine
- Deoxyuridine 3′,5′-*bis*phosphate; one phosphate group at C(3′) and one phosphate group at C(5′) of deoxyuridine

Note that we use the prefixes *bis* and *tris* if two or three phosphate groups are esterified to different carbons, but we use the prefixes *di* and *tri* if the groups are attached through anhydride bonds to a single carbon. Of greatest importance in biochemical systems are 5′-nucleotides and related compounds. For these, we customarily delete the 5′ designation unless special emphasis is needed. Thus, ATP is adenosine 5′-triphosphate, and GDP is guanosine 5′-diphosphate.

As a final note to this discussion, we need to point out that biochemists use the term *nucleotide* in two ways. Primarily, we use the term to describe a phosphorylated nucleoside, particularly one phosphorylated at the 5′ position and serving as a nucleic acid building block. Additionally, we use the term for a number of compounds that have *some* of the chemical parts of nucleotides in their structure. Examples are *nicotinamide adenine dinucleotide (NAD⁺)* and *flavin mononucleotide (FMN),* both of which you will encounter in Section 11.1.

7.1.4. Basic Nucleic Acid Structure

7.1.4A. 3′,5′-Phosphodiester Bond. Nucleic acids form by sequential linkage of nucleotides, much as proteins form by sequential linkage of amino acids. Nucleotides are joined by means of their phosphate groups. Each phosphate group has two functions: it constitutes a structural part of a given nucleotide, and it serves to link that nucleotide to its nearest neighbor. Phosphoric acid groups function this way by being esterified twice—once to the 5′ position of a given nucleotide and once to the 3′ position of the adjacent nucleotide (Figure 7.11).

Each phosphoric acid residue forms a *phosphodiester,* and we call the internucleotide linkage a **3′,5′-phosphodiester bond.** The role of the 3′,5′-phosphodiester bond in nucleic acid chemistry is analogous to that of the peptide bond in protein chemistry. Successive linking of nucleotides by means of 3′,5′-phosphodiester bonds produces chains or **strands.** We designate polymers of increasing length by the prefixes "oligo" and "poly," as for proteins and carbohydrates. Thus, we refer to polymers consisting of 2–10 nucleotides as oligonucleotides, and those consisting of >10 nucleotides as polynucleotides.

We call large, naturally occurring polynucleotides *nucleic acids. Transfer RNAs,* composed of about 80 nucleotides, represent some of the smallest known nucleic acids. Since the average molecular weight of a nucleotide is approximately 300, transfer RNAs have molecular weights of about 24,000.

7.1.4B. Shorthand Notations. It is obviously impractical to write out lengthy sequences of nucleotides

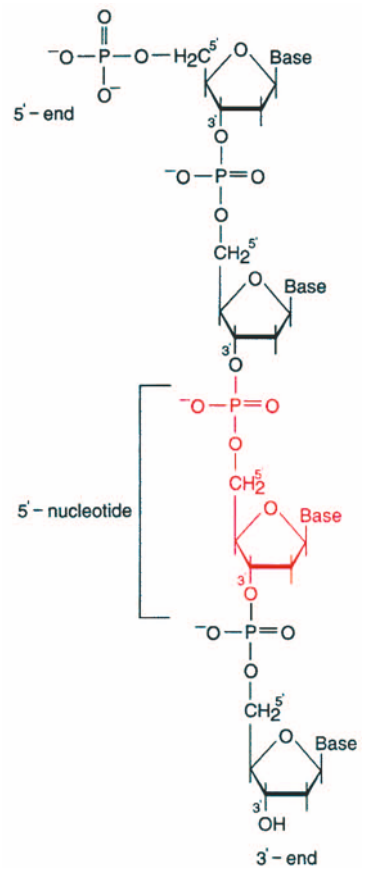


Figure 7.11. The polynucleotide strand in ribo- and deoxyribonucleic acids. The 3′- and 5′-ends may carry a hydroxyl or a phosphate group.

using the format of Figure 7.11. Hence, several shorthand notations have come into use. The convention depicted in Figure 7.12 is useful for describing the action of enzymes that catalyze hydrolysis of the phosphodiester bond on either the 3′ or the 5′ side (type *a* or type *b* cleavage).

For large polynucleotides, we further simplify the representation of nucleotide sequences. To do so, we write the nucleotide sequence shown in Figure 7.12 as

ApCpGpTpA

When using this notation, base symbols represent the corresponding *nucleosides.* Phosphoester bonds are indicated by writing the "p" to the right or left of the base symbol, so that

(Base)p = 3′-linkage

p(Base) = 5′-linkage

and a "p" between two base symbols represents a 3′,5′-phosphodiester bond. Thus, ApU refers to the dinucleotide (A-3′)-p-(5′-U), and pApGpTpCpCpApA describes a heptanucleotide having a free 5′-phosphate group. We always write oligo- and polynucleotide sequences with the 5′-end on the left and the 3′-end on the right (similar to the N- and C-termini of a polypeptide chain). The 5′-end of a polynucleotide strand may or may not have a phosphate group attached. The 3′-end of the strand generally carries a free 3′-OH group.

We simplify the notation even further by altogether omitting the "p" between bases. Biochemists understand that writing AU does not mean that *adenine* is linked directly to *uracil* but rather that *adenosine* is linked to *uridine* via a 3′,5′-phosphodiester bond. Specifically, the phosphate group links the 3′ position of *adenosine* and the 5′ position of *uridine.* When using this shorthand representation, you can indicate deoxyribonucleotide sequences by using the prefix "deoxy" or "d." However, the prefix is often omitted when you are clearly dealing with DNA sequences.

7.1.4C. Sizes of Nucleic Acids.

Nucleic acids occur as single-stranded (ss) and double-stranded (ds) molecules. A nucleic acid strand can be linear or circular. These terms do not refer to straight lines or geometrical circles but rather to open-chain and closed-chain conformations.

Nucleic acids vary greatly in their size. DNA size increases from viruses to prokaryotes to eukaryotes. Some DNAs contain an enormous number of base pairs and are very long (Table 7.3). To get a feel for these numbers, consider that the length of DNA in a human (haploid) cell is about one meter. Since the human body contains approximately 10^{13} cells, the total length of an individual's DNA is about 1×10^{13} m or 1×10^{10} km. By comparison, the circumference of the Earth has a length of 4×10^4 km, and the distance from the Earth to the Sun equals 1.44 ×

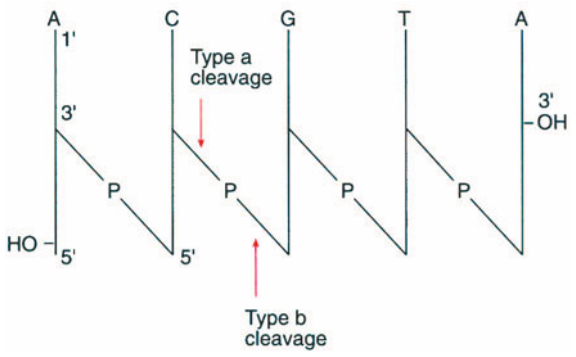


Figure 7.12. A shorthand notation for oligonucleotides. A vertical line represents a pentose, P designates a phosphate group, and bases are designated by their one-letter abbreviations.

Table 7.3. Characteristics of Some DNA Molecules[a]

Source	Number of base pairs (kbp)[b]	Contour length (μm)[c]	Structure
Viruses			
Polyoma, SV40	5.1	1.7	Circular ds[d]
Phage lambda	48.6	17	Linear ds
Phage T2, T4, T6	166	55	Linear ds
Bacteria			
Mycoplasma hominis	760	260	Circular ds
Escherichia coli	4,000	1,360	Circular ds
Eukaryotes			
Yeast	13,500	4,600	17 Chromosomes (haploid)
Fruit fly	165,000	56,000	4 Chromosomes (haploid)
Human	2,900,000	990,000	23 Chromosomes (haploid)
Lungfish	102,000,000	34,700,000	19 Chromosomes (haploid)

[a]Adapted with permission from A. Kornberg, *DNA Replication*, 2nd ed., W. H. Freeman, New York (1990).
[b]bp, Base pair; kbp, kilo base pair = 1000 base pairs.
[c]Contour length, End-to-end length of the extended molecule.
[d]ds, Double-stranded.

10^8 km. Thus, your total DNA, if stretched out, will have a length equal to about 70 times the distance between the Earth and the Sun! Compared to DNA, RNA usually contains a smaller number of nucleotides and covers a smaller range of sizes (see Table 7.4).

7.1.5. Occurrence and Functions of Nucleic Acids

Both DNA and RNA occur in free form and in association with proteins. Because of their acidic nature (resulting from the many phosphate groups per strand), nucleic acids associate preferentially with basic proteins. Nucleic acids and proteins become linked via ionic interactions to form a variety of **nucleoproteins.**

7.1.5A. Nucleoproteins. One major group of nucleoproteins results from association of DNA with proteins. The association is relatively nonspecific in prokaryotes and produces a non-membrane-bound nuclear region, or *nucleoid.* In eukaryotes, the association is specific, involves basic proteins called histones, and gives rise to *nucleosomes,* located in the nucleus (more in Section 7.5). Nucleosomes serve as structural elements of *chromatin,* the threadlike material of the chromosomes. Most of cellular DNA is located in either the nuclear region or the nucleus.

Association of RNA with proteins in the form of *ribosomes* constitutes a second main group of nucleoproteins. Ribosomes are subcellular particles, located mainly in the cytoplasm, that serve as sites for protein synthesis. Most of cellular RNA is located in the cytoplasm in the form of ribosomal RNA.

Viruses constitute a third group of important nucleoproteins. The *genome* (complete set of genes) of some viruses consists of DNA; that of others consists of RNA. Viral nucleic acid cores are surrounded by *protein coats.* Scientists consider viruses a border form of life because viruses cannot replicate independently, only in a suitable host. Some viruses infect animals, others infect plants, and still others infect bacteria. A virus that infects bacteria is called a *bacteriophage* or *phage.*

7.1.5B.The Varied Functions of RNA. While DNA has a single critical function as repository and carrier of genetic information, RNA has a number of functions. RNA occurs in three major forms, all of which participate in protein biosynthesis but have different roles. **Ribosomal RNA (rRNA)** serves as a structural component of ribosomes; **transfer RNA (tRNA)** functions as an adapter that transports amino acids to the ribosomes for incorporation into protein; and **messenger RNA (mRNA)** is a transcribed copy of DNA that contains the information for synthesizing specific proteins (Table 7.4).

As noted above, RNA also serves as the genetic material of some viruses. One of the best studied RNA viruses is the *tobacco mosaic virus (TMV),* a virus that infects tobacco leaves (Figure 7.13). This virus has a molecular weight of 40 million and contains one single-stranded RNA molecule (6400 nucleotides) surrounded by a protein coat (capsid) composed of 2130 protein molecules (MW = 17,500 each).

Table 7.4. Sizes of Some RNA Molecules

Type of RNA	Number of nucleotides
Viral RNA	
Avocado sun blotch viroid	247
Potato spindle tuber viroid	359
Tobacco mosaic virus (TMV)	6400
Polio virus	7440
Transfer RNA	
$tRNA^{Leu}$ (*Escherichia coli*)	87
$tRNA^{Phe}$ (yeast)	76
$tRNA^{Ser}$ (rat)	85
Ribosomal RNA	
rRNA (*Escherichia coli*)	120; 1542; 2904
rRNA (rat)	160; 1874; 4718
Messenger RNA	
mRNA (lysozyme)	~600
mRNA (serum albumin)	~2000

Recently, researchers have identified small, viruslike infectious particles that consist of single-stranded RNA and do not have a protein coat. These "naked" RNA molecules, called **viroids,** represent the smallest known infectious nucleic acids. Viroids cause several plant diseases.

A number of RNAs, termed **ribozymes,** possess enzymatic activities. *Ribonuclease P* and *precursor ribosomal RNA* of the protozoan *Tetrahymena thermophila* are two examples. Ribonuclease P is a bacterial "enzyme" that catalyzes the hydrolysis of an oligonucleotide fragment from precursor tRNA, thereby producing the 5′-terminus of mature tRNA. Ribonuclease P contains 86% RNA and 14% protein, and the catalytic activity resides entirely in the RNA; the protein serves merely as molecular scaffolding. Precursor ribosomal RNA of *Tetrahymena* catalyzes a "self-splicing" reaction. A 413-residue polynucleotide fragment is excised from the precursor, followed by a joining of the ends of the remaining molecule to produce a mature ribosomal RNA. The entire reaction proceeds in the absence of protein. The discovery of ribozymes supports the view of some researchers that life began with nucleic acids, with proteins evolving later.

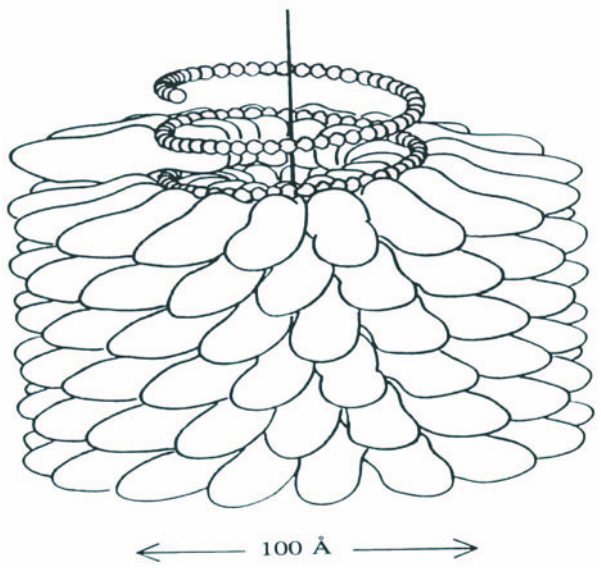


Figure 7.13. The tobacco mosaic virus (TMV). [Reprinted, with permission, from A. Klug and L. D. Caspar, *Adv. Virus Res.* 7:225–325 (1960).]

7.2. PRIMARY STRUCTURE

We describe the structure of nucleic acids in terms of different levels like those used for proteins. Three levels of structure—primary, secondary, and tertiary—are applicable. The primary structure comprises the *type, number, and sequence of nucleotides* in the polynucleotide strand. Because the sugar and phosphate components are fixed for a given type of nucleic acid, the *sequence of the bases* along the strand constitutes the actual variable of the primary structure. Investigators determine the type and number of bases in a nucleic acid by hydrolysis, followed by chromatography. Determining the base sequence is more complex, and we will discuss it in Section 7.7.

We can hydrolyze nucleic acids by means of hydroxides, acids, and enzymes. Alkaline hydrolysis of RNA and oligo- and polyribonucleotides proceeds through formation of nucleoside 2′,3′-cyclic monophosphate intermediates. In DNA and oligo- and polydeoxyribonucleotides, these cyclic intermediates cannot form because of the absence of a 2′-OH in deoxyribose. As a result, DNA and deoxynucleotide strands exhibit resistance to alkaline hydrolysis.

Acid hydrolysis, however, is effective with both RNA and DNA and corresponding oligo- and polynucleotides. When carried out under suitable conditions, hydrolysis produces free bases that can be characterized and quantified by chromatography. We generally express the data in terms of *mole percent* (moles of base per 100 moles of total bases). Such experiments yield the **base composition** of the nucleic acid, a term analogous to the amino acid composition of a protein.

Various enzymes, called **nucleases,** catalyze a hydrolytic cleavage of internucleotide bonds in nucleic acids. We refer to these enzymes as *ribonucleases* or *deoxyribonucleases,* depending on whether they catalyze RNA or DNA degradation. Nucleases, like proteases, can

Table 7.5. Characteristics of Some Nucleases

Enzyme	Substrate(s)	Cleavage[a]	Specificity[b]
Endonucleases			
Ribonuclease A (bovine pancreas)	ssRNA	b	$Py_p \downarrow B$
Ribonuclease T_1 (*Aspergillus oryzae*)	ssRNA	b	$G_p \downarrow B$
Deoxyribonuclease I (bovine pancreas)	ssDNA, dsDNA	a	$B \downarrow_p B$ (random)
Deoxyribonuclease II (bovine spleen)	ssDNA, dsDNA	b	$B_p \downarrow B$ (random)
Nuclease S_1 (*Aspergillus oryzae*)	ssRNA, ssDNA	a	$B \downarrow_p B$ (random)
Exonucleases			
Phosphodiesterase I (snake venom)	ssRNA, ssDNA	a	$B \downarrow_p B$ (from 3′-end; needs free 3′-OH)
Phosphodiesterase II (bovine spleen)	ssRNA, ssDNA	b	$B_p \downarrow B$ (from 5′-end; needs free 5′-OH)

[a]See Figure 7.12.
[b]B, Any base.

be of the *endo* type (attacking the strand internally) or the *exo* type (attacking the strand from one end). Nucleases catalyze hydrolysis of phosphodiester bonds in one of two ways, resulting in *type a* or *type b* cleavage (Figure 7.12). Some nucleases show specificity for bonds involving particular nucleotides; others are nonspecific (Table 7.5). The only nucleases known to be *sequence-specific* are the *restriction enzymes,* discussed in Section 7.6.

7.3. SECONDARY STRUCTURE

7.3.1. Watson and Crick Double Helix

Secondary structure refers to the *regular or periodic folding of the polynucleotide strand along one axis.* The **double helix** (*duplex*) of DNA (Figure 7.14), proposed by Watson and Crick, exemplifies the importance of the secondary structure.

Watson and Crick arrived at their proposal on the basis of two main lines of evidence. First, Rosalind Franklin had succeeded in obtaining an X-ray diffraction pattern using a DNA fiber. The X-ray diffraction data provided information about interatomic distances in DNA and indicated that the overall molecule was helical. Watson and Crick proceeded to build various molecular models that would agree with the X-ray diffraction data. Their breakthrough came when they realized that the correct forms for the bases were their *keto* rather than their *enol* forms.

The second line of evidence came from a series of experiments conducted by Erwin Chargaff in the late 1940s. When Chargaff isolated DNA from a large number of sources, hydrolyzed it, and determined its base composition, he found a striking regularity, described by two equations that we now term **Chargaff's rules:**

$$[A] = [T]$$
$$[G] = [C]$$

That is, the concentration (*in mole percent*) of adenine in a given DNA equals that of thymine, and the concentration of guanine equals that of cytosine. Chargaff's rules lead to three corollary relationships:

$$[A]/[T] = [G]/[C] = 1$$

$$\Sigma \text{ Pu} = \Sigma \text{ Py}$$
$$([A] + [G]) = ([T] + [C])$$

$$\Sigma \text{ Amino bases} = \Sigma \text{ keto bases}$$
$$([A] + [C]) = ([G] + [T])$$

The third corollary relates to one position, C(6) of purines and the corresponding C(4) of pyrimidines.

The DNA double helix has the following features:

1. It consists of two polynucleotide strands, each of which is a **right-handed helix** having dimensions different from those of the α-helix of proteins. The two helices are *intertwined,* resulting in **plectonemic coiling;** one strand cannot be unwound without unwinding the other at the same time.

2. The sugar–phosphate residues constitute the *backbone* of each strand and are located on the outer surface of the double helix. The negatively charged phosphate

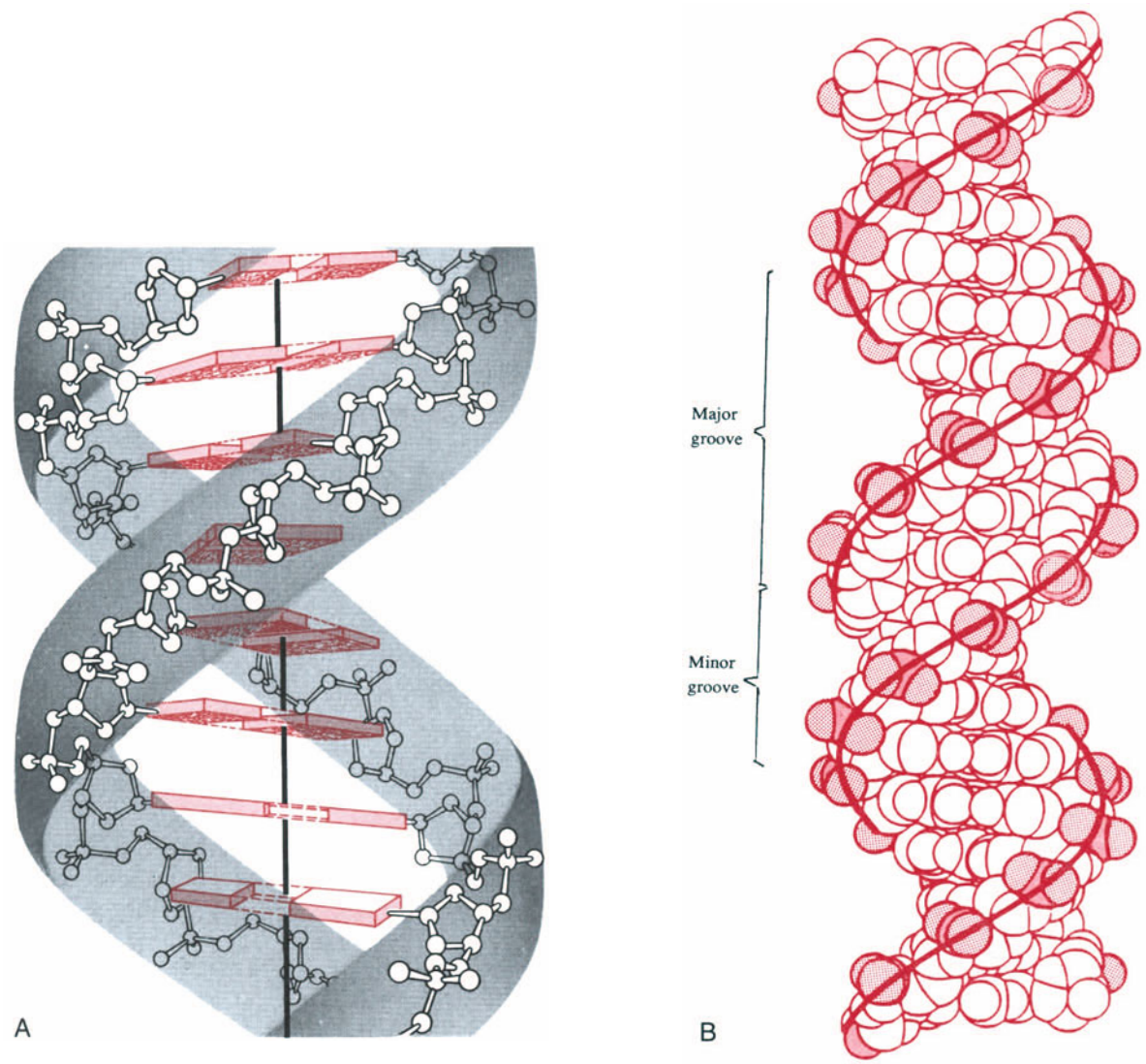


Figure 7.14. Models of the double helix of Watson–Crick type DNA. Exposed parts of the base pairs define a major and a minor groove. (A) Drawing that shows the stacked base pairs. (Figure copyrighted © by Irving Geis. Reproduced with permission.) (B) Space-filling model showing the helical path traced on the duplex surface by a line connecting the phosphate groups. [Reprinted, with permission, from A. Rich, A. Nordheim, and A. H. J. Wang, *Annu. Rev. Biochem.* 53:791–846 (1984). © 1984 by Annual Reviews, Inc.]

groups, and their positive counterions, make the exterior of the double helix polar.

3. The bases, attached to the sugars, are located *inside* the double helix. The planes of the bases are nearly perpendicular to the axis of the helix (see B-DNA in Table 7.6). Glycosidic bonds between bases and sugars have the *anti* conformation.

4. Juxtaposition of the bases leads to every A in one strand being opposite a T in the other strand, and vice versa; likewise for every G and C. We say that bases in one strand are *complementary* to those in the other strand. **Complementary base pairing** explains the data exemplified by Chargaff's rules. Although other types of base pairs are theoretically possible, only the AT and GC base pairs (*Watson–Crick base pairs*) have the appropriate dimensions to fit into the Watson–Crick double helix.

5. Every base in one strand is *hydrogen-bonded* to the complementary base in the other strand. The bonding consists of A and T being linked via *two hydrogen bonds* and G and C via *three hydrogen bonds* (Figure 7.15):

A ⋯ T

G ⋮⋮ C

6. The flat base pairs are *stacked* inside the double helix much like layers of wood in plywood. Stacking of

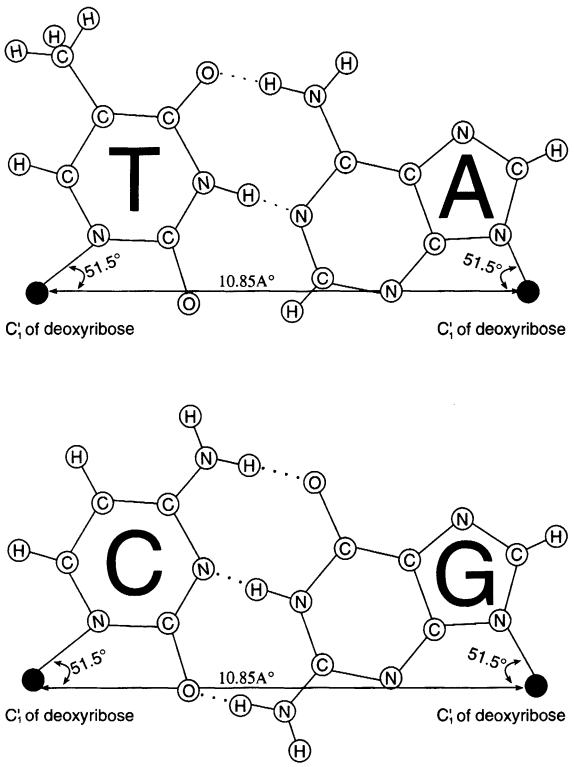


Figure 7.15. Complementary base pairing in DNA: Watson–Crick type base pairs in which adenine and thymine are linked via two H-bonds, and guanine and cytosine are linked via three.

the bases involves hydrophobic interactions that contribute to the overall stability of the double helix.

7. The double helix has two grooves along its outer surface, defined by exposed parts of the base pairs. A *major groove* is wide and deep, and a *minor groove* is narrow and deep. Both grooves are large enough to allow protein molecules to come into contact with the bases.

8. The two polynucleotide strands are **antiparallel.** Proceeding along the DNA in one direction, one strand runs from its 5′-end to its 3′-end while the other runs in the opposite direction, from its 3′-end to its 5′-end:

5′ A—C—G—T—T—A 3′

3′ T—G—C—A—A—T 5′

In this example, traversing the upper strand from left to right, you come to the 5′ position of a given deoxyribose first, and to the 3′ position subsequently. Traversing the lower strand *in the same direction,* you encounter the 3′ position of a deoxyribose first, followed by the 5′ position. Thus, the sugar residue in one strand appears to be rotated by 180° relative to that in the other strand (Figure 7.16).

It is interesting to note that both the α-helix for proteins and the double helix for DNA were proposed without a consideration of hydrophobic interactions. We now realize that hydrophobic interactions are involved in both protein and DNA structure and contribute to the overall stability of these macromolecules.

We refer to the Watson–Crick double helix of DNA as **B-DNA.** X-ray diffraction patterns of DNA fibers indicate the presence of this form when the counterion of the phosphate groups is an alkali metal ion such as sodium, and when the relative humidity is high (92%). We regard B-DNA as the native form of DNA, because its X-ray diffraction pattern is similar to that of DNA in intact sperm cells.

7.3.2. Other Nucleic Acid Duplexes

7.3.2A. A-DNA. At lower relative humidities (75%), B-DNA undergoes a reversible conformational change to **A-DNA** (Figure 7.17). In this double helix, both strands are right-handed, but the duplex is flatter and wider than that of B-DNA and has more base pairs per turn (Table 7.6). Like B-DNA, A-DNA has two grooves, but they differ in their characteristics from those in B-DNA: The major groove is narrow and deep, and the minor groove is wide and shallow. Additionally, A-DNA has an

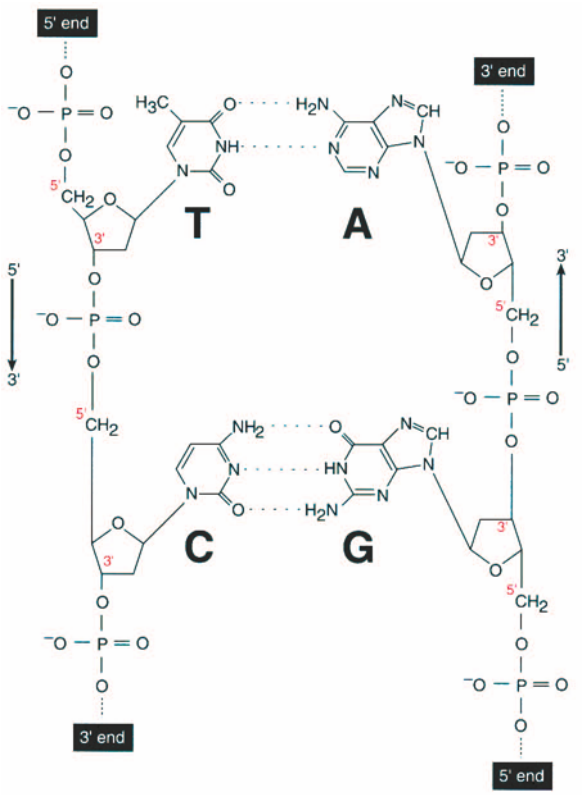


Figure 7.16. Antiparallel strands in DNA. The sugar moiety in one strand has undergone an apparent 180° rotation relative to that in the other strand.

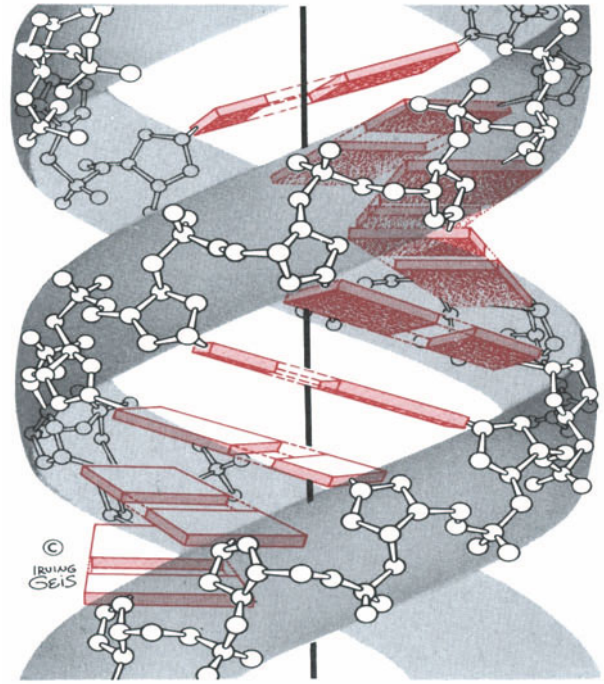


Figure 7.17. A-DNA. The bases are tilted from the perpendicular to the helix axis. (Figure copyrighted © by Irving Geis. Reproduced with permission.)

empty core around the helix axis, and the bases are inclined at an angle of 20° to the perpendicular to this axis. Glycosidic bonds between bases and sugars have the *anti* conformation.

7.3.2B. Z-DNA. Yet a different DNA double helix was discovered by Andrew Wang and Alexander Rich some 25 years after the introduction of the Watson–Crick model. This is a **left-handed helix** in which the base pairs have been flipped by 180° relative to those in B-DNA. As a result, the repeat unit in the DNA structure is a dinucleotide, rather than a mononucleotide as in B-DNA and A-DNA. The line joining successive phosphate groups follows a zigzag path on the outer surface of the helix, hence the name **Z-DNA** (Figure 7.18). Z-DNA

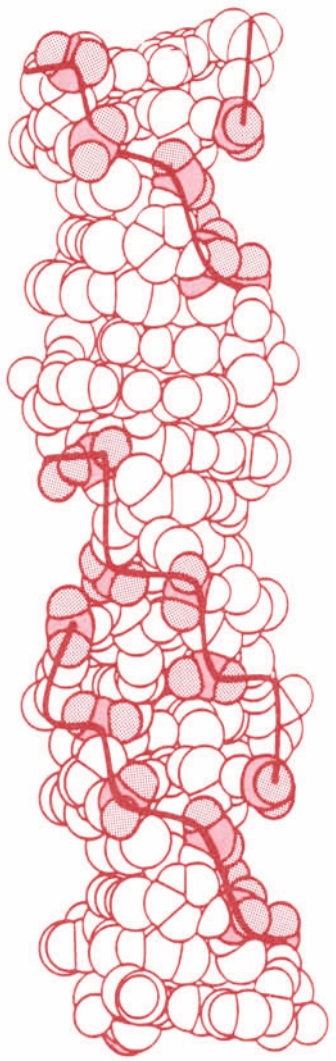


Figure 7.18. Z-DNA. The line connecting the phosphate groups traces a zigzag path on the surface of the duplex. [Reprinted, with permission, from A. Rich, A. Nordheim, and A. H. J. Wang, *Annu. Rev. Biochem.* 53:791–846 (1984). © 1984 by Annual Reviews, Inc.]

Table 7.6. Some Features of Double Helices

	A-DNA	B-DNA	Z-DNA	dsRNA
Handedness of strands	Right	Right	Left	Right
Helix diameter (nm)	~2.6	~2.0	~1.8	~2.6
Helix pitch (nm)[a]	2.8	3.4	4.5	3.0
Helix rise (nm)[b]	0.26	0.34	0.37	0.27
Base tilt[c]	20°	6°	7°	14°
Base pairs per helix turn	11	10	12	11

[a]Distance along the helix per one helix turn.
[b]Distance along the helix per one base pair.
[c]From perpendicular to the helix axis.

has been shown to occur *in vivo,* but its biological function remains unclear. Some have proposed that the reversible interconversion of B- and Z-DNA may function as a switch in regulating genetic expression. Z-DNA has a narrow and deep minor groove but no discernible major groove. Glycosidic bonds between bases and sugars are unusual. The bond is *anti* for pyrimidines but *syn* for purines. Formation of Z-DNA appears to be facilitated by alternating sequences of purines and pyrimidines.

7.3.2C. Double-Stranded RNA. Most RNAs occur as single-stranded molecules. Some viral RNAs, however, are double-stranded. Additionally, rRNA and tRNA contain some double-stranded segments. We will discuss the different types of RNA in Section 16.2. At this point we only want to deal with RNA's double-stranded aspect.

A double-stranded RNA segment forms by having a single ribonucleotide strand *fold back on itself* (Figure 7.19). This conformation provides juxtaposition of *two antiparallel strands* linked, as in DNA, by means of complementary base pairing. Guanine–cytosine base pairs (three H-bonds) occur in both DNA and RNA. However, since RNA does not contain thymine, the adenine–thymine base pair of DNA is replaced by an adenine–uracil base pair (two H-bonds) in RNA. In double-stranded RNA, the bulky OH group at the C(2′) position of ribose causes steric hindrance that prevents formation of a Watson–Crick-type double helix. Instead, double-stranded RNA has a structure that resembles that of *A-DNA* (Table 7.6).

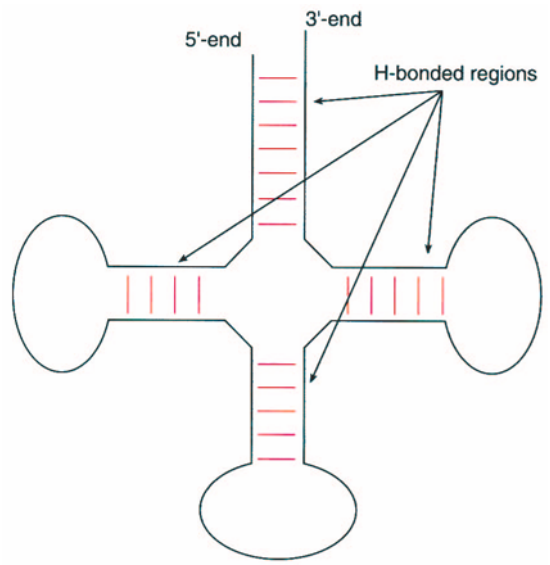


Figure 7.19. Double-stranded, H-bonded segments in RNA. H-bonds form between adenine and uracil (two bonds) and between guanine and cytosine (three bonds). The structure represents the cloverleaf model of transfer RNA.

7.4. TERTIARY STRUCTURE OF DNA—THE SUPERHELIX

Some double-stranded circular DNAs have a twisted appearance when examined with the electron microscope. Such structures form by the folding of an already folded structure (the double helix) and represent DNA's tertiary structure. We call the twisted double helix a **superhelix** or **supercoil** (Figure 7.20).

You can easily demonstrate formation of a superhelix with a piece of rubber tubing or a thick rope. If you hold the tubing with both hands and bring the two ends together, you will form a closed and *relaxed* circle. However, if you first rotate one end of the tubing by 360°, while holding on to the other end, and then bring the two ends together, a figure eight will result as the tubing compensates for the strain imposed on it. If, before bringing the two ends together, you rotate one end a number of times by 360°, the tubing will take on the shape of multiple figure eights; the more you rotate the tubing, the larger the number of loops formed. The superhelix of DNA forms likewise; a relaxed closed circle changes to a linked string of figure eights.

When DNA is twisted in an opposite sense to that of the double helix, we term the superhelix *negative,* or *underwound.* If the superhelix forms by twisting the DNA in the same sense as that of the double helix, we term the superhelix *positive,* or *overwound.* Naturally occurring DNA is a negative superhelix.

Supercoiling of DNA has two important functions. First, it decreases the occupied space and thus allows for more efficient *packaging* of DNA inside the cell. Second, supercoiling may allow for specific enzyme reactions and protein binding. Interaction with enzymes and other proteins may depend on the precise topological state of the DNA and proceed only with either relaxed, negatively supercoiled, or positively supercoiled DNA.

A special class of enzymes, *topoisomerases,* catalyzes the incremental interconversion of topological isomers of DNA. Topoisomerases fall into two classes: *type I enzymes* catalyze the relaxation of negatively supercoiled DNA, whereas *type II enzymes* catalyze the induction of negative supercoiling in relaxed DNA.

A type I enzyme, or *nicking–closing enzyme,* catalyzes the hydrolysis of a phosphodiester bond and produces a transient *break in one strand (a nick)* of supercoiled DNA. Rotation of one strand around the other releases the strain, and the enzyme then catalyzes a re-

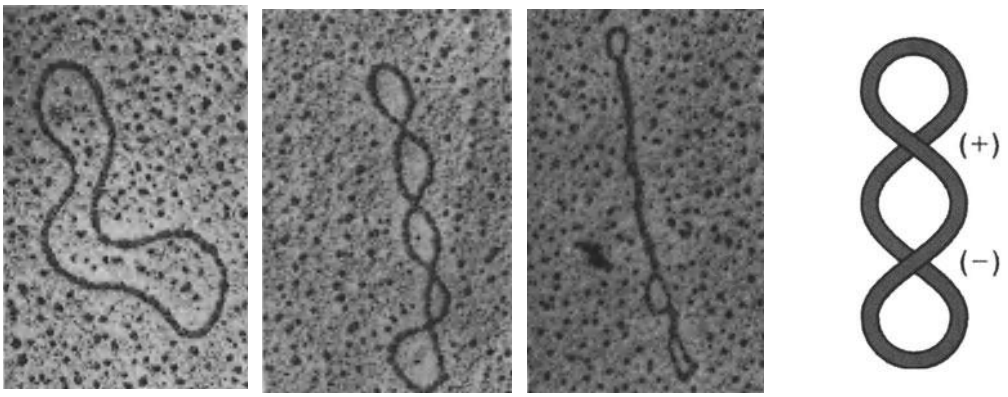


Figure 7.20. (A) Electron micrographs of relaxed, moderately supercoiled, and highly supercoiled DNA. (Courtesy of Dr. L. Polder.) (B) Negative and positive supercoils.

sealing of the break to produce a relaxed DNA segment. A type II enzyme, or *DNA gyrase,* catalyzes the hydrolysis of two phosphodiester bonds and produces a transient *break in both strands (a cut),* followed by resealing in such a fashion that a relaxed DNA segment is converted to a negatively supercoiled one (see Figure 17.13).

Note that tertiary structure is not limited to DNA but occurs in RNA as well. The L-shaped form of transfer RNA (see Section 16.2) provides an example of a specific tertiary RNA structure.

7.5. DNA—THE GENETIC MATERIAL

The identification of DNA as the genetic material of the cell evolved from an accumulation of both indirect and direct evidence. As studies of DNA spread, researchers discovered that DNA possesses the kinds of properties one might reasonably associate with a substance serving as genetic material. Such indirect evidence was augmented by direct evidence, derived from several ingenious experiments involving DNA. These milestones in the development of molecular biology showed conclusively that DNA serves as the carrier of genetic information.

7.5.1. Indirect Evidence

Indirect evidence for the role of DNA consisted of a number of findings. Investigators noted, for example, that the amount of DNA was constant for cells of a given organism and that it generally increased with the complexity of the organism (Table 7.7); a human cell contained considerably more DNA than a bacterial cell. Researchers also found that the size of DNA molecules generally increased with the complexity of the organism (Table 7.3). We now know that when an exception to these generalizations occurs, it does not mean that the less complex organism contains more genetic information. Rather, the increased amount of DNA represents sections of *nontranscribed* or *nontranslated DNA* (DNA without genetic information) or *repetitive DNA* (multiple copies of the same genes).

Another property that we expect a substance controlling heredity to exhibit is invariance. That, too, was shown to apply to DNA. Changes in the environment, metabolism, or nutrition did not affect either the amount of DNA per cell or its composition. DNA base composition was shown to vary from species to species but to be constant for any given species. In the case of bacteria, base composition [%(G + C)] varies to such an extent that it can be used as a criterion for species classification.

7.5.2. Direct Evidence

7.5.2A. Transforming Principle. The first direct evidence for DNA's function in heredity was provided by an experiment conducted by O. T. Avery, C. M. MacLeod, and M. McCarty in 1944 (Figure 7.21) and based on prior work by R. Griffith. These workers used

Table 7.7. Amounts of DNA in Different Organisms

Type of cell	Organism	Mass of DNA per cell (pg)[a]
Bacteriophage	T4	2.4×10^{-4}
Bacterium	*Escherichia coli*	4.4×10^{-3}
Fungus	*Neurospora crassa*	1.7×10^{-2}
Erythrocyte	Chicken	2.5
Plant	Tobacco	2.5
Protozoan	Euglena	3.3
Leukocyte	Human	3.4

[a] 1 pg (picogram) = 10^{-12} g.

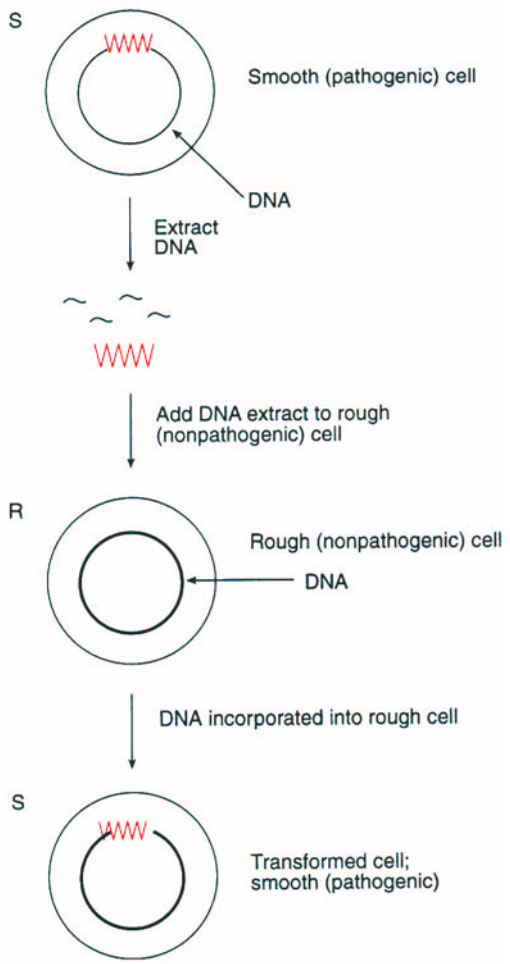


Figure 7.21. Avery, MacLeod, and McCarty's experiment on the transformation of rough, nonpathogenic *Pneumococcus* cells to smooth and virulent cells.

two species of the bacterium *Pneumococcus.* One, designated *S,* grows as a *smooth*-looking colony, has a carbohydrate capsule, and is pathogenic (virulent). The other, designated *R,* grows as a *rough*-looking colony, does not have a capsule, and is nonpathogenic.

Avery and colleagues grew S-type cells, extracted their DNA, and added it to cultures of R-type cells. In the process of incorporating the added DNA, R-type cells were *transformed* into S-type cells; their colony had a smooth-looking appearance, and the organisms were now pathogenic. Evidently, the genetic information required to produce the carbohydrate capsule and to endow the organism with pathogenicity had been transferred from one organism (S) to another (R) by means of the added DNA. A debate ensued as workers tried to establish the nature of the **transforming principle** that brought about this change. Was it the DNA itself or was it a contaminating protein? When they finally showed that the transforming principle was pure DNA, clear proof had been advanced for DNA's role as genetic material.

For some reason, this important work was largely ignored for eight years until the Hershey–Chase experiment (see below) led to its reevaluation. Possibly, the fact that the research was published during World War II, and in a medical rather than a biochemical journal, was partly responsible for the delayed recognition.

7.5.2B. Bacteriophage Infection. A second key experiment was carried out by A. D. Hershey and M. Chase in 1952. These workers used *bacteriophage T2,* a phage that infects *Escherichia coli.* T2 is a large, DNA-containing phage with an icosahedral capsid (head) that encloses a linear, double-stranded DNA molecule (Figure 7.22). The phage adsorbs to the bacterial cell wall by means of *tail fibers* and injects its DNA into the host cell. Once inside, the viral DNA replicates and leads to synthesis of viral protein coats. Assembly of many intact viral particles inside the bacterium ultimately results in cell *lysis* (rupture) and cell death. Hershey and Chase prepared a phage whose DNA was labeled with ^{32}P and whose protein coat was labeled with ^{35}S. They showed that, during infection, only the labeled DNA entered the cell while the protein coat remained outside. They concluded that the genetic information for making intact viral particles must all be contained within the DNA.

7.5.2C. *In Vitro* Synthesis. In a third important experiment, conducted in 1967, M. Goulian, A. Kornberg, and R. L. Sinsheimer, using a number of purified enzymes, were able to carry out the complete *in vitro* synthesis of an intact phage DNA molecule. This phage, designated *ϕX174,* consists of a circular, single-stranded DNA molecule (5386 nucleotides) enclosed in a protein coat. The synthetic preparation of the phage DNA was shown to have the same properties as the native viral DNA, including identical base sequence and full infectivity. Because only nucleic acid synthesis took place during the experiment, it was impossible that contaminating proteins could account for the results. The scientists concluded that the synthetic DNA contained all of the genetic information of the native phage.

These three critical experiments were followed by a burst of research activity that has resulted in an accumulation of molecular-biological data and the development of many new techniques. Particular achievements include synthesis of specific genes, insertion of genes into cellular DNAs, and other aspects of recombinant DNA technology. The role of DNA as genetic material is now firmly established.

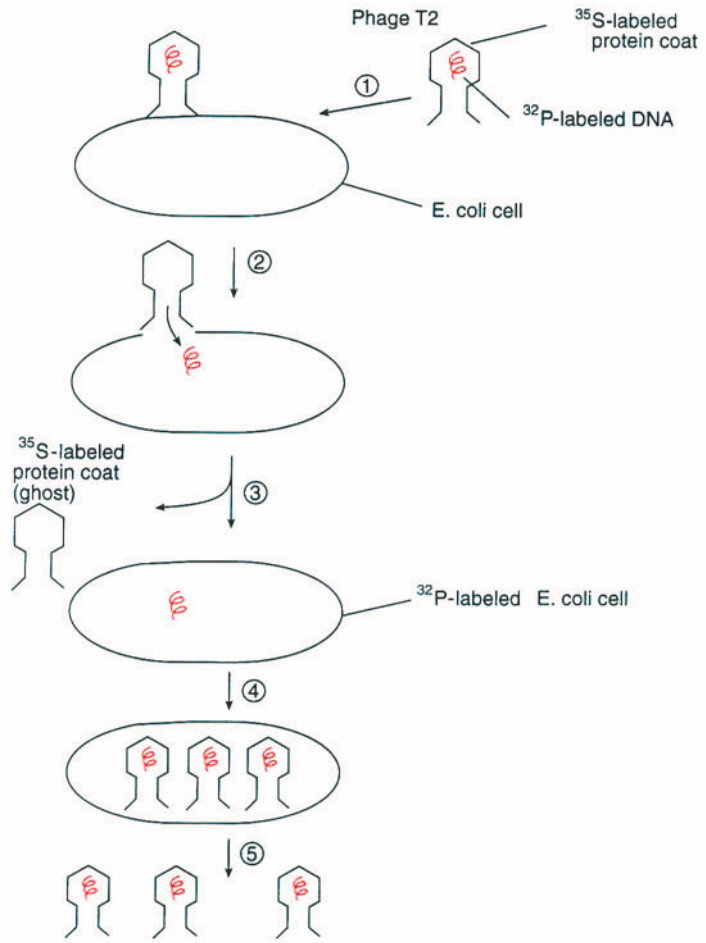


Figure 7.22. Hershey–Chase experiment on the infection of *E. coli* cells with bacteriophage T2. (1) Phage attaches to cell wall and (2) injects its DNA into the cell while its protein coat (capsid) stays outside; (3) agitation with a blender removes capsids from the cells; (4) phage DNA gives rise to new intracellular complete viral particles; (5) cell lyses and releases phage particles that can infect other cells.

7.5.3. Prokaryotic and Eukaryotic DNA

Although prokaryotic and eukaryotic DNAs have the same basic structure, their size, organization, and other properties differ significantly (Table 7.8). In both prokaryotes and eukaryotes, most of the DNA is located in the chromosomes. The term **chromosome** describes any large DNA molecule that constitutes a discrete unit of the genome and carries many genes. Chromosomes function in the storage and transmission of genetic information.

7.5.3A. Prokaryotes. The bacterial chromosome consists of one very large, double-stranded, circular DNA molecule; the molecular weight of *E. coli* DNA, for example, is 2.6×10^9. The DNA is located in a nuclear region (*nucleoid*) but is not surrounded by a nuclear membrane. Because there is just one molecule of DNA per cell, preparations of bacterial DNA exhibit great homogeneity as compared to those from humans and other eukaryotes.

Storing the DNA in the nucleoid requires *packaging*. A simple calculation can explain this concept. An *E. coli* cell has a length of about 2 μm (1 μm = 10^{-6} m), but the DNA has a length of 1400 μm. Hence, even if the DNA were deposited throughout the entire *E. coli* cell, its storage would require a packaging ratio of about 700/1. In other words, 700 μm of DNA must be folded and contained in 1 μm of cell space. The packaging ratio is actually much larger because the nucleoid occupies only a fraction of the total cell volume. Efficient DNA storage is made possible by the superhelical configuration of DNA.

Table 7.8. Comparison of Prokaryotic and Eukaryotic DNA

Property	Prokaryotes	Eukaryotes
Molecular weight	*E. coli*: 2.6×10^9 (4×10^6 bp; 1.4 mm)	Human: 1.8×10^{12} (2.9×10^9 bp; 0.99 m)
Number of molecules per chromosome	1	1
Number of molecules per cell	1	Many
Configuration	ds circular[a]	ds linear
Supercoiling	Yes	Yes
Associated with basic proteins	Some; no specific structure	Histones; nucleosome, nucleosome packing
Repetitive DNA	No	Yes
Number of times a gene occurs	Once	Once or more
Genes continuous	Yes	No (introns, exons)
Nontranslated DNA	Few, small regulatory sequences	Many, often long sequences
Palindromes	Few, small	Many, large
Restriction/modification system	Yes	No
Plasmid DNA (ds circular)	Yes	No
Organelle DNA (ds circular)	No	Yes

[a]ds, Double-stranded.

In addition to chromosomal DNA, bacteria contain *extrachromosomal* genetic elements called **plasmids.** A plasmid is a circular, double-stranded DNA molecule that usually provides the organism with some evolutionary advantage such as resistance to antibiotics or production of bactericidal substances. Plasmids replicate independently of the bacterial chromosome and constitute a useful tool in recombinant DNA technology (see Appendix C). Some plasmids, called *episomes,* can become integrated reversibly into the host bacterial chromosome.

7.5.3B. Eukaryotes. Eukaryotic DNA is organized into morphologically distinct structures, the chromosomes, that you can see with a light microscope during a certain part of the cell cycle. Eukaryotic DNA is much larger than that of prokaryotes. The DNA of each chromosome constitutes an enormous, double-stranded, linear molecule. Because a eukaryotic cell contains different chromosomes, preparations of eukaryotic DNA generally exhibit heterogeneity.

The occurrence of giant DNA molecules in microscopic chromosomes represents a formidable packaging problem. The DNA contained in the 23 chromosomes of the human (haploid) cell has a total length of about one meter. The 23 chromosomes themselves have a total length of about 100 μm. Therefore, the packaging ratio must be

$$1/(100 \times 10^{-6}) = 10{,}000/1$$

This means that 10,000 μm of DNA must be contained in 1 μm of chromosome. Moreover, the packaging must be such that all functional elements of the DNA (all the genes) are accessible for transcription into RNA and for other regulatory functions. The great efficiency of DNA storage is achieved in three ways. First, the DNA exists in supercoiled configuration. Second, the supercoiled DNA is wound around globular cores of protein, resulting in beadlike structures named *nucleosomes.* Lastly, the nucleosomes are closely packed within the chromosome.

A **nucleosome** consists of DNA and basic proteins, **histones,** held together by electrostatic attraction. Histones are rich in the basic amino acids arginine and lysine. We classify histones into five categories based on their content of these amino acids (Table 7.9).

The nucleosome contains eight histones (two each of H2A, H2B, H3, and H4) around which are wound approximately 1.75 turns of superhelical DNA (Figure

Table 7.9. Classification of Histones

Type	Arginine content (%)	Lysine content (%)
H1	1	29
H2A	9	11
H2B	6	16
H3	13	10
H4	14	11

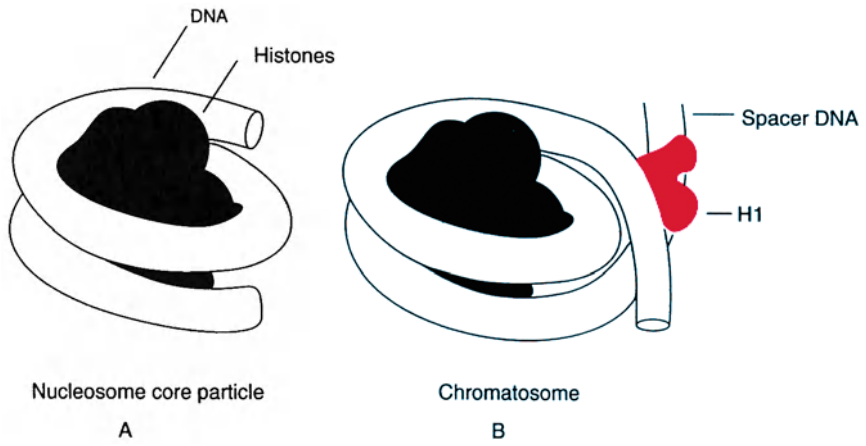


Figure 7.23. Proposed structure of nucleosomes and chromatosomes. (A) A nucleosome consists of a core of eight histones (two each of H2A, H2B, H3, and H4), surrounded by 1.75 turns of superhelical DNA. (B) A chromatosome consists of a nucleosome plus spacer DNA with its attached histone H1.

7.23). The entire complex has a diameter of about 10 nm and contains 146 base pairs (bp) of DNA. Nucleosomes are separated by sections of "naked" DNA, amounting to some 10–100 bp and called **spacer DNA** or **linker DNA.** A molecule of histone H1 attaches to the spacer DNA. The nucleosome plus the section of spacer DNA (with histone H1 attached) forms a **chromatosome.** Chains of chromatosomes form **chromatin,** the substance from which chromosomes are constructed. Because it contains so many chromatosomes, chromatin has an appearance of "beads on a string." In the chromosome, these beads are closely packed. According to one proposal, chromatin forms a compact structure in which the zigzag pattern of nucleosomes closes up to form a helical solenoid with about six nucleosomes per turn.

Portions of eukaryotic DNA may occur a large number of times; some genes occur in the form of multiple copies or **repetitive DNA.** Eukaryotic DNA also contains sections that do not carry genetic information (noncoding) and do not appear in the final RNA transcript. We call these sections **introns** (*in*tervening sequences). Introns flank sections that carry genetic information (coding) and form part of the final RNA transcript. These sections we call **exons** (*ex*pressed sequences). In prokaryotes, a gene occurs as one continuous coding base sequence, but in eukaryotes many genes occur as exons separated by introns. Before such genes can be expressed, the original RNA transcript of both introns and exons must be processed. Introns must be removed, and exons must be spliced together (more in Section 18.3). Note that the terms *intron* and *exon* refer to both DNA and the corresponding sections of the RNA transcript.

Eukaryotes, like prokaryotes, also contain **extrachromosomal DNA.** Such DNA occurs in both *mitochondria* and *plastids.* Mitochondrial DNA occurs as a double-stranded, circular, histone-free molecule that replicates independently of the nuclear DNA. It also codes for components of mitochondrial protein synthesis and for proteins that function in the electron transport system and oxidative phosphorylation. The genetic code of mitochondrial DNA differs slightly from the "universal" code (see Section 16.3). These findings support the theory, discussed in Section 1.1, that mitochondria represent an early bacterial form that began a symbiotic relationship with eukaryotic cells.

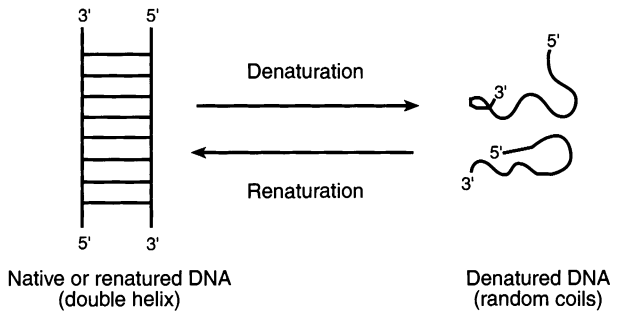
In many plants, some DNA occurs in plastids, which are self-replicating subcellular organelles. Plastids function in photosynthesis (*chloroplasts*), starch storage, and flower pigmentation (*chromoplasts*).

7.6. PROPERTIES OF DNA

7.6.1. Degradation and Denaturation

Nucleic acids are degraded when hydrolysis of 3′,5′-phosphodiester bonds breaks the polynucleotide strand into smaller fragments. **Degradation** can be produced by nuclease action or by physical means, especially in the case of DNA. Molecules of DNA are both very long and very thin; *E. coli* DNA, for example, has a length-to-diameter ratio of 730,000. Consequently, DNA molecules are fragile and easily broken. The shear produced by common laboratory practices such as shaking, stirring, and pipetting may be sufficient to degrade DNA, so special care must be exercised in handling DNA solutions.

By contrast, **denaturation** of nucleic acids, much like that of proteins, involves breaking noncovalent bonds and loss of secondary structure. For nucleic acids, denaturation describes a transition from a double-stranded helical configuration to one of single-stranded random coils:



The transition produces a decrease in optical rotation because the double helix of nucleic acids, like the α-helix of proteins, contributes to the overall optical rotation of the molecule. Additionally, and especially in the case of DNA, solution viscosity decreases upon denaturation, as long helical DNA molecules are converted to globular random coils. Lastly, denaturation results in pronounced changes in absorbance (see below).

Nucleic acid denaturation involves breaking both hydrogen and hydrophobic bonds. Agents that produce denaturation are, for the most part, the same as those that denature proteins. These agents include H-bond-forming compounds such as urea and guanidine hydrochloride and hydrophobic-bond-forming compounds such as sodium dodecyl sulfate and other detergents. Nucleic acids (especially DNA) can also be denatured by extremes of pH, low ionic strength, and high temperature.

7.6.1A. Extremes of pH. Conditions of *low or high pH* tend to destabilize DNA and may lead to denaturation. At low pH, protonation of adenine's N(1) and cytosine's N(3) eliminates these as proton acceptor sites for H-bonding. Additionally, protonation of all bases leads to their acquiring a relative positive charge so that their mutual repulsion increases.

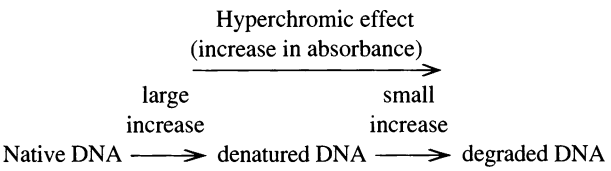
At high pH, deprotonation of guanine's N(1) and thymine's N(3) eliminates these as proton donor sites for H-bonding. Additionally, deprotonation of all bases also leads to their acquiring a relative negative charge so that, once again, their mutual repulsion increases.

7.6.1B. Low Ionic Strength. Mere *suspension in distilled water* may lead to denaturation if the concentration of dsDNA is low. Under those conditions, a given volume of solution contains only a few DNA molecules together with their positive counterions. The latter, while somewhat clustered around the DNA, also diffuse away and distribute themselves throughout the solution. This leaves DNA's negatively charged phosphate groups partially unshielded. Repulsive forces between adjacent phosphate groups on a DNA strand become pronounced and may be sufficient to cause the two strands to separate.

We can minimize the chance of such denaturation by using a higher concentration of DNA, thereby increasing the total counterion concentration to provide sufficient electrostatic shielding. Alternatively, we can increase the ionic strength of the solution by adding salt, thereby dampening the repulsive forces between adjacent phosphate groups.

7.6.1C. High Temperature. Nucleic acids, like proteins, undergo denaturation when heated. We commonly follow the **thermal denaturation** of nucleic acids by measuring the changes that occur in the ultraviolet absorbance of their solutions. Changes in absorbance are especially large for solutions of dsDNA. By measuring absorbance as a function of temperature, we can follow the transition from native to denatured DNA.

To do this, we place a DNA solution in a spectrophotometer and measure the absorbance while changing the solution temperature. We find that the absorbance *increases* by about 40% as the temperature rises from room temperature to around 100°C. We refer to the increase in absorbance as a *hyperchromic effect.* If, following denaturation, we *degrade* the denatured DNA to mononucleotides, there occurs an additional, though small, increase in absorbance:



Proceeding in the reverse sense, from degraded to denatured DNA and then to the native state, leads to a *decrease* in absorbance or a *hypochromic effect.* The absorbance changes occur at all wavelengths at which DNA absorbs (Figure 7.24), but we generally determine them at 260 nm (Figure 7.25).

The absorbance changes result from alterations in base–base interactions. In native DNA, each base participates in a large number of electronic interactions with other bases: H-bonding to the complementary base and hydrophobic interactions with the two adjacent base pairs. Because of these multiple interactions, a given base can-

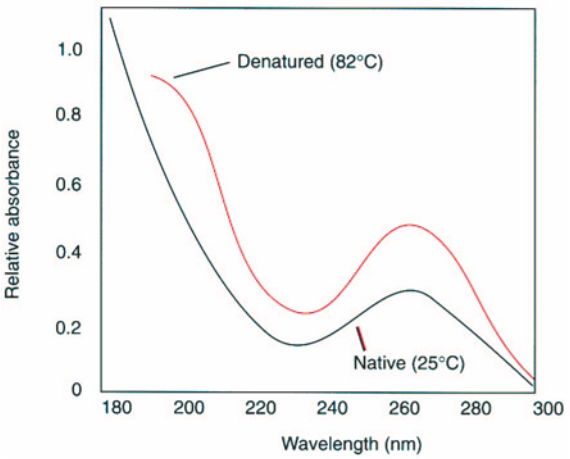


Figure 7.24. Hyperchromic effect of DNA.

not assume many resonance structures by shifting electrons within the molecule. However, when electronic interactions between bases are disrupted by denaturation (and, to a lesser extent, by degradation), the base can assume a variety of resonance structures. Since an increase in the number of resonance structures leads to an increase in ultraviolet absorbance, denatured and degraded DNA absorb more strongly than native DNA.

We call a plot of absorbance as a function of temperature a *thermal denaturation profile.* Because the transition from double helix to random coil usually occurs over a relatively narrow range, we refer to the process as *melting* by analogy with the melting of a crystal in organic chemistry. The midpoint of the transition, T_m, constitutes the **melting temperature** of the DNA. It is *the temperature at which one-half of the maximum change in absorbance is observed.*

Melting temperatures are related to DNA base composition. Biochemists have isolated a large number of DNAs, denatured them thermally, and determined their T_m values. A plot of T_m as a function of base composition (in terms of mole percent of guanine plus cytosine) yields a straight line such that T_m increases with increasing mole percent of (G + C) (Figure 7.26). Since guanine and cytosine are linked via three H-bonds while adenine and thymine are linked by only two, an increase in the content of guanine and cytosine allows for more extensive H-bonding, resulting in a DNA that has greater thermal stability.

The straight-line equation derived from the data of Figure 7.26, and applicable to DNA solutions containing 0.15*M* NaCl and 0.015*M* sodium citrate (pH 7), takes the form

$$T_m = 69.3 + 0.41(G + C)$$

where T_m is in degrees Celsius and (G + C) is in mole percent. The fit of the points to the line in Figure 7.26 is so good that we can use the equation to determine the base composition of an unknown DNA from the spectrophotometric measurement of its T_m.

The sigmoidal thermal denaturation profile indicates that denaturation consists of a *cooperative* process in which H-bonds constitute the "binding sites." Breaking

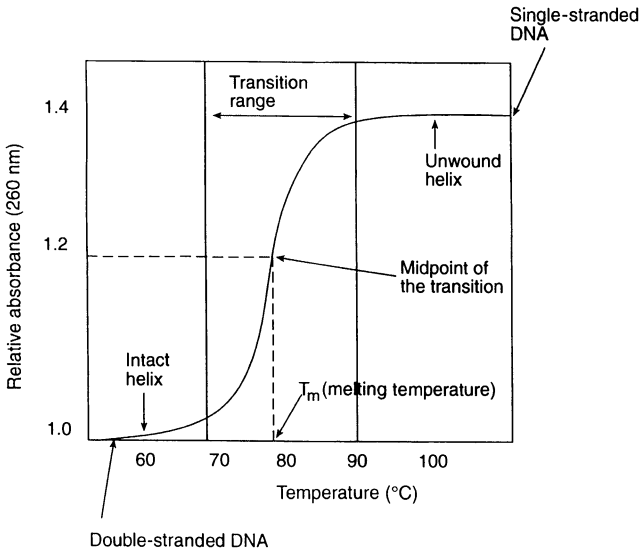


Figure 7.25. Thermal denaturation profile (melting curve) of DNA.

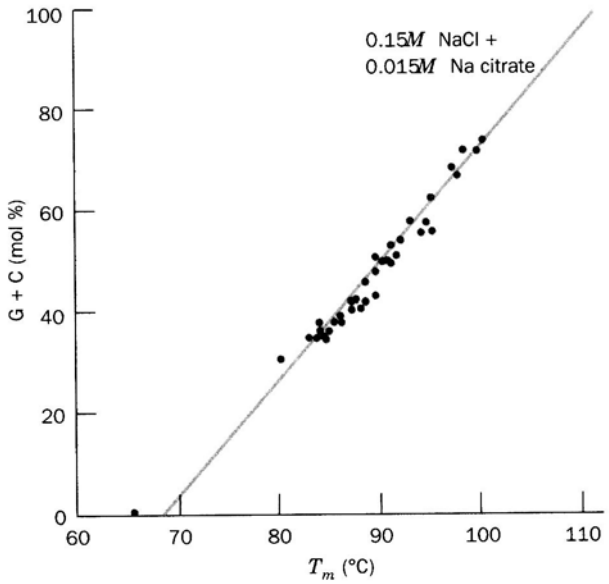


Figure 7.26. Variation of melting temperature (T_m) with DNA base composition [mole percent of (G + C)]. [Reprinted, with permission, from J. Marmur and P. Doty, *J. Mol. Biol.* 5:109–118 (1962).]

the first H-bond by increasing the temperature proceeds with relative difficulty because extensive H-bonding confers stability on the DNA. However, once the first H-bond has been disrupted, adjacent H-bonds become destabilized so that the second H-bond breaks more easily, and so on. The process resembles that of opening a zipper. Once you overcome the initial resistance, you can move a zipper easily. Because of this analogy, we frequently refer to unwinding of DNA as *unzippering*.

RNAs that contain double-stranded segments also show a hyperchromic effect. However, absorbance changes for RNA are usually much smaller than those for DNA. Whereas dsDNA has essentially 100% of its structure in double-helical form, most RNAs possess only a few, short double-stranded segments.

7.6.2. Renaturation and Hybridization

If we immerse a solution of thermally denatured DNA in an ice-water bath to *cool it rapidly,* the strands remain separated, except for a small amount of nonspecific pairing; we consider the DNA to be in its denatured state. If, however, we allow the solution to *cool slowly* (called "annealing"), the strands slide past each other, forming and breaking H-bonds repeatedly. Ultimately, they line up properly, re-form the original H-bonds, and restore the DNA to its native double-stranded state. The DNA has undergone **renaturation.**

Other conditions that favor reassociation of single strands are high concentrations of DNA (to increase the chance of strand–strand interactions), high ionic strength (to allow close approach of strands by minimizing repulsion between charged phosphate groups), and a pH around neutrality (to minimize electrical repulsion among bases). Lastly, reassociation proceeds more rapidly, the less complex the DNA; we can renature bacterial or viral DNA more readily than eukaryotic DNA. Renaturation, like denaturation, is a cooperative process. Formation of the first H-bond proceeds with relative difficulty, but once that has been achieved, and the strands have begun aligning, formation of subsequent H-bonds proceeds more easily.

Renaturation provides the basis for a useful analytical technique called *hybridization.* Assume, for example, that you prepare a solution containing a mixture of two DNAs. After heating to denature the DNAs, you allow the solution to cool slowly. Renaturation will lead to reassociation of the single strands, forming the

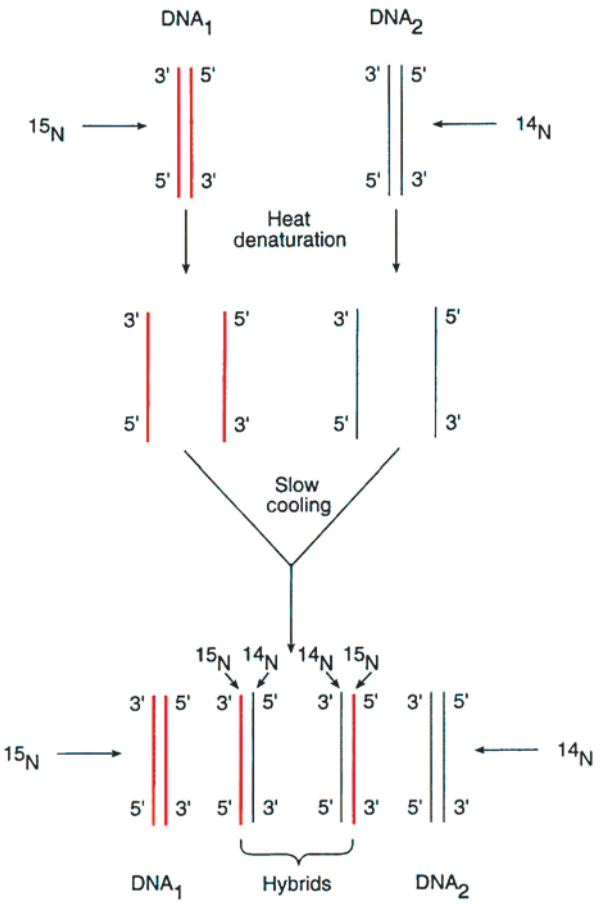


Figure 7.27. Principle of hybridization. Strands H-bond to varying degrees, depending on the extent of complementary base sequences present.

two original dsDNAs. In addition, varying amounts of two *hybrid* double-stranded structures will form, depending on the degree of similarity between the base sequences of the two DNAs (Figure 7.27): the greater the similarity between the two DNAs, the greater the amount of hybrid formation. Typically, we measure the extent of hybrid formation by using one DNA in fragmented and radioactively labeled form. Hybridization yields information about the similarity of base sequences of different DNAs but does not yield the actual base sequence.

We can also use hybridization to study the binding between two ssDNAs, two ssRNAs, or ssDNA and ssRNA. Double-stranded RNA–DNA hybrids have dimensions close to those of A-DNA, and their H-bonding is a combination of that in dsDNA and that in dsRNA:

DNA	G	C	A	T
	⋮	⋮	⋮	⋮
RNA	C	G	U	A

7.6.3. Restriction and Modification

Prokaryotic organisms possess sets of enzymes that produce specific chemical changes in dsDNA. Each set of enzymes consist of a **restriction endonuclease** (**restriction enzyme**) and a corresponding **modification methylase.** A restriction endonuclease catalyzes the hydrolysis of both strands of dsDNA at specific sites that usually consist of four to six base pairs, typically in the form of a **palindrome** or *inverted repeat.* In general terms, a palindrome is a number, word, or sentence that reads the same forward and backward. Some examples are "18181," "noon," and "Madam, in Eden I'm Adam." An example of a palindromic sequence in DNA is

5′ A A G C T T 3′

3′ T T C G A A 5′

A given restriction endonuclease "recognizes" specific palindromic sites on DNA. We use the term *recognize* to indicate a unique binding interaction in a biochemical system. In the present case, the term refers to the binding of a restriction endonuclease to DNA. Cleavage of the strands as a result of endonuclease action can occur in two ways. Some enzymes catalyze hydrolysis of both strands at the same point, resulting in a "flush" cut, whereas others create "staggered" cuts (Figure 7.28). Staggered cuts result in formation of **cohesive ends.** These single-stranded segments have a great tendency to form a double strand because they possess complementary base sequences. Depending on the specificity of the enzyme, the actual cuts in the DNA occur either in the region of the palindrome or at some distance from it.

Biochemists have identified three types of restriction endonucleases (Table 7.10). Type II endonucleases have been very useful in molecular biology because they produce the cut *within* the recognition site, so that DNA cleavage becomes sequence-specific. Treating a DNA molecule with a type II endonuclease produces a small number of fragments. If we know the base sequence of the DNA, we can predict the number and structure of the fragments.

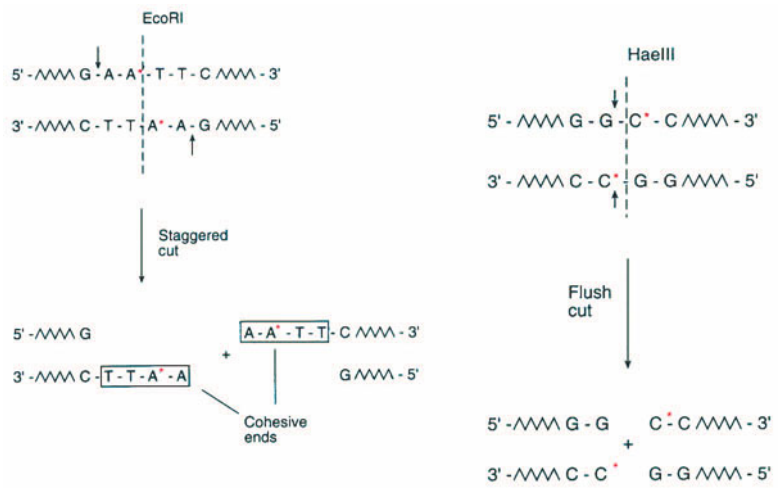


Figure 7.28. Mode of action of restriction endonucleases. Arrows indicate points of hydrolysis on either side of the palindrome axis (dashed line). Asterisks designate bases altered by modification methylases.

Table 7.10. Types of Restriction Endonucleases

	Type I	Type II	Type III
Cleavage site	1–10 kbp from recognition site	Within recognition site	24–26 bp from recognition site
Nuclease and methylase in one enzyme	Yes	No	Yes
ATP required for cleavage	Yes	No	No

Restriction endonucleases constitute powerful enzymes for the specific degradation of DNA. In the absence of safety precautions, these enzymes would degrade the cell's own DNA. Protection against such self-destruction is provided by counterpart enzymes—*modification methylases*. A modification methylase catalyzes the methylation of DNA at precisely the same palindromic sequences recognized by the corresponding restriction endonuclease. Once the palindromic sites have been methylated, they can no longer be "recognized" by the endonuclease and are not cleaved. Methylation thus protects the cell's DNA against digestion by the cell's endonucleases. At times, both enzymatic activities reside on the same protein molecule; at other times, they occur in the form of two separate proteins.

The restriction–modification system protects cells against foreign DNA. When DNA from a different bacterium or a phage enters a cell, it is rarely modified by methylation because modification methylases are generally species-specific. Hence, the foreign DNA is unprotected and serves as a substrate for restriction endonucleases that catalyze its degradation to a number of smaller fragments. Other intracellular nucleases subsequently degrade these fragments further.

7.7. SEQUENCING NUCLEIC ACIDS

We have seen that hybridization provides some clues as to the similarity or dissimilarity of two nucleic acids. Comparable information can also be obtained by constructing **nucleotide maps** analogous to the peptide maps described for proteins (Section 3.2). To prepare nucleotide maps, we subject nucleic acids to partial hydrolysis, followed by two-dimensional separation of the fragments by chromatography and/or electrophoresis. Although these methods are useful, they do not yield the actual base sequence of a nucleic acid.

Determining base sequences (**sequencing**) was for many years a much more difficult undertaking than determining amino acid sequences. Two reasons account for this. First, nucleic acids frequently consist of much larger molecules than proteins, having molecular weights in the millions or billions. Second, nucleic acids consist of only 4 different building blocks (4 bases/nucleic acid) whereas proteins contain 20 different blocks (20 amino acids/protein). To determine the unique sequence of an extremely long chain with only four variables represents a difficult task.

Until about 20 years ago, the only method available for sequencing nucleic acids was the equivalent of the *overlap method* described for proteins in Section 3.2. This approach was first used successfully by Robert Holley in 1965 to sequence one of the smaller RNAs, specifically a 77-base transfer RNA molecule (yeast $tRNA^{Ala}$). Holley chose transfer RNA not just because of its size, but also because it contains many unique nucleotides that made the sequencing somewhat easier.

Following that achievement, nucleic acid sequencing progressed slowly until the mid-1970s, when the ability to determine base sequences changed drastically. At that time, several methods were developed, both chemical and enzymatic, that revolutionized the task of sequencing. In fact, the tables have now been turned so that sequencing a nucleic acid has become easier and faster than sequencing a protein. At times, investigators actually prefer to deduce an amino acid sequence from the base sequence of the corresponding gene rather than determine it experimentally.

In this chapter we discuss one chemical and one enzymatic method for sequencing DNA. Similar methods are available for sequencing RNA. Both Frederick Sanger and Walter Gilbert received the Nobel Prize in 1980 for their work on sequencing DNA. For Sanger, this was the second Nobel Prize, the first having been awarded for his work on sequencing proteins.

7.7.1. Maxam–Gilbert Method

The **Maxam–Gilbert method** represents a chemical method for sequencing, based on reactions that permit the specific modification of one type of base (A, T, G, or C). Chemical modification of a base is followed by selective cleavage of the polynucleotide strand at the modification sites. The method, also called the *chemical cleavage method,* can be divided into the following steps:

1. The DNA is cut into a number of specific fragments by means of one or more restriction enzymes. The double-stranded segments are denatured to separate the two strands, and a specific single-stranded segment is isolated.

2. The 5′-end of this single-stranded segment is labeled with ^{32}P. If the 5′-end carries no phosphate group, the label is added by means of *polynucleotide kinase.* If the 5′-end carries an unlabeled phosphate group, it is necessary to remove this group by means of *alkaline phosphatase* prior to labeling with ^{32}P by use of polynucleotide kinase (Figure 7.29). The 5′-labeled oligonucleotide is divided into four batches.

3. One batch is treated to modify G, one to modify C, one to modify T, and one to modify A. The modification for each base is random and has identical frequency for all bases of one type. Thus, when modifying G in the 10-nucleotide segment of Figure 7.30, an equal chance exists that any one of the four Gs may be modified. Experimenters choose conditions such that, on average, *only one base per segment undergoes modification.* For a batch consisting of 100 molecules that have the sequence ^{32}P-AGTGACGTCG, 25 molecules have the first G altered, 25 the second G, 25 the third G, and 25 the fourth G.

4. Modification of a base renders the phosphodiester bond of the corresponding nucleotide susceptible to hydrolysis. Following base modification, the polynucleotide is *selectively cleaved* at the sites of the modifed bases, yielding mixtures of labeled and unlabeled fragments.

5. The mixture of labeled and unlabeled fragments is separated by *polyacrylamide gel electrophoresis* (*PAGE;* see Appendix C). In this procedure, fragments separate on the basis of their size (number of nucleotides), regardless

Figure 7.30. Principle of the Maxam–Gilbert method for sequencing DNA. Chemical modification of a single strand involves one type of base, for example, G. Modification of G residues is random, and only one residue per strand is altered.

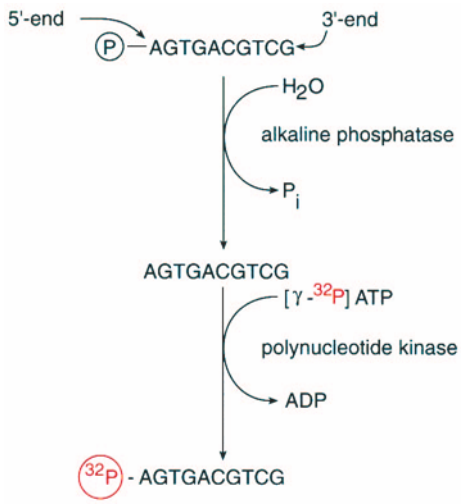


Figure 7.29. Labeling of the 5′-end of ssDNA with ^{32}P. The first step can be omitted if the strand does not carry a 5′-phosphate group.

of the type or sequence of the bases. *The smallest fragment moves the fastest and the farthest* in the gel. All of the fragments carry a negative charge due to the phosphate groups and move toward the positive electrode.

6. The gel is placed against a photographic film, and the radioactively labeled fragments expose the film, marking their positions (*autoradiography*). Only the locations of labeled fragments will be recorded on the film. Unlabeled fragments are not detected, nor do they interfere with detection of the labeled fragments.

7. The base sequence is deduced from the patterns on the films. One can essentially "read" the sequence of bases from the gel patterns. In the example of Figure 7.31, the smallest fragment (the one that has moved farthest) is found in the "G cleavage gel." That fragment must be a mononucleotide. It follows that G must be residue number 2 in the original segment. Proceeding with sequencing, we look for a dinucleotide and find it in the "T cleavage gel" so that residue number 3 must be T, and so on.

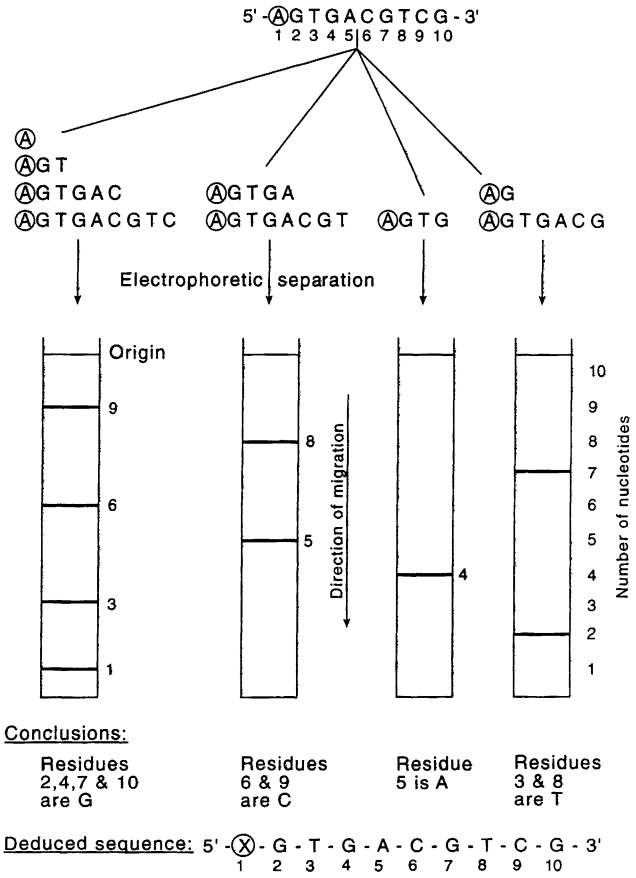


Figure 7.31. Outline of the Maxam–Gilbert method for sequencing DNA. Circles indicate radioactively labeled nucleotides. Base sequence is "read off" by proceeding from bottom to top in the gel patterns. The first base must be identified by other means.

Note that the first residue (A in our example) cannot be identified by this method. To determine it, we can treat the original segment with an exonuclease specific for removal of a 5′-nucleotide and then identify the nucleotide by chromatography. Alternatively, we can find the first residue either by determining the sequence of a strand complementary to the original segment or by using an overlapping fragment of the original segment.

8. In practice, interpreting the "ladder patterns" of the gels is a little more complex. Two reactions modify only G or only C, respectively, and the residues are "read off" as described. However, two other reactions modify both A and G in one case, and both C and T in another. One reaction leads to cleavage of (A + G), and the other to cleavage of (C + T). A residues are identified by comparing the "(A + G) cleavage gel" with the "G cleavage gel." Likewise, T residues are identified by comparing the "(C + T) cleavage gel" with the "C cleavage gel."

7.7.2. Sanger–Coulson Method

The **Sanger–Coulson method** represents an enzymatic method for sequencing, based on the action of *DNA polymerase.* This enzyme catalyzes the synthesis of a single strand of DNA, complementary to a second strand, by using deoxynucleoside triphosphates as substrates. The original method has been modified by use of 2′,3′-dideoxynucleoside triphosphates and is also known as the *Sanger dideoxynucleotide* or *chain terminator method.*

2′,3′-Dideoxynucleoside triphosphates are nucleoside triphosphates that lack a hydroxyl group at both the 2′ and 3′ positions (Figure 7.32). Whenever a dideoxynucleoside triphosphate is incorporated by the enzyme into a growing strand, the next 3′,5′-phosphodiester bond cannot be formed, and the *nascent* polynucleotide strand cannot be extended further. Accordingly, 2′,3′-dideoxynucleoside triphosphate acts as chain terminator. The method can be divided into the following steps (Figure 7.33):

1. The DNA is cut into a number of specific fragments by means of one or more restriction enzymes. The double-stranded segments are denatured to separate the two strands, and a specific single-stranded segment is isolated.

2. A small section of a DNA *primer* is hydrogen-bonded to the 3′-end of the single-stranded segment to be sequenced (*template*). The primer/template complex is divided into four batches.

3. Each batch is treated with DNA polymerase, which catalyzes extension of the primer strands; the nucleotide sequence synthesized is complementary to the template strand. The reaction mixture contains all the necessary four deoxynucleoside triphosphates plus a small amount of a single dideoxynucleoside triphosphate that acts as chain terminator. A radioactive label is included in the reaction mixture so that all newly synthesized fragments will be labeled. The label can be in the form of the dideoxynucleoside triphosphate used for each batch. Alternatively, a labeled primer (TAT) can be used throughout. A recent advance in this method employs fluorescent derivatives of the primer.

4. Each reaction mixture results in a mixture of double-stranded segments in which the primer has been ex-

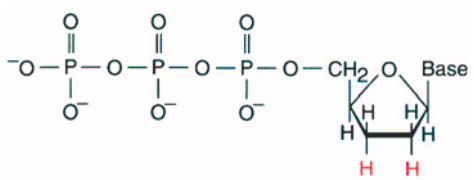


Figure 7.32. A dideoxynucleotide (2′,3′-dideoxynucleoside triphosphate).

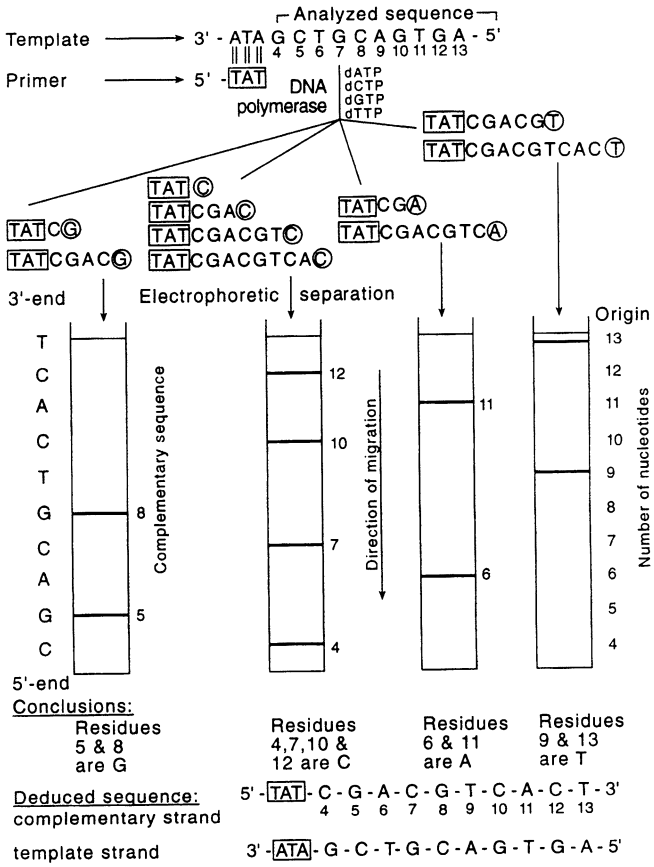


Figure 7.33. Outline of the Sanger–Coulson method for sequencing DNA. Circles indicate dideoxynucleotides (dd). Base sequence is "read off" by proceeding from bottom to top in the gel patterns.

ments are linked covalently to the primer, and the extended primer is linked noncovalently to the template. Each extended primer strand is labeled and is automatically terminated whenever a dideoxynucleoside triphosphate is incorporated in place of the normal nucleotide.

5. The mixture of double-stranded fragments is denatured in order to separate the strands and then subjected to gel electrophoresis, followed by autoradiography as in steps 5 and 6 of the Maxam–Gilbert method.

6. The base sequence is deduced from the patterns on the films. Note that in this procedure the entire segment, including the first position, can be read off from the gel.

SUMMARY

DNA and RNA are polymers, composed of nucleotide building blocks. Each nucleotide consists of a sugar (ribose for RNA, deoxyribose for DNA), linked via a β-glycosidic bond to a base (purine or pyrimidine), with a phosphate group esterified at the C(5′) position of the sugar. The phosphate group of each nucleotide serves to link individual nucleotides together in polynucleotides and nucleic acids by being esterified to the 3′-OH of an adjacent nucleotide. The bond formed constitutes a 3′,5′-phosphodiester bond. Nucleases catalyze the hydrolysis of this bond on either the 3′- or the 5′-side. The purines adenine and guanine occur in

both DNA and RNA. The pyrimidines cytosine and thymine occur in DNA whereas cytosine and uracil occur in RNA.

Nucleic acids vary greatly in their size and associate with proteins noncovalently to form nucleoproteins like chromatin, ribosomes, and viruses. Major forms of RNA are ribosomal, transfer, and messenger RNA, all of which function in protein synthesis. DNA is the carrier of genetic information in all bacteria, plants, and animals. Some viruses have DNA as their genetic material, whereas others have RNA. Generally, the amount of DNA and the size of the DNA molecule increase with the complexity of the organism. DNA is packaged efficiently in the nucleoid region of prokaryotes and in the chromosomes of eukaryotes.

The DNA double helix (B-DNA) consists of two intertwined, antiparallel, right-handed helices of polynucleotides. Sugar and phosphate form the backbone of each strand, with the phosphate groups exposed at the surface of the helix. The bases are stacked in the interior of the helix, forming flat base pairs with one base contributed by each strand. Base pairing is complementary: adenine is linked to thymine (two H-bonds), and guanine is linked to cytosine (three H-bonds). Hydrophobic stacking interactions between base pairs contribute to helix stability. Other duplex configurations of DNA are known, and a segment of single-stranded RNA can fold back on itself to take on a double-stranded configuration.

DNA can be denatured by agents that break hydrogen or hydrophobic bonds, by extremes of pH, by suspension at low ionic strength, and by heat. Thermal denaturation is commonly followed by measuring changes in ultraviolet absorbance. These can be used to determine the base composition [mole percent of (G + C)] of the DNA. Denatured DNA can be renatured by slow cooling. A mixture of denatured nucleic acids, upon slow cooling, forms some hybrid double-stranded structures. Hybridization provides information about the degree of similarity among nucleic acids.

Prokaryotes possess restriction endonucleases that catalyze cleavage of foreign DNA at a few specific sites. Modification methylases catalyze methylation at these same sites in cellular DNA, thereby protecting the cell's DNA against restriction endonucleases. Both DNA and RNA can be sequenced rapidly by either chemical or enzymatic methods.

SELECTED READINGS

Adams, R. L., Knowler, J. T., and Leader, D. P., *The Biochemistry of the Nucleic Acids,* 11th ed., Chapman & Hall, London (1992).

Agris, P. F., The importance of being modified: Roles of modified nucleosides and Mg^{2+} in RNA structure and function, *Prog. Nucleic Acid Res. Mol. Biol.* 53:79–129 (1996).

Avery, O. T., MacLeod, C. M., and McCarty, M., Studies on the chemical transformation of Pneumococcal types, *J. Exp. Med.* 79:137–158 (1944).

Cold Spring Harbor Symposia on Quantitative Biology, *DNA and Chromosomes,* Vol LVIII, Cold Spring Harbor Laboratory Press, Plainview, New York (1993).

Gesteland, R. F., and Atkins, J. F. (eds.), *The RNA World,* Cold Spring Harbor Laboratory Press, Plainview, New York (1993).

Glick, B. R., and Pasternak, J. J., *Molecular Biotechnology: Principles and Applications of Recombinant DNA,* ASM Press, Washington, D.C. (1994).

Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M., Diversity of oligonucleotide functions, *Annu. Rev. Biochem.* 64:763–797 (1995).

Herbert, A., and Rich, A., The biology of left-handed Z-DNA, *J. Biol. Chem.* 271:11595–11598 (1996).

Holley, R. W., The nucleotide sequence of a nucleic acid, *Sci. Am.* 214:30–46 (1966).

Nicholson, A. W., Structure, reactivity, and biology of double-stranded RNA, *Prog. Nucleic Acid Res. Mol. Biol.* 52:1–65 (1996).

Rennie, J., DNA's new twists, *Sci. Am.* 268:122–132 (1993).

Watson, J. D., *The Double Helix,* Atheneum, New York (1968).

Watson, J. D., and Crick, F. H. C., Molecular structure of nucleic acid. A structure for deoxyribose nucleic acid, *Nature (London)* 171: 737–738 (1953).

Von Hippel, P. H., Protein–DNA recognition: New perspectives and underlying themes, *Science* 263:769–770 (1994).

Zlatanova, J., and Van Holde, K., The linker histones and chromatin structure, *Prog. Nucleic Acid Res. Mol. Biol.* 52:217–259 (1996).

REVIEW QUESTIONS

A. Define each of the following terms:

Plectonemic coiling
Ribozyme
Transforming principle
Complementary base pairing
Renaturation
Nucleotide map
Base composition
Antiparallel strands
Chargaff's rules
Thermal denaturation
Palindrome
Melting temperature

B. Differentiate between the two terms in each of the following pairs:

Restriction enzyme/modification methylase
RNA/DNA
B-DNA/Z-DNA
Double helix/superhelix
Nuclease/nucleosome
Right-handed helix/left-handed helix
Base/modified base
Nucleoside/nucleotide
Purine/pyrimidine
Chromosome/chromatosome

C. (1) Outline the steps of the chemical cleavage and chain terminator methods of DNA sequencing.
(2) What are some major properties of (a) purines and pyrimidines and (b) RNA and DNA?
(3) Draw the following structures: (a) 5-Methylcytosine; (b) N^6-dimethyladenine; (c) 5′-guanylic acid; (d) deoxycytidine 3′,5′-bisphosphate; (e) thymidine 5′-triphosphate; (f) pApTpC.
(4) Outline the experiment of Avery, MacLeod, and McCarty and that of Hershey and Chase. What could be concluded from each?
(5) What are the main features of the Watson–Crick DNA double helix?

PROBLEMS

7.1. Propose a hypothetical mechanism whereby a mutated DNA, in which a specific base has been chemically altered, could be repaired.

7.2. What is the mole percent of adenine in a double-stranded DNA that contains 20 mol % of cytosine?

7.3.* An analytical determination can detect as little as 2.0 μg of phosphorus (atomic weight = 31). What is the minimum number of micromoles of polyadenylic acid (pApApA . . .) you must use to be able to detect phosphorus by this method? The molecular weights of AMP and polyadenylic acid are 345 and 10,000, respectively.

7.4. Write out the base sequence for the complementary strand to the following ssDNA fragment:

7.5. Consider the ribonucleotide pCGUACGAUGAGU–OH. What fragments do you obtain when you hydrolyze this compound by means of each of the following? (a) ribonuclease A; (b) ribonuclease T_1; (c) nuclease S_1

7.6. Given that the pK'_a values of H_3PO_4 are 2.12, 7.21, and 12.4, what is the approximate pK'_a of the first proton to be dissociated from the phosphate group of 5′-AMP?

7.7. Draw all possible enol forms of thymine.

7.8. Calculate the base composition of the longest double-stranded fragment that can be formed from the template/primer shown at the top of Figure 7.33.

7.9. We call organisms that grow at moderate temperatures (about 20–45°C) mesophiles, and those that grow at high temperatures (about 45–70°C or higher) thermophiles. What difference might you expect to find between the average DNA base composition of a number of mesophiles and that of a number of thermophiles? Why is this so?

7.10.* Plot the expected changes in viscosity and optical rotation for dsDNA as a function of increasing temperature.

7.11. The ultraviolet absorbance at 260 nm of a solution of dsDNA increases by 40% as the temperature rises from 25 to 95°C. A solution of ssRNA, of identical concentration, shows an increase in absorbance of 5% under the same conditions. Estimate the fraction of bases in the RNA that are in double-stranded configuration.

7.12.* A dsDNA segment has 120 base pairs. Calculate the length of the segment and the number of helical turns that are in it if the segment is: (a) A-DNA; (b) B-DNA; (c) Z-DNA.

7.13.* What are the minimum and maximum numbers of H-bonds that could be formed in a segment of double-stranded B-DNA that has a length of 34 nm?

7.14. You can see from Table 7.3 that lungfish DNA is longer than human DNA. Does this mean that a lungfish carries more genetic information than a human? Explain your answer.

7.15. The mitochondrial human genome is a circular dsDNA that has a length of 16,569 bp. How do these

facts tie in with the evolution of eukaryotes discussed in Section 1.1?

7.16. A viral DNA has the following base composition (in mole percent): A = 15, G = 32, C = 13, and T = 40. What can you conclude from these data?

7.17.* After exposing a bacterial culture to several mutagens, a researcher isolated the DNA and determined its base composition. The double-stranded DNA contained 30.2 mol % A and 15.8 mol % G. The [A]/[T] ratio was 0.95, and the [C]/[G] ratio was 0.93. Calculate the mole percent of cytosine and thymine in the DNA.

7.18. Calculate the packaging ratio of phage T2 DNA, given that the phage head is about 100 nm long.

7.19. A dsDNA has a melting temperature (T_m) of 80°C when measured in a solution containing 0.15*M* NaCl and 0.015*M* sodium citrate (pH 7). What is the mole percent of the DNA's (A + T)?

7.20. The first four nucleotides of a palindrome, consisting of eight base pairs, are 5′-ACGG-3′. What is the palindrome's complete base sequence?

7.21. The phosphate groups of nucleoside mono-, di-, and triphosphates have pK'_a values of 1.0 and 6.0 for the first and second protons, respectively. On this basis, determine the net charge of: (a) ATP at pH 7.0; (b) TTP at pH 12; (c) UDP at pH 11; (d) GMP at pH 2.0. (Refer to Figure 7.3.)

7.22. What is the mass, in grams, of a single base pair (MW = 670; $N = 6.02 \times 10^{23}$)? What is the mass of the complete genome of a single human diploid cell (5.8×10^9 bp)?

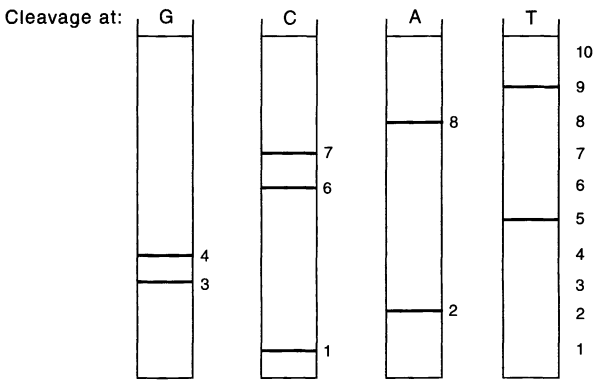
7.23.* Sketch the gel electrophoresis patterns that you would obtain by using the Sanger–Coulson method to sequence the following template/primer fragment:

3′-GCATCGACTGATGCA-5′

5′-CGTAG-3′

7.24.* You isolate ribosomal RNA (rRNA) from *E. coli,* fragment it, and label it with a radioactive isotope. You then allow the rRNA to hybridize with denatured ds-DNA from the same organism. You find that about 2% of the DNA forms hybrid structures with the rRNA. What fraction of the DNA genome in *E. coli* carries the genetic information for making ribosomal RNA?

7.25. A 10-nucleotide fragment has adenine at its 5′-end and is sequenced by the chemical cleavage method. The patterns obtained are shown below. What is the sequence of the fragment?



Metabolism

This part of the book deals with the multitude of reactions whereby biomolecules are synthesized, degraded, and interconverted in living systems. Energy transformations constitute integral aspects of these processes. Some reactions produce usable energy, whereas others require an input of energy to proceed. To provide an overall view, we consider both synthesis and degradation of a given type of biomolecule in a single chapter. We will cover major pathways of carbohydrate, lipid, amino acid, and nucleotide metabolism. Because of their unique links to DNA, we will discuss the metabolism of proteins and nucleic acids separately in Part IV.

Introduction to Metabolism

8

Metabolism comprises all of the chemical reactions that take place in a living system, be it a cell, a tissue, an organ, or an organism. Metabolic reactions are almost all enzyme-catalyzed and include transformations of energy and nutrients, syntheses and degradations, and excretions of waste products. Chemical changes concerned with the production, storage, and utilization of metabolic energy for biosynthesis are known as **intermediary metabolism.** Typically, intermediary metabolism includes all aspects of metabolism (including digestion) except those involved in the transfer of genetic information—replication, transcription, and translation.

At the level of an individual organism, metabolic reactions form the basis of all the functions—growth, movement, reproduction, and the like—characteristic of living systems. On a larger scale, the combined metabolism of microorganisms, plants, and animals leads to the great natural cycles of the biosphere (Figure 8.1).

The Sun is the ultimate source of energy for metabolism in all organisms; it constitutes the source of life on Earth. Plants trap solar energy during photosynthesis, the single most important chemical reaction in the biosphere. Photosynthesis results in conversion of carbon dioxide to carbohydrates, which then provide carbon skeletons for the synthesis of proteins, lipids, and nucleic acids. Nonphotosynthetic organisms feed on plants and use plant products for their growth and reproduction.

8.1. CATABOLISM AND ANABOLISM

We divide metabolism into two parts, **catabolism,** or degradative reactions, and **anabolism,** or synthetic reactions. Catabolism consists of three stages (Figure 8.2). In the first stage, polymeric nutrients break down to small monomeric building blocks. In animals, this stage comprises digestion, as a result of which carbohydrates, proteins, lipids, and nucleic acids are degraded to, respectively, monosaccharides (principally glucose), amino acids, fatty acids and glycerol, and nucleotides.

The second stage produces only a few different types of molecules. These products have even simpler structures than the building blocks from which they derive and include the two key compounds *pyruvate* and *acetate.* Acetate enters metabolism in the form of *acetyl coenzyme A.* Protein catabolism also yields ammonia, formed by deamination of amino acids.

In the third stage, acetyl coenzyme A enters the *cit-*

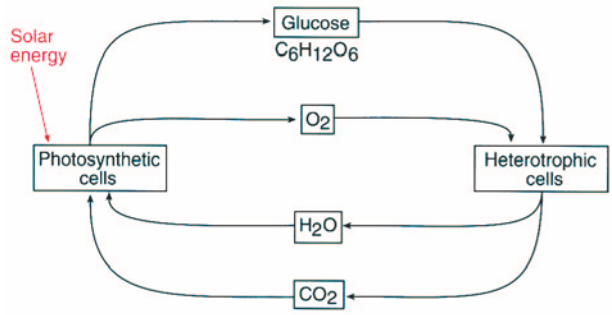


Figure 8.1. The carbon and oxygen cycles of the biosphere.

ric acid cycle, where its acetyl group is oxidized to carbon dioxide and water. All told, catabolism yields only three major end products—CO_2, H_2O, and NH_3—in addition to the capture of chemical energy in the form of ATP. The number of end products is small because of the *convergent* nature of catabolic reactions; multiple pathways of nutrient degradation come together and ultimately enter one major metabolic system.

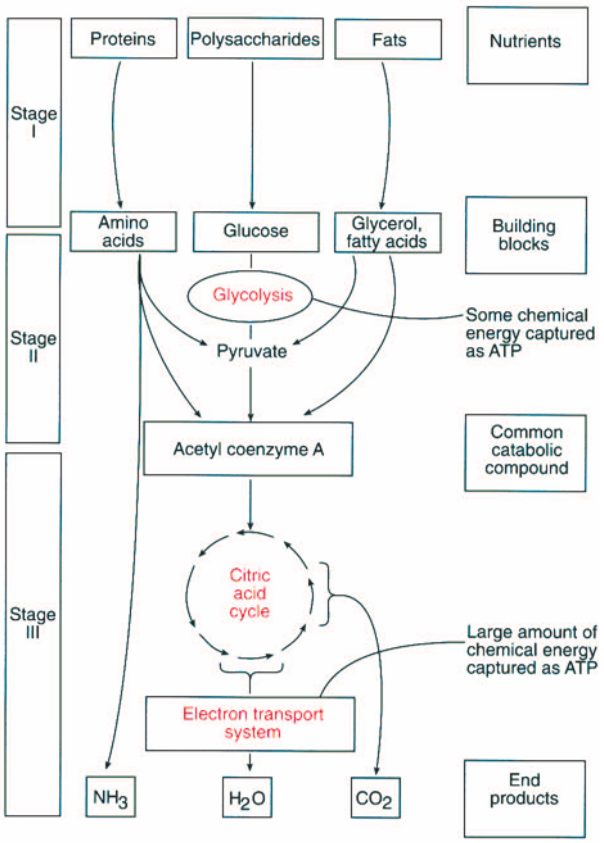


Figure 8.2. The three stages of catabolism for the major energy-yielding nutrients. Nucleic acids are also degraded in three stages, but they have been omitted from the figure because their breakdown does not contribute significantly to the cell's energy supply.

By contrast, anabolic reactions have *divergent* characteristics; a simple precursor gives rise to numerous larger molecules by reactions that fan out from the initial step (Figure 8.3).

Both catabolism and anabolism involve hundreds of different reactions, all interconnected in a giant network (Figure 8.4). As you can see, most metabolic intermediates can be converted to several compounds by reactions that proceed along different paths. Not surprisingly, therefore, interconversions between major nutrients occur readily. Carbohydrates can yield proteins, proteins can form lipids, and so on. We will discuss specific examples of such interrelationships later on.

Studying this multitude of reactions may appear to be an insurmountable task, but that is not so. Without having to discuss each and every reaction, you will find that it is possible to understand the operation of the major metabolic systems by focusing on a few groups of closely linked reactions. In the following chapters, we deal with these unifying concepts of metabolism.

8.1.1. Fermentation

Sets of some catabolic reactions constitute fermentations. We define **fermentation** as an energy-yielding catabolic pathway that proceeds without net oxidation; oxidation of one intermediate is balanced by reduction of another. In fermentation, organic compounds act as both donors and acceptors of electrons. Fermentation reactions do not require molecular oxygen and yield energy in the form of *ATP,* an *energy-rich compound,* discussed in Section 9.2. Two examples will illustrate the process.

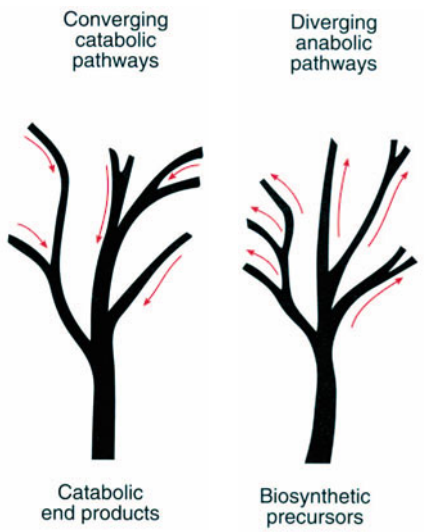


Figure 8.3. Converging and diverging metabolic pathways.

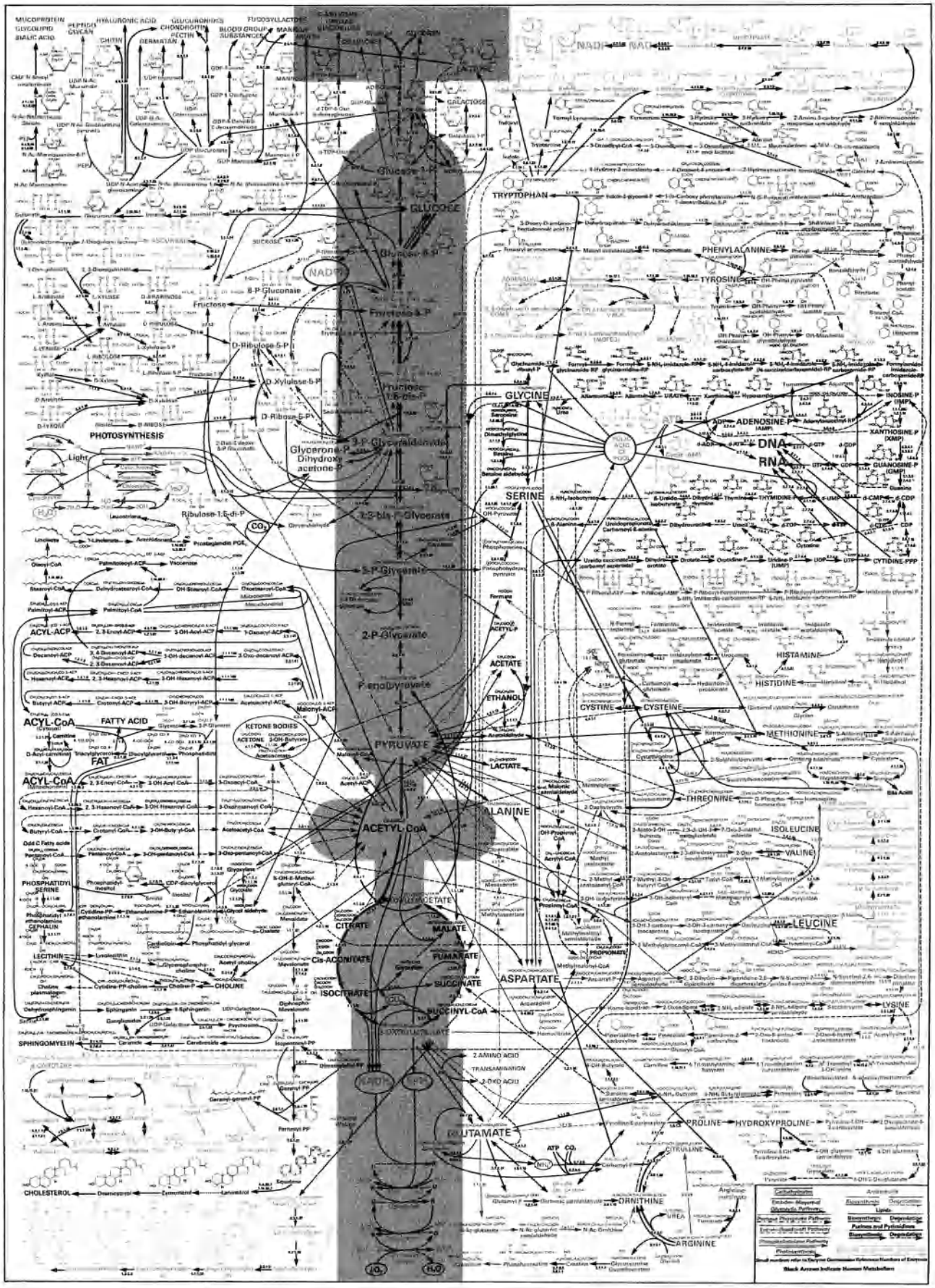


Figure 8.4. Schematic map of the major metabolic pathways in a typical cell. The shaded region encompasses the main reactions of glucose metabolism. (Designed by D. E. Nicholson. Published by BDH Ltd., Poole, Dorset, England.)

8.1.1A. Alcoholic Fermentation. In **alcoholic fermentation,** yeast converts glucose to ethanol. *Alcohol dehydrogenase* catalyzes the last reaction in this sequence:

$$\underset{\text{Acetaldehyde}}{CH_3-\overset{\overset{O}{\|}}{C}H} + NADH + H^+ \rightarrow \underset{\text{Ethanol}}{CH_3CH_2OH} + NAD^+$$

Acetaldehyde is *reduced* to ethanol while NADH is *oxidized* to NAD^+. The overall reaction of alcoholic fermentation includes the synthesis of ATP from ADP:

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + 2P_i^{2-} + 2ADP^{3-} + 2H^+ \rightarrow \underset{\text{Ethanol}}{2CH_3CH_2OH} + 2CO_2 + 2ATP^{4-} + 2H_2O$$

8.1.1B. Lactate Fermentation. **Lactate fermentation** involves the catabolism of glucose to lactate (the anionic form of lactic acid). In microorganisms, lactate is formed by lactic acid bacteria. In humans, lactate forms under relatively anaerobic conditions, such as those that occur during strenuous exercise. *Lactate dehydrogenase* catalyzes the last reaction in this process:

$$\underset{\text{Pyruvate}}{CH_3-\overset{\overset{O}{\|}}{C}-COO^-} + NADH + H^+ \rightarrow \underset{\text{Lactate}}{CH_3-CHOH-COO^-} + NAD^+$$

Pyruvate is *reduced* to lactate while NADH is *oxidized* to NAD^+. The overall reaction of lactate fermentation also includes the synthesis of ATP from ADP:

$$\text{Glucose} + 2P_i^{2-} + 2ADP^{3-} + 2H^+ \rightarrow 2 \text{ lactate} + 2ATP^{4-} + 2H_2O$$

8.2. REGULATION OF METABOLISM

Because almost all metabolic reactions are enzyme-catalyzed, the primary regulation of metabolism occurs by controlling enzymatic activity. This control, as you saw earlier, involves many different factors and mechanisms (Figure 4.6). Recall that a typical cell contains thousands of different kinds of substrates and enzymes (Table 1.1). Regulating the multitude of intracellular reactions that use these substrates and enzymes amounts to an astounding feat. When you consider that in large multicellular organisms there exists the added element of coordinating the reactions of billions of cells, the regulation of metabolism emerges as a staggering evolutionary achievement.

In addition to regulating individual enzymes, metabolic control involves three major elements—the existence and regulation of specific *metabolic pathways,* the occurrence of *compartmentation,* and the action of *hormones.*

8.2.1. Metabolic Pathways

A **metabolic pathway** consists of a sequence of steps by which a metabolite is synthesized, degraded, or transformed. The sequence proceeds from some key intermediate to a specific terminal product. Pathways may be *linear, branched,* or *cyclic* (Figure 8.5).

Metabolic pathways typically involve a large number of steps, best described by steady-state rather than by equilibrium conditions (Section 4.5). Multiple steps mean that alterations of intermediates generally occur in small increments. In this way, catabolism releases energy in amounts that can be used for efficient ATP synthesis. Metabolic pathways frequently commence by effectively *trapping* the first key intermediate inside the cell or by *activating* it to a chemically more reactive form. All metabolic pathways have three important characteristics in common.

First, *metabolic pathways are irreversible.* Catabolic pathways constitute "downhill" sequences in which one or more reactions are strongly *exergonic,* resulting in the release of free energy ($\Delta G'$ is negative). Hence, the reverse pathway, the anabolic sequence, constitutes an "uphill" process in which these same reactions are strongly *endergonic* ($\Delta G'$ is positive). However, endergonic reactions are thermodynamically not feasible and cannot pro-

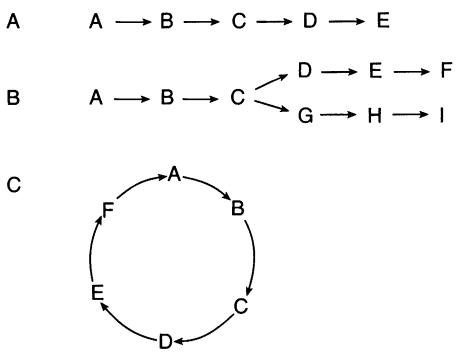


Figure 8.5. Types of metabolic pathways: (A) linear; (B) branched; (C) cyclic.

ceed as written. Accordingly, these reactions serve as energy barriers for the anabolic pathway. Because of them, the pathway cannot occur even though some of its steps may be readily reversible.

In order for the anabolic sequence to become feasible, the pathway must be altered (Figure 8.6). Each highly endergonic step must be changed by replacing it with one or several different reactions. Anabolic and catabolic sequences must differ in the reactants and products of these altered steps since only then will the steps have different free energy changes associated with them. Recall that a reaction's free energy change is a function of the reactants and the products but does not depend on the path whereby reactants are converted to products.

Second, *metabolic pathways have committed steps.* A **committed step** is a unique step of a pathway. It may be a step producing a metabolite that has no other role than to serve as an intermediate in this pathway. In most instances, the committed step constitutes a highly exergonic and essentially irreversible reaction. The committed step ensures that, once it has taken place, all subsequent reactions in the sequence will also take place; it "commits" the intermediate it produces to continue down the pathway. Committed steps occur early in metabolic pathways and frequently represent rate-determining steps.

Lastly, *metabolic pathways are regulated.* Regulation of pathways may be accomplished by multienzyme systems, allosteric enzymes, activators and inhibitors, and the like. More than one factor may be operative at a given step, and frequently more than one step of the pathway is subject to control. A committed step usually represents the most important control element of the pathway. Because catabolic and anabolic sequences differ, they can be regulated independently of each other. Separate regulation of degradation and synthesis of key intermediates allows for fine-tuned control of metabolism.

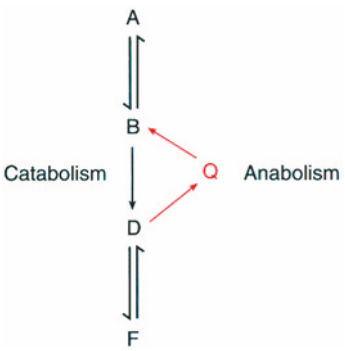


Figure 8.6. Irreversibility of catabolic pathways. The catabolic sequence A → B → D → F has two readily reversible steps but one exergonic and irreversible step (B → D). To overcome the energy barrier presented by this step, the anabolic sequence must proceed via a different path, at least at this step. Introduction of the Q bypass allows anabolism to proceed from F to A.

8.2.2. Compartmentation

Compartmentation represents a second important element in the regulation of overall metabolism. The term denotes an unequal distribution of a metabolite, an enzyme, a pathway, or some other biomolecule or system within a cell or organism. Compartmentation affects metabolic reactions in several ways. In eukaryotes, it frequently accomplishes complete segregation of entire metabolic pathways in specific subcellular locations. For example, fatty acid catabolism occurs in the mitochondria, but fatty acid biosynthesis takes place in the cytoplasm. Likewise, production of ATP occurs in the mitochondria, but utilization of ATP takes place largely in the cytoplasm.

Separation of degradative and synthetic processes is advantageous because otherwise the two opposing activities would cancel each other out, in part or in entirety. We term such a set of opposing reactions a *futile cycle* (see Figure 10.35); it achieves nothing except the dissipation of free energy and, possibly, the generation of some heat.

Compartmentation also regulates enzymatic activity via the permeability properties of the compartment. By being selectively permeable, a compartment membrane controls the entry of substrates into and the exit of products from the compartment. Relative concentrations of substrates and products in turn influence enzymatic activities.

Lastly, compartmentation is linked to the action of several hormones that affect the transfer of metabolites across cell and subcellular membranes. In animals, all of these different aspects of compartmentation are related in a complex fashion. Compartmentation involves not only specific organs such as kidney, liver, stomach, and intestine, each of which is responsible for particular aspects of metabolism, but also specific subcellular locations of biochemical functions.

Mitochondria exhibit a high degree of compartmentation. Mitochondria possess two membranes, an inner and an outer one, separated by an intermembrane space (Figure 8.7). The inner membrane has many infoldings called *cristae.* Filling the spaces between cristae and surrounded by the inner membrane is the *matrix,* a gel-like substance. Thus, there exist four compartments in a mitochondrion: *outer membrane, intermembrane space, inner membrane,* and *matrix.* Each compartment contains specific metabolic systems and one or more enzymes that are located only in that compartment. We call such enzymes,

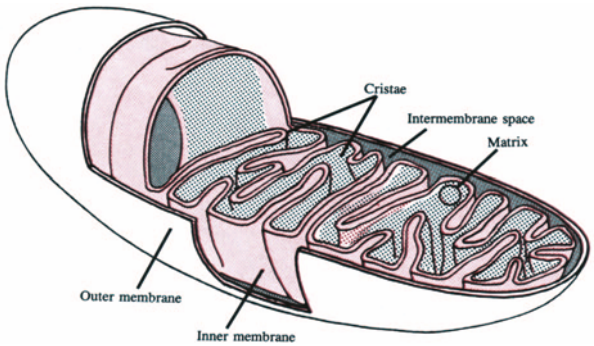


Figure 8.7. Schematic structure of a mitochondrion. (Adapted, with permission, from S. L. Wolfe, *Biology of the Cell,* © 1972 by Wadsworth Publishing Co., Belmont, California.)

located *exclusively* in certain parts of a cell, **marker enzymes** (Table 8.1). Because of its exclusive location, a marker enzyme can serve to identify an experimentally obtained cellular fraction. Thus, if a fraction possesses cytochrome oxidase activity, you could assume that it contains inner mitochondrial membrane fragments.

8.2.3. Hormones

Hormone action constitutes a third key factor in the regulation of overall metabolism. According to their classic definition, hormones comprise regulatory substances in plants and animals that are synthesized by specialized cells, are active at low concentrations, and affect either all cells of an organism or only certain target cells in specific organs. Biochemists have expanded this definition to include any substance that carries a signal to produce some change at the cellular level. We now classify hormones based on the distance over which they act: an *autocrine hormone* acts on the same cell that produced it (e.g., *interleukin*); a *paracrine hormone* acts on cells in close proximity to the one that produced it (e.g., *prostaglandins*); and an *endocrine hormone* acts on cells remote from the one that produced it (e.g., *insulin*).

Table 8.1. Location of Some Metabolic Activities and Marker Enzymes in Mitochondria

Location	Metabolic activities or marker enzymes
Outer membrane	Monoamine oxidase[a]; phospholipid synthesis; fatty acid elongation and desaturation
Intermembrane space	Adenylate kinase[a]; nucleoside diphosphate kinase
Inner membrane	Electron transport system; oxidative phosphorylation; succinate dehydrogenase[a]; cytochrome oxidase[a]; ATP synthase
Matrix	Citric acid cycle; β-oxidation of fatty acids; citrate synthase[a]; fumarase[a]; pyruvate dehydrogenase complex; urea cycle

[a]Marker enzyme.

Animal hormones are secreted principally by ductless glands and transported via the circulation to target tissues or organs. There they exert their effects either directly or indirectly and help regulate overall physiological processes like metabolism, growth, and reproduction. Plant hormones are organic compounds that control growth or some other function at a site removed from their place of production in the plant.

Chemically speaking, there exist four main classes of hormones: polypeptides, steroids, amino acid derivatives (amines), and eicosanoids. Table 8.2 lists major human hormones.

Many hormonal systems in animals originate in the brain, where an environmental or internal stimulus triggers the first signal (Figure 8.8). From the central nervous system the signal may be transmitted as an electrical pulse or as a chemical signal—or both—and then lead to release of a hormone. Frequently, a number of hormones act on each other in a *cascade mechanism* (Figure 4.24), thereby *amplifying* the initial signal. Reaction of the *ultimate hormone* with target cells produces particular metabolic changes. Hormones act on these target cells by binding to specific sites, called **receptors,** located on the cell membrane or in the interior of responsive cells.

Biochemists recognize four classes of hormone receptors that differ in their location and in the mechanism of action of the corresponding hormone. Some receptors are located on the cell periphery and bind hormones there. Binding leads to synthesis and/or release of a compound inside the cell that evokes a metabolic change by acting on a specific enzyme. We refer to the compound released inside the cell as a **second messenger** because we consider the hormone binding to the cell membrane as the **first messenger.** We will discuss specific hormones and second messengers as we encounter them in the text.

The second class of receptors comprises receptors that are located on cytoplasmic or nuclear proteins and bind hormones capable of traversing the cell or nuclear membrane. The remaining two classes of receptors constitute transmembrane proteins. One type of transmembrane protein has a hormone binding site on its extracellular side and a catalytic site on its intracellular side; the cytoplasmic site becomes activated when the hormone binds to the external site. The second type of transmembrane protein functions as an ion channel. Upon binding of the hormone to the cell periphery, the channel opens and specific ions flow across the membrane.

Table 8.2. Selected Human Hormones

Hormone	Source	Major effects
Polypeptides		
Corticotropin-releasing factor (CRF)	Hypothalamus	Stimulates ACTH release
Gonadotropin-releasing factor (GnRF)	Hypothalamus	Stimulates LH and FSH release
Growth hormone-releasing factor (GRF)	Hypothalamus	Stimulates GH release
Thyrotropin-releasing factor (TRF)	Hypothalamus	Stimulates TSH release
Adrenocorticotropic hormone (ACTH)	Anterior pituitary	Stimulates release of adrenocorticosteroids (glucocorticoids and mineralocorticoids)
β-Endorphin	Anterior pituitary	Exerts opioid effects on central nervous system
Follicle-stimulating hormone (FSH)	Anterior pituitary	Stimulates ovulation and estrogen synthesis in ovaries; stimulates spermatogenesis in testes
Growth hormone (somatotropin, GH)	Anterior pituitary	Stimulates bone growth and release of insulin and glucagon
Leu-enkephalin	Anterior pituitary	Exerts opioid effects on central nervous system
Luteinizing hormone (LH)	Anterior pituitary	Stimulates estrogen and progesterone synthesis in ovaries; stimulates androgen synthesis in testes
Met-enkephalin	Anterior pituitary	Exerts opioid effects on central nervous system
Prolactin (PRL)	Anterior pituitary	Stimulates lactation of mammary gland
Thyrotropin (TSH)	Anterior pituitary	Stimulates T_3 and T_4 release
Oxytocin	Posterior pituitary	Stimulates uterine contractions
Vasopressin	Posterior pituitary	Stimulates water resorption by the kidney and increases blood pressure
Glucagon	Pancreas (α cells)	Stimulates glucose release by glycogen breakdown
Insulin	Pancreas (β cells)	Stimulates glucose uptake by cells from the blood
Steroids		
Glucocorticoids (e.g., cortisol)	Adrenal cortex	Decrease inflammation, increase resistance to stress
Mineralocorticoids (e.g., aldosterone)	Adrenal cortex	Maintain salt and water balance
Androgens (e.g., testosterone)	Gonads and adrenal cortex	Promote development of secondary sex characteristics, particularly in males
Estrogens (e.g., estrone)	Gonads and adrenal cortex	Promote development of secondary sex characteristics, particularly in females
Amino acid derivatives		
Epinephrine (adrenaline)	Adrenal medulla	Increases heart rate and blood pressure
Norepinephrine	Adrenal medulla	Decreases peripheral circulation
Thyroxine (T_4)	Thyroid	Stimulates metabolism
Triiodothyronine (T_3)	Thyroid	Stimulates metabolism
Eicosanoids		
Prostaglandins	Cell membranes	Stimulate smooth muscle contraction; regulate inflammatory reactions

8.2.4. Genetic Diseases

Although metabolism is carefully regulated, abnormalities do occur. Some abnormalities result from major dietary changes or specific pathological states. Other abnormalities result from *mutations* in the genetic makeup of an organism and constitute **genetic** or **hereditary diseases.** Each genetic disease reflects a *deficiency* of a specific enzyme or other protein.

The deficiency may result from one of two causes. First, the specific enzyme or other protein *may not be synthesized at all.* In this case, the gene coding for the enzyme or protein is either missing or not expressed. Second, the enzyme or other protein may be synthesized, but *in an altered and ineffective form.* For example, an enzyme may be formed with a nonfunctional active site.

In some instances, we know the nature of the defect resulting in the genetic disease. We might be able to distinguish between the two causes of deficiency by determining whether or not antibodies prepared against the protein produce an antigen–antibody reaction with cell-free extracts expected to contain the protein. A positive antigen–antibody reaction indicates that the enzyme or other protein (the antigen) is probably present but does not react properly with its substrate.

Whatever the underlying reason, an enzyme or pro-

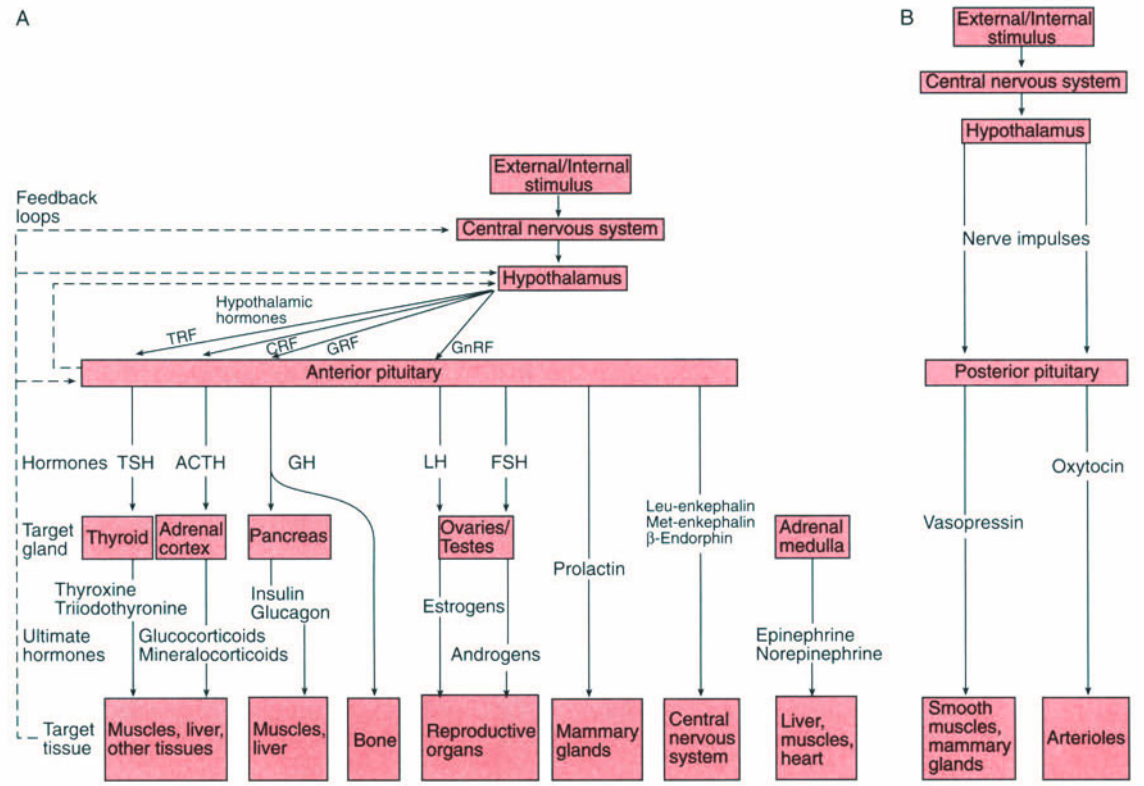


Figure 8.8. The hormonal control network in animals involves both the anterior (A) and the posterior (B) pituitary gland.

Table 8.3. Some Human Genetic Diseases

Disease	Symptoms	Defective enzyme or protein
Acanthocytosis	Abnormal erythrocytes with projecting spines	Low-density lipoprotein
Albinism	Lack of skin pigmentation	Tyrosinase
Analbuminemia	Impaired synthesis of serum albumin	Serum albumin
Argininemia	Elevated levels of arginine in blood and urine	Arginase
Diabetes	Deranged carbohydrate metabolism	Insulin
Dwarfism	Premature arrest of skeletal growth	Growth hormone
Galactosemia	Inability to convert galactose to glucose metabolites	Galactokinase
Gaucher's disease	Accumulation of cerebrosides in tissues	Glucocerebrosidase
Glycogen storage disease II	Accumulation of normal glycogen in all organs	α-1,4-Glucosidase
Glycogen storage disease IV	Glycogen with very long unbranched chains in liver and spleen	Branching enzyme
Hemolytic anemia	Excessive destruction of red blood cells	Glucose 6-phosphate dehydrogenase
Lactose intolerance	Inability to metabolize lactose	Lactase
Methylmalonic acidemia	Massive ketosis	Methylmalonyl CoA mutase
Niemann–Pick disease	Sphingomyelin accumulation and mental retardation	Sphingomyelinase
Phenylketonuria	Inability to hydroxylate phenylalanine, associated with mental retardation	Phenylalanine hydroxylase
Sickle-cell anemia	Hemolysis of sickle-shaped red blood cells	Hemoglobin
Tangier disease	Almost complete absence of high-density lipoprotein (HDL) and accumulation of cholesterol	High-density lipoprotein
Tay–Sachs disease	Accumulation of gangliosides, blindness, and brain deterioration	*N*-Acetylhexosaminidase

tein deficiency results in a **metabolic block** so that a specific reaction cannot proceed. Because of the block, products of the reaction do not form, reactants accumulate, and a set of disease symptoms may develop. Attempts are currently under way to "cure" various genetic diseases by replacing the defective or missing gene with a normal gene from some other source. Such attempts, called *gene therapy,* require *genetic engineering* or *recombinant DNA technology* (see Appendix C).

Scientists have identified a large number of genetic diseases and characterized their specific enzymatic defects. Table 8.3 lists a number of human genetic diseases and their known defects.

8.3. EXPERIMENTAL APPROACHES TO METABOLISM

8.3.1. Levels of Cellular Organization

We can investigate metabolism at different levels of organization, from the intact organism (*in vivo,* i.e., within a living organism) to the isolated molecule (*in vitro,* i.e., outside a living organism). Biochemists face an unavoidable dilemma: the more closely a system reflects its natural state, the more complex its level of organization and the less the amount of specific information we can derive from it. Thus, an entire animal or plant constitutes a fully organized biological system but, unfortunately, lends itself to only a few, relatively simple studies. Conversely, a purified enzyme or other protein can be studied in great detail. However, in the process of isolating and purifying the enzyme or protein, organization at all levels had to be destroyed.

In one type of whole-organism study, we determine the effects on metabolism produced by surgically removing an endocrine gland. Another experiment with whole organisms consists of performing a **balance study.** Here we measure the difference between intake of a substance and output of the same or a different substance. *Nitrogen balance,* for example, represents the difference between the amount of nitrogen ingested and the amount of nitrogen excreted. *Respiratory quotient* represents the ratio of moles of carbon dioxide produced to moles of oxygen consumed.

Physicians also employ measurements performed on intact organisms when they diagnose and treat medical problems. Some typical examples are construction of a *glucose tolerance curve* (see Section 10.1); determinations of cholesterol, triglycerides, and urea in blood (BUN, blood urea nitrogen); and determinations of urea, uric acid, and phenylpyruvate in urine.

Bacteria are particularly useful for studies requiring an intact organism. We can easily grow bacteria in large numbers because they are small, require relatively simple growth media, and multiply rapidly. Thus, *Escherichia coli (E. coli),* commonly found in mammalian intestines, has a generation time of about 30 minutes in the laboratory so that the number of cells doubles every 30 minutes. We can grow *E. coli* cells on a defined medium containing only glucose, citrate, K_2HPO_4, $MgSO_4$, and $NaHNH_4PO_4$. Using a growth medium composed of only a few components is advantageous for metabolic studies because it simplifies the task of isolating and determining specific metabolites. Because of its short generation time, *E. coli* lends itself well to genetic studies; researchers can investigate a large number of generations in a limited time.

An organizational level less complex than an intact organism consists of an isolated and perfused organ. When we *perfuse* an organ, we pass blood, plasma, or other fluids through the organ's blood vessels. With this technique, we can measure the metabolism of a particular substance by the isolated organ. For example, passing a known amount of glucose through a perfused liver and analyzing the effluent allows us to determine the rate of glucose oxidation by the liver. Surgeons use perfusion to keep organs "alive" during organ transplantation.

A level of organization below that of a perfused organ comprises several types of tissue preparations. These include thin sections of tissue (*tissue slices*), specific types of cells maintained in an artificial medium (*tissue culture*), and tissue disrupted in a blender (*tissue homogenate*). Researchers have used tissue slices for many years to study aerobic metabolism by determining oxygen uptake and carbon dioxide evolution from manometric measurements.

Studies at the lowest level of cellular organization involve those of a purified enzyme or other protein. Such studies can provide a wealth of information about the isolated molecule. Keep in mind, though, that the properties determined may differ from those the molecule has inside the cell, where it may be located in a specific compartment, be adsorbed to a particular subcellular structure, or be subject to unique control mechanisms.

8.3.2. Radioactive Labels

Studies at all levels of cellular organization can be enhanced by using radioactive isotopes (see Appendix C). Unlabeled and radioactively labeled compounds have identical chemical properties and undergo the same reactions in metabolism. Using compounds that carry radioactive labels provides ways of tracing their metabolic

fates in a cell or in an organism. We can illustrate this by asking the following question: Can the carbons of glucose, administered to a rat, be converted to carbons of exhaled CO_2?

$$\text{Glucose} \xrightarrow[\text{(Rat)}]{??} CO_2$$

Without using isotopes, you would find it very difficult to investigate this problem. However, by means of isotopes you can readily arrive at an answer to the question. To do so, you would administer some ^{14}C-labeled glucose to a rat, collect the exhaled CO_2 as a function of time, and analyze the gas for radioactivity. If the exhaled CO_2 becomes radioactive during the experiment, you can conclude that at least some of the carbons of glucose can be converted to those of CO_2. Glucose, in other words, can serve as a metabolic **precursor** of CO_2 (a precursor is a substance that precedes another in a metabolic pathway). Whether the conversion of glucose to CO_2 occurs by means of a single step or via a long series of reactions cannot be answered by your experiment. Note that even though glucose undergoes numerous reactions in metabolism, the use of radioisotopes allows you to study one particular reaction (or set of reactions) in a very complex system.

The use of isotopes has been invaluable in elucidating metabolic pathways. Continuing with the above example, suppose you have reason to suspect that *pyruvate* serves as an intermediate in the conversion of glucose to CO_2. To investigate your hypothesis, you repeat the experiment, administer radioactive glucose to a rat, isolate pyruvate, and determine whether or not it contains any ^{14}C. If the isolated pyruvate is radioactive, you conclude that it, too, must serve as a precursor of CO_2 so that you can expand the pathway to read:

$$\text{Glucose} \longrightarrow \text{pyruvate} \longrightarrow CO_2$$

To illustrate another use of isotopes, assume that you wanted to determine where proteins are being synthesized in the bacterial cell. To find the answer, you could add radioactively labeled amino acids to a growing culture of cells and allow the cells to incorporate the label for a very short time. You would then stop protein synthesis, harvest the cells, lyse them, collect various subcellular fractions (mitochondria, cell walls, etc.), and determine their radioactivity. You would find that the ribosome fraction has the greatest concentration of radioactive amino acids. You conclude that ribosomes very likely represent the sites of intracellular protein synthesis.

8.3.3. Mutants

Mutants provide another useful tool for studying metabolism. Bacterial mutants have similar genetic defects to the defects found in humans. However, whereas human genetic diseases are difficult to study because they occur at low frequencies, bacterial mutants can be produced in the laboratory at high frequencies and studied with ease. One way to produce mutants consists of exposing bacteria to mutagenic agents, such as X rays or specific chemicals. The resultant mutations in DNA lead to changes in bacterial metabolism.

Biochemists find *nutritional (auxotrophic) mutants* particularly useful for metabolic studies. An auxotroph is a microorganism that has a block in a metabolic pathway due to lack of a specific enzyme or presence of a defective enzyme. Such mutants require for their growth either the product of the enzymatic reaction or other metabolites not required by the normal, *wild-type* organism.

As an example, suppose you know that a given microorganism can synthesize the amino acid lysine when provided with compound A. There must, therefore, exist a pathway in this organism such that

$$\text{A} \ldots\ldots\ldots\ldots\ldots\ldots \longrightarrow \text{Lys}$$

Assume further that, when you expose the wild-type organism to mutagenic agents, several mutants (I–III) form. When you isolate the mutants and determine their growth requirements, you find that they will not grow on the medium of the wild type unless you supplement that medium with additional nutrients.

You discover that you can get mutant I to grow only by addition of lysine. When you add lysine, the organism grows and produces an excess of compound D. You may conclude that D serves as an intermediate in the pathway of lysine biosynthesis. Moreover, because it is likely that only one enzyme has been damaged in mutant I, you deduce that the conversion of D to lysine probably represents the last step in the pathway and that mutant I has a block (X) at this step.

$$\text{Mutant I} \quad \text{A} \ldots\ldots\ldots\ldots \xrightarrow{X} \text{D} \longrightarrow \text{Lys}$$

You interpret the results obtained with the other mutants in a similar fashion. Mutant II grows when you add *either* lysine or D and produces an excess of compound C so that

$$\text{Mutant II} \quad \text{A} \ldots\ldots\ldots\ldots \text{C} \xrightarrow{X} \text{D} \longrightarrow \text{Lys}$$

Mutant III grows when you add lysine, D, or C and produces an excess of compound B. It follows that

$$\text{Mutant III} \quad A \xrightarrow{X} B \longrightarrow C \longrightarrow D \longrightarrow \text{Lys}$$

You can see how you begin to elucidate the pathway from A to lysine. At this point, you must still determine whether the conversion of A to B proceeds in a single step or via a multiple-step sequence.

8.4. NUTRITIONAL ASPECTS

8.4.1. Digestion and Absorption

Digestion represents the first stage in metabolism in humans and other animals. It comprises the hydrolysis of food macromolecules to smaller molecules in the digestive tract, followed by their absorption across the intestinal membrane into the bloodstream or lymph and their transport to the tissues. We refer to the agents of digestion as *digestive fluids* and distinguish five such fluids (Table 8.4).

Hydrolysis during digestion is enzymatic in nature, except for the acid hydrolysis of proteins in the stomach. Strictly speaking, *bile* is not a *digestive* fluid since it does not contain enzymes. However, it does play an important role in lipid digestion. Bile empties into the intestine and acts as an *emulsifying agent,* hence its inclusion with the digestive fluids.

Saliva contains an *amylase* that catalyzes starch hydrolysis. Because of food's short exposure to the enzyme, only partial degradation of starch to oligosaccharides takes place in the mouth. *Gastric juice* has a pH of about 1–2 (it represents a 0.10–0.15*M* HCl solution), low enough to result in some hydrolysis of peptide bonds. Additionally, stomach secretions contain *pepsin,* an unusual enzyme that not only functions at the low pH of the stomach but actually has its optimum pH value in that range. Owing to the combined action of HCl and pepsin, a significant amount of protein digestion takes place in the stomach. Gastric juice also contains an acid-stable *lipase* that results in a small amount of lipid digestion.

Table 8.4. Digestion of Carbohydrates, Lipids, and Proteins

	Compounds digested		
Digestive fluid	Carbohydrates	Lipids	Proteins
Saliva	Some	None	None
Gastric juice	None	Some	Yes
Pancreatic fluid	Yes	Yes	Yes
Intestinal fluid	Yes	Yes	Yes
Bile	None	Yes	None

Most of the digestion of the three major nutrients—carbohydrates, lipids, and proteins—occurs through the action of the *pancreatic* and *intestinal fluids.* Nucleic acids as well are digested primarily by action of these digestive fluids. Both fluids contain many hydrolytic enzymes that catalyze the degradation of oligo- and polysaccharides, oligo- and polypeptides, lipids, and oligo- and polynucleotides. Pancreatic fluid contains large amounts of protease zymogens that undergo conversion to active enzymes in the intestine.

The combined action of the digestive fluids and bile results in degradation of nutrient macromolecules to small building blocks—proteins to amino acids and di- and tripeptides; carbohydrates to mono- and oligosaccharides; and lipids to fatty acids, monoacylglycerols, cholesterol, and other lipid components.

Protein digestion also destroys the antigenic character of foreign proteins taken in through the diet. If foreign proteins were not degraded in this fashion and were absorbed intact into the bloodstream, humans would experience a powerful, and possibly fatal, immunological reaction called *anaphylactic shock.* When such a reaction occurs, the immune system overreacts, and antigen–antibody complexes trigger the release of harmful substances, notably histamine, by a class of body cells termed mast cells. The amino acids produced from dietary proteins are, of course, indistinguishable from those found in humans and do not trigger any immunological response.

Following the digestive process, nutrients are absorbed across the membranes of the epithelial cells of the intestine. Free amino acids are absorbed via a carrier-mediated transport. Small peptides are taken up by means of specific transport systems and are generally hydrolyzed by cytoplasmic enzymes of these cells to amino acids. The portal vein then transports all of the amino acids to the liver (Figure 8.9).

Oligosaccharides are hydrolyzed to monosaccharides by surface enzymes of the epithelial cells. All of the monosaccharides are absorbed by means of carrier-mediated transport and moved via the portal vein to the liver.

Bile salts solubilize lipid components by forming micelle-like structures that diffuse into the epithelial cells. This absorption is virtually complete for free fatty acids and monoacylglycerols but is less efficient for other lipids. Only 30–40% of dietary cholesterol, for example, is absorbed.

The fate of the absorbed fatty acids inside the cells depends on the chain lengths of the fatty acids. Fatty acids of medium chain length (C_6–C_{10}) pass through the cells into the portal blood and are carried directly to the liver.

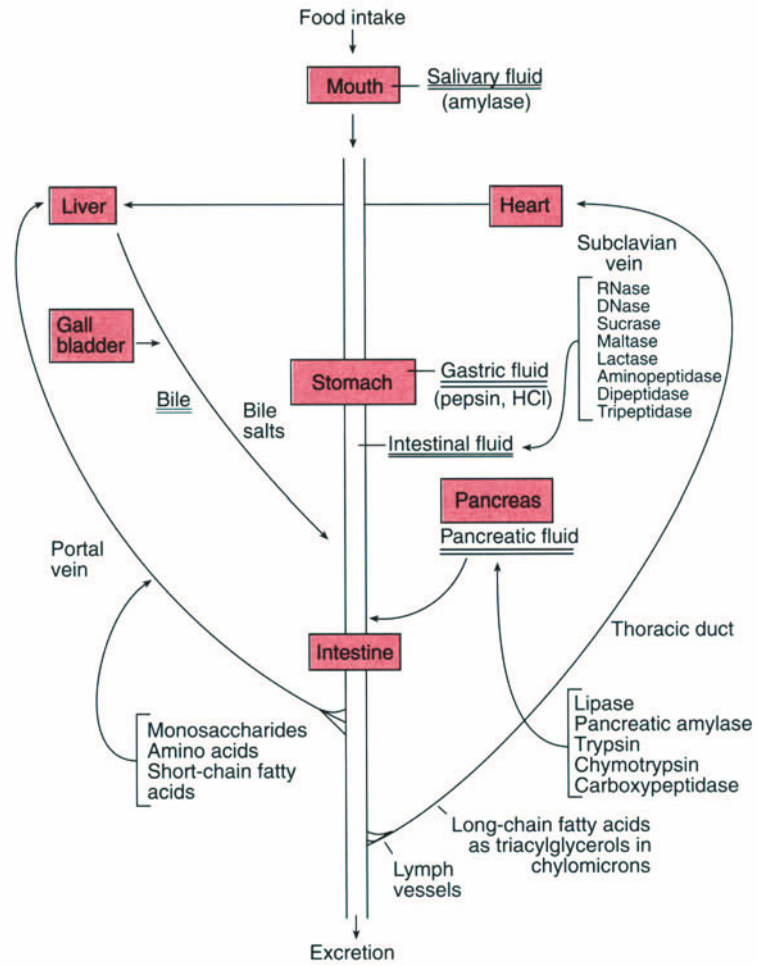


Figure 8.9. Some components of the digestive system in animals.

Long-chain fatty acids ($>C_{12}$) are resynthesized into triacylglcyerols that form parts of *chylomicrons,* which leave the intestine via the lymphatic system and bypass the liver. We refer to the suspension of chylomicrons in the lymphatic system present after a meal as *chyle.* Intestinal lymph vessels drain into large body veins via the thoracic duct. Blood in large veins first reaches peripheral tissues, including adipose and muscle tissue, before coming in contact with the liver.

8.4.2. Dietary Nutrients

An organism's nutritional requirement usually reflects its biosynthetic capabilities. Many prokaryotes require only carbon, energy, and a few simple inorganic salts for growth. As you read earlier, *E. coli* can grow when provided with a few simple inorganic salts and glucose as a carbon and energy source. From these nutrients, the bacterium can synthesize all 20 amino acids and many other compounds. Animals, on the other hand, require a larger number of nutrients and have a more limited capacity for amino acid synthesis. Humans require some 50 different nutrients (Table 8.5) and can only synthesize 10 amino acids (see Table 2.4).

We classify nutrients required by humans as *macronutrients,* required in gram quantities per day, and *micronutrients,* needed only in milligram or microgram amounts. Carbohydrates, lipids, and proteins constitute macronutrients. Vitamins and minerals represent micronutrients. Minerals comprise *macro-* and *microminerals.* We require macrominerals (calcium, chlorine, magnesium, phosphorus, potassium, sodium, and sulfur) in the diet in relatively large amounts. Microminerals, or *trace elements,* are required only in small amounts.

Table 8.6 shows the *recommended dietary allowances (RDAs)* of some minerals and vitamins. These daily allowances are not minimal values needed to avoid clinical symptoms of deficiency but rather amounts that provide an

ample safety margin. RDAs are established by the *Food and Nutrition Board* of the *National Research Council* (an agency of the *National Academy of Sciences*) for a normal individual engaged in average activity and living in a temperate climate. RDAs represent general guidelines that vary with age, sex, physiological condition, and activity.

For some vitamins and minerals, less information exists on which to base dietary allowances. Instead, recommended amounts are quoted in terms of ranges of *estimated safe and adequate daily dietary intakes (ESADDIs).* Table 8.7 lists a number of examples.

Two other general nutritional recommendations have emerged in recent times: to decrease fat intake, especially intake of saturated fat, and to include *dietary fiber*—components of food that cannot be degraded by human digestive enzymes. Dietary fiber consists primarily of cellulose and other nondigestible cell-wall polymers of plants. Cellulose and hemicellulose enhance intestinal motility, increase stool bulk, and decrease the stool's transit time through the colon. Researchers believe that dietary fiber may reduce the risk of colon cancer because the fiber lowers the concentration of a potential carcinogen

Table 8.5. Nutrients Requied by Humans

Energy sources
- Carbohydrates[a]
- Fats
- Proteins

Essential amino acids
- Arginine (in young)
- Histidine
- Isoleucine
- Leucine
- Lysine
- Methionine
- Phenylalanine
- Threonine
- Tryptophan
- Valine

Essential fatty acids
- Linoleic acid
- Linolenic acid

Vitamins (fat-soluble)
- Vitamin A
- Vitamin D
- Vitamin E
- Vitamin K

Vitamins (water-soluble)
- Thiamine (vitamin B_1)
- Riboflavin (vitamin B_2)
- Niacin
- Pyridoxine (vitamin B_6)
- Pantothenic acid
- Folic acid
- Biotin
- Vitamin B_{12}
- Ascorbic acid

Water

Minerals
- Arsenic[b]
- Calcium
- Chlorine
- Chromium
- Copper
- Fluorine
- Iodine
- Iron
- Magnesium
- Manganese
- Molybdenum
- Nickel[b]
- Phosphorus
- Potassium
- Selenium
- Silicon
- Sodium
- Tin[b]
- Vanadium[b]
- Zinc

[a]Recommended carbohydrate intake is 100 g/day.
[b]Believed to be required by humans.

Table 8.6. Recommended Dietary Allowances (RDAs) for the Average Human between Ages 19 and 24[a]

Nutrient	RDA Male	RDA Female
Protein	58 g	46 g
Fat-soluble vitamins		
Vitamin A	1000 μg RE[b]	800 μg RE[b]
Vitamin D	10 μg[c]	10 μg[c]
Vitamin E	10 mg α-TE[d]	8 mg α-TE[d]
Vitamin K	70 μg	60 μg
Water-soluble vitamins		
Vitamin C	60 mg	60 mg
Thiamine (vitamin B_1)	1.5 mg	1.1 mg
Riboflavin (vitamin B_2)	1.7 mg	1.3 mg
Vitamin B_6	2.0 mg	1.6 mg
Vitamin B_{12}	2.0 μg	2.0 μg
Niacin	19 mg NE[e]	15 mg NE[e]
Folate	200 μg	180 μg
Minerals		
Calcium	1200 mg	1200 mg
Phosphorus	1200 mg	1200 mg
Magnesium	350 mg	280 mg
Zinc	15 mg	12 mg
Iron	10 mg	15 mg
Iodine	150 μg	150 μg
Selenium	70 μg	55 μg
Energy[f]	12,134 kJ (2,900 kcal)	9,205 kJ (2,200 kcal)

[a]Source: National Research Council (1989).
[b]RE, Retinol equivalent (1 RE = 1 μg of retinol or 6 μg of β-carotene).
[c]As cholecalciferol (10 μg of cholecalciferol = 400 IU of vitamin D).
[d]α-TE, α-Tocopherol equivalent (1 α-TE = 1 mg of *d*-α-tocopherol).
[e]NE, Niacin equivalent (1 NE = 1 mg of niacin or 60 mg of dietary tryptophan).
[f]These values may be compared with those required to maintain the *basal metabolic rate (BMR)*, the rate of energy utilization in the resting state. Energy requirements for BMR are about 7531 kJ (1800 kcal) for males and 5439 kJ (1300 kcal) for females.

Table 8.7. Estimated Safe and Adequate Daily Dietary Intakes (ESADDIs) of Selected Vitamins and Minerals[a]

Nutrient	ESADDI Children (1–3 yr)	ESADDI Adolescents (11+ yr)	ESADDI Adults
Vitamins			
Biotin	20 μg	30–100 μg	30–100 μg
Pantothenic acid	3 mg	4–7 mg	4–7 mg
Minerals			
Copper	0.7–1.0 mg	1.5–2.5 mg	1.5–3.0 mg
Manganese	1.0–1.5 mg	2.0–5.0 mg	2.0–5.0 mg
Fluoride	0.5–1.5 μg	1.5–2.5 μg	1.5–4.0 μg
Chromium	20–80 μg	50–200 μg	50–200 μg
Molybdenum	25–50 mg	75–250 mg	75–250 mg

[a]Source: National Research Council (1989).

(by increasing stool bulk) and accelerates its passage through the colon. Some water-soluble types of fiber, such as pectin and gums, form viscous solutions in the stomach and the intestine. This decreases the rate of carbohydrate digestion and absorption. Consequently, the level of blood sugar rises more slowly after a meal. Soluble fibers also help to lower the level of serum cholesterol in most people. Researchers have proposed that this effect results from increased fecal excretion of bile acids that are major degradation products of cholesterol.

8.4.3. Vitamins

Vitamins are integral components of the diet of humans and other animals, and their importance for general health and growth is well known. We define a **vitamin** as an organic compound that (a) occurs in natural foods in extremely small concentrations and is distinct from carbohydrates, lipids, proteins, and nucleic acids; (b) is required by the organism (generally restricted to animals) in minute amounts for normal health and growth; (c) when absent from the diet, or improperly absorbed from the intestine, leads to development of a specific deficiency disease; and (d) cannot be synthesized by the organism or is not synthesized in sufficient amount and must, therefore, be obtained through the diet. The term *vitamin* was proposed by Casimir Funk in 1912 and was originally spelled "vitamine" to designate a "vital amine" from rice husks that cured the disease beriberi. We now know that this "vital amine" is thiamine or vitamin B_1. Funk thought that all vitamins were amines, but other scientists showed later that that was incorrect, and the "e" was omitted from the name.

Because all vitamins occur naturally, there clearly exist organisms capable of synthesizing each vitamin. However, a compound that is a vitamin for one organism may not be a vitamin for another organism. We classify vi-

Table 8.8. Water-Soluble Vitamins and Their Coenzymes

Coenzyme	Reaction mediated	Vitamin precursor	Deficiency disease	Section
L-Ascorbic acid	Hydroxylation	L-Ascorbic acid (vitamin C)	Scurvy	5.1
Biotin attached to ϵ-NH_2 of Lys (biocytin)	Carboxylation	Biotin (vitamin B complex)	Dermatitis (humans)	13.5
Cobamide coenzymes	Alkylation	Cobalamin (vitamin B_{12})	Pernicious anemia	13.3
Coenzyme A	Acyl transfer	Pantothenic acid (vitamin B complex)	Dermatitis (chickens)	11.1
Flavin coenzymes	Oxidation–reduction	Riboflavin (vitamin B_2)	Growth retardation	11.1
Folate coenzymes	One-carbon group transfer	Folic acid (vitamin B complex)	Macrocytic anemia	14.4
Lipoic acid[a] attached to ϵ-NH_2 of Lys (lipoamide, lipoyllysine)	Acyl transfer	Lipoic acid[a] (vitamin B complex)	Growth deficiencies	11.1
Nicotinamide coenzymes	Oxidation–reduction	Niacin (nicotinic acid; vitamin B complex)	Pellagra	11.1
Pyridoxal phosphate	Amino group transfer (transamination)	Pyridoxine, pyridoxal, and pyridoxamine (vitamin B_6)	Dermatitis (rats); neurological symptoms	14.2
Thiamine pyrophosphate	Aldehyde transfer and decarboxylation	Thiamine (vitamin B_1)	Beriberi	11.1

[a]Lipoic acid is a vitamin-like compound but not a vitamin.

tamins into two broad classes, *fat-soluble* and *water-soluble vitamins*

Fat-soluble vitamins are soluble in organic solvents and insoluble in water. They are structurally related but have diverse functions. The group includes vitamins A, D, E, and K and was discussed in Chapter 6.

Water-soluble vitamins include the B vitamins, or *vitamin B complex,* as well as some others, such as vitamin C. They have diverse structures but show functional similarity; all participate in transfer reactions involving protons, electrons, acyl groups, *one-carbon fragments* (HCHO, $-CH_2OH$, $-CH_3$, and the like), or other metabolic groups.

As you read in Section 4.4, water-soluble vitamins form parts of **coenzymes.** Some vitamins represent small parts of coenzymes; others constitute major portions. In one case, ascorbic acid, the vitamin and the coenzyme are one and the same. Sometimes the vitamin functions as the "business end" of the coenzyme, site of the chemical reaction. In other instances, the vitamin serves as a structural component but does not participate directly in the reaction. Whatever the particulars, each water-soluble vitamin is required for the biosynthesis of a coenzyme; each serves as a coenzyme precursor. Table 8.8 lists coenzymes and the types of reactions they mediate. We will examine the structures and functions of these coenzymes as we encounter them in later chapters.

Because of the high efficiency of enzymes, most intracellular enzyme and coenzyme concentrations are low. Accordingly, vitamins are required in only small amounts in the diet. For a healthy individual living under suitable conditions, a well-balanced diet probably provides all the necessary types and amounts of vitamins required. However, people suffering from specific ailments or living under special conditions may benefit from supplemental vitamin intake.

Usually, additional moderate intake of water-soluble vitamins poses no harm because they are readily excreted. Fat-soluble vitamins, on the other hand, tend to sequester in lipid-containing structures like membranes and fat cells and may accumulate to toxic levels. Therefore, present-day "megavitamin" regimens must be carried out with caution.

SUMMARY

Metabolism comprises a network of interconnected reactions that consists of two parts: catabolism, or degradative processes, and anabolism, or synthetic processes. Degradation and synthesis of intermediates proceeds via specific metabolic pathways. Because an exergonic reaction is endergonic in the opposite direction, a "downhill" catabolic pathway cannot be simply reversed. Instead, changes must occur at one or more steps of the sequence to circumvent these energy barriers. All metabolic pathways are regulated, and each has a committed step that ensures completion of the sequence of reactions.

Metabolism is regulated by control of individual enzymes; by control of metabolic pathways; by compartmentation of substrates, enzymes, and pathways; and by hormone action. Hormones bind to receptors on cell membranes or on cytoplasmic or nuclear proteins. Some hormones cause release of a second messenger inside the cell.

We can study metabolism at the level of whole organisms, perfused organs, tissue preparations, or purified proteins. The more closely a system reflects its natural state, the more complex its level of organization and the less the amount of specific information that we can derive from it. Biochemists use radioactively labeled compounds and bacterial mutants to elucidate metabolic pathways.

Five digestive fluids accomplish the digestion of macromolecules to smaller building blocks. Most digestion of the major fooodstuffs occurs via the pancreatic and intestinal fluids.

Humans require both macro- and micronutrients, including vitamins, in the diet. Water-soluble vitamins have functional similarity but diversity of structure. All form parts of coenzymes that function in group transfer reactions.

SELECTED READINGS

Beishir, L., *Microbiology in Practice,* 6th ed., HarperCollins, New York (1996).

Bender, D., *Nutritional Biochemistry of the Vitamins,* Cambridge University Press, Cambridge (1992).

Cowan, J. A., *Inorganic Biochemistry—An Introduction,* VCH Publishers, New York (1993).

Ford, T. C., and Graham, J. M., *An Introduction to Centrifugation,* Bios Scientific Publishers, Oxford (1991).

Freshney, R. I., *Culture of Animal Cells: A Manual of Basic Technique,* 3rd ed., Wiley, New York (1993).

Liscum, L., and Underwood, K. W., Intracellular cholesterol transport and compartmentation, *J. Biol. Chem.* 270:15443–15446 (1995).

Matthews, J. C., *Fundamentals of Receptor, Enzyme, and Transport Kinetics,* CRC Press, Boca Raton, Florida (1993).

Norman, A. W., and Litwack, G., *Hormones,* Academic Press, Orlando (1987).

Russell, J. B., and Cook, G. M., Energetics of bacterial growth: Balance of anabolic and catabolic reactions, *Microbiol. Rev.* 59:48–62 (1995).

Scriver, C. R. *The Metabolic and Molecular Bases of Inherited Disease,* 7th ed., McGraw-Hill, New York (1995).

Weindruch, R., Caloric restriction and aging, *Sci. Am.* 274:46–52 (1996).

REVIEW QUESTIONS

A. Define each of the following terms:

Marker enzyme	Digestion
Mitochondria	Balance study
Committed step	Precursor
Compartmentation	Fermentation
Receptor	Genetic (hereditary) disease

B. Differentiate between the two terms in each of the following pairs:

Catabolism/anabolism	Digestion/fermentation
First messenger/second messenger	Metabolism/intermediary metabolism
Alcoholic fermentation/lactate fermentation	Metabolic block/metabolic pathway

C. (1) Describe the levels at which metabolism can be studied. How does the organization of the system being studied relate to the amount of information that you can derive from it? Why is this so?
(2) What are the three general characteristics of metabolic pathways?
(3) What are the three stages of catabolism?
(4) Outline the main features of the digestion of carbohydrates, lipids, and proteins, beginning with the intake of food through the mouth and ending with absorption of components across the membranes of intestinal cells.
(5) How is the hypothalamus–pituitary hormonal control network organized?

PROBLEMS

Why is it advantageous to control a metabolic pathway by means of its committed step?

8.2. The 12-hour human gastric secretion of HCl typically varies from about 150 to 1000 ml. Calculate the number of protons in these two volumes of output on the assumption that the secretion is 0.150*M* in HCl.

8.3. Assume that you have disrupted mitochondria and obtained three fractions designated A, B, and C. You then assay the three fractions for a number of different enzymes with the following results (for each enzyme the total activity in all three fractions is 100%):

	Percent of total activity		
	A	B	C
Monoamine oxidase	5	82	13
Cytochrome oxidase	10	11	79
Fumarase	85	7	8

On this basis, identify fractions A, B, and C as corresponding primarily to matrix, inner membrane, or outer membrane preparations.

8.4.* You can see that each mitochondrial preparation in the previous problem is contaminated by material from the other two fractions. Assuming that a unit of enzymatic activity reflects the same amount (in milligrams) of subcellular fraction for all three enzymes, estimate the approximate contamination of each fraction by the other two.

8.5.* A liver is perfused with a $1.50 \times 10^{-3}M$ solution of glucose. When the flow rate is adjusted to 0.500 ml/min, the exiting solution has a glucose concentration of $1.30 \times 10^{-3}M$. What is the rate of glucose metabolism by the liver in terms of milligrams of glucose per hour? The molecular weight of glucose is 180.

8.6. You incubate aliquots of a cell-free extract with a radioactively labeled compound (X) for various

lengths of time. Following the incubation, you fractionate the aliquots and identify the labeled compounds. You obtain the following results:

Incubation time (min)	Labeled compounds found
5	K
10	K, G
25	K, G, T

Deduce the pathway whereby compound X is metabolized in this cell-free extract.

8.7. In the illustration on the use of radioactively labeled amino acids to identify the sites of protein synthesis (Section 8.3), the bacterial culture was exposed to the label for only a *short time.* Why does this need to be stressed? What would happen if the organisms are allowed to grow the entire time in the presence of labeled amino acids? (Hint: Once a polypeptide chain has been completed, it is released from the ribosome.)

8.8. A wild-type bacterium can synthesize compound X when compound Y is present in the growth medium. An auxotrophic mutant grows only if X is present in the medium, and during its growth compound Q accumulates. A second auxotrophic mutant can grow when *either* X or Q is present in the medium, and during its growth compound S accumulates. Sketch the metabolic pathway that these data indicate.

8.9.* A solution of a ^{14}C-labeled compound that is not metabolized contains 50.0 mmol/ml and has a specific activity of 2.40×10^3 cpm/μmol (cpm = counts per minute; see Appendix C). A patient receives an intravenous injection of 2.00 ml of this solution. After a short time, the radioactivity is evenly distributed throughout the blood. At that point, 5.00 ml of blood is collected from the patient. The protein in the blood is removed by precipitation (the precipitate is not radioactively labeled), and the supernatant contains a total of 4.00×10^5 cpm. What is the patient's total blood volume?

8.10. In a laboratory experiment, two rats are injected with labeled compounds that are not metabolized and that are excreted at identical rates. One rat receives 1.80 millicuries (see Appendix C) of a ^{14}C-labeled compound, and the other receives 20.0 millicuries of a ^{35}S-labeled compound. The half-life (see Appendix C) of ^{14}C is 5568 years, and that of ^{35}S is 87.1 days. Which animal will have the larger number of millicuries left in its system after 3, 6, and 9 months?

8.11.* What is the half-life of an isotope if 50 atoms out of 1000 disintegrate in one year? (See Appendix C.)

8.12. Damage to the pituitary gland results in a condition known as *diabetes insipidus* and is characterized by a massive flow of dilute urine and an unquenchable thirst. Why is this so?

8.13. Based on the data in Table 8.6, how many calories above those needed to maintain the basal metabolic rate do men and women aged 19–24 need each day?

Bioenergetics

9

Broadly speaking, metabolism deals with transformations of substances and transformations of energy. How usable energy is derived from nutrients and how it is used to drive metabolic processes are questions fundamental to understanding the workings of metabolism.

Historically, little progress was made in biochemistry until a key principle of biochemical energetics had been formulated. Antoine Lavoisier recognized in 1777 that cellular respiration was slower than combustion but not essentially different from it. Both processes accomplish the oxidation of foodstuffs to carbon dioxide and water. Lavoisier's observation led to the realization that biochemical reactions have the same characteristics as ordinary chemical reactions performed in the laboratory.

Specifically, investigators concluded that life plays by the rules of thermodynamics. A reaction in a living system has the same free energy change associated with it as the corresponding *in vitro* reaction. We now understand that this must be so because free energy, like other thermodynamic functions, *depends only on the final and initial states* of the system and not on the mechanism, or pathway, whereby the system proceeds from the initial to the final state. Hence, the conversion of glucose to CO_2 and water must have the same free energy change associated with it whether it is carried out by combustion in a calorimeter or by enzymatic reactions in a living cell.

9.1. FREE ENERGY

Free energy (**G** for **Gibbs free energy**) is the key thermodynamic function in biochemical systems. It allows us to deduce the direction that a spontaneous reaction will take and indicates the energetic yield or "cost" of a process. A reaction with a negative free energy change (an *exergonic* reaction) proceeds spontaneously (Figure 4.4); a reaction with a positive free energy change (an *endergonic* reaction) does not proceed spontaneously. An endergonic reaction is thermodynamically not feasible unless energy is applied to it.

We cannot determine the actual free energy of any process or substance; we can only determine the *change* in free energy (ΔG). The symbol delta (Δ) always represents the difference between *final* and *initial* states. Thus, ΔG of a chemical reaction represents the difference between the free energies of the *products* and the *reactants:*

$$\Delta G_{\text{reaction}} = G_{\text{products}} - G_{\text{reactants}}$$

The free energy change is related to two other thermodynamic functions—enthalpy and entropy—at constant temperature and pressure:

$$\Delta G = \Delta H - T\Delta S \qquad (9.1)$$

where ΔH is the change in **enthalpy** (or **heat content**) of the reaction, T is the absolute temperature in degrees Kelvin (K = 273.2 +°C), and ΔS is the change in **entropy,** a measure of randomness. You can infer from Eq. (9.1) that free energy constitutes the difference between the total heat content of a reaction and the energy lost by randomnization ($T\Delta S$); it represents the fraction of the total energy of a system that can be used to do work (hence, the term *free* energy).

The units of free energy are joules (J) per mole or kilojoules (kJ) per mole [calories (cal) per mole or kilocalories (kcal) per mole; 1 cal = 4.184 J]. As these units indicate, the free energy yield of a reaction depends on the *number of moles* of reactants that undergo reaction.

9.1.1. Standard and Actual Free Energy Changes

9.1.1A. Standard Free Energy Change.

We base all thermodynamic functions on arbitrarily selected reference states, called *standard states,* that represent particular sets of conditions. For solutions, we use as standard state a solute concentration of 1.0M, a temperature of 25°C, and a pressure of 1.0 atm. We call the free energy change that corresponds to this state the **standard free energy change** and designate it $\Delta G°$.

You can calculate the standard free energy change of a reaction from its equilibrium constant (K'_{eq}) by means of the equation

$$\Delta G° = -RT \ln K'_{eq} \qquad (9.2)$$

where ln is the natural logarithm (ln = 2.303 $\log_{10}$), T is the absolute temperature, and R is the gas constant (8.314 J deg^{-1} mol^{-1}; 1.987 cal deg^{-1} mol^{-1}). Table 9.1 illustrates the quantitative relationship between the standard free energy change and the equilibrium constant.

You can see from Table 9.1 that the larger the equilibrium constant is (product concentrations greatly exceed reactant concentrations at equilibrium), the more negative the free energy change. Remember, however, that the *magnitude of the free energy* change tells you nothing about the *rate of the reaction.* A reaction with a highly negative free energy change does not necessarily proceed rapidly; the reaction rate depends on the magnitude of the *energy of activation* (Figure 4.4).

Table 9.1. Relationship between the Equilibrium Constant (K'_{eq}) and the Standard Free Energy Change ($\Delta G°$) at 25°C[a]

K'_{eq}	log K'_{eq}	$\Delta G°$ (J mol^{-1})
0.001	−3	+17,117
0.01	−2	+11,410
0.1	−1	+5,707
1	0	0
10	1	−5,707
100	2	−11,410
1000	3	−17,117

[a]The same relationship holds between the biochemical equilibrium constant (K'_{bio}) and the biochemical standard free energy change ($\Delta G°'$), both of which are discussed in Section 9.1.2.

We can define $\Delta G°$ in two ways. One represents a *mathematical definition* and consists of Eq.(9.2). Given a reaction for which you know the equilibrium constant, you can compute the *value* of $\Delta G°$ from Eq.(9.2).

But what exactly does this calculated value of $\Delta G°$ mean in physical terms? To what specific reaction conditions does it refer? We might call the answer to these questions a *conceptual definition* of $\Delta G°$ and can phrase it as follows:

> The standard free energy change of a reaction is the free energy change associated with the reaction when all reactants and all products are at an initial concentration of 1.0M each, the temperature is 25°C, the pressure is 1.0 atm, and the reaction is allowed to proceed to equilibrium.

It is immaterial how, or whether, this set of conditions can be attained. A 1.0M concentration, for example, might exceed the solubility of a given compound. These specifications serve merely as a *reference for defining* the free energy change. *If* we mix reactants and products under these conditions, *then* the free energy change associated with the reaction has the value denoted as $\Delta G°$.

We must distinguish clearly between the two ways of defining $\Delta G°$. We base our mathematical evaluation of $\Delta G°$ on the equilibrium constant of the reaction. By contrast, the physical meaning of $\Delta G°$ does not pertain to a state of equilibrium but rather to the reaction proceeding under specified conditions. Let us illustrate these concepts by considering the ionization of acetic acid:

$$CH_3COOH \rightleftarrows CH_3COO^- + H^+$$

This reaction has an equilibrium constant of 1.76×10^{-5} at 25°C. To calculate the *numerical value* of $\Delta G°$, we use Eq.(9.2):

$$\Delta G^\circ = -(8.31)\,(298.2)\,(2.303)\log(1.76 \times 10^{-5})$$
$$= 27,150 \text{ J mol}^{-1}$$
$$= 27.2 \text{ kJ mol}^{-1}$$

Now ask yourself what this calculated quantity *means* in physical terms. The answer is that if you were to mix CH_3COOH, CH_3COO^-, and H^+ in such a fashion that each species had an initial concentration of 1.0*M* in the reaction mixture, the temperature was kept at 25°C, and the pressure was kept at 1.0 atm, then the tendency of the reaction to proceed to equilibrium *under these conditions* is described by ΔG°. Because ΔG° is very positive, the reaction is strongly endergonic and will not proceed as written. (In fact, the reaction will proceed spontaneously in the opposite direction, from right to left.) At the low pH of zero, acetic acid will not ionize; the high proton concentration ($[H^+] = 1.0M$) depresses the ionization of the acid.

9.1.1B. Actual Free Energy Change. Clearly, not all reactions commence with standard 1.0 molar concentrations of reactants and products. What free energy changes are associated with initial nonstandard reaction conditions? We describe such reactions by a free energy change that, to emphasize its distinction from ΔG°, we term an **actual free energy change (ΔG).** It is defined by the equation

$$\Delta G = \Delta G^\circ + RT \ln \frac{[\text{products}]}{[\text{reactants}]} \qquad (9.3)$$

where the ratio of products to reactants is that under *actual initial reaction conditions.* This ratio is *not an equilibrium constant,* so that the last term of Eq. (9.3) *cannot* be replaced by $RT \ln K'_{eq}$!

Only if the reaction is already at equilibrium are the initial concentrations of reactants and products identical to their equilibrium concentrations. *Then, and only then,* can the ratio in Eq. (9.3) be replaced by the equilibrium constant. At equilibrium (and since $\Delta G^\circ = -RT \ln K'_{eq}$), Eq. (9.3) becomes

$$\Delta G = \Delta G^\circ + RT \ln K'_{eq} = -RT \ln K'_{eq} + RT \ln K'_{eq} = 0$$

It follows that *at equilibrium, and only at equilibrium,* ΔG is equal to zero. We can define the actual free energy change in two ways, much as we did for the standard free energy change. Equation (9.3) represents the *mathematical definition.* The *conceptual definition* of ΔG states that ΔG constitutes the free energy change of a reaction in which reactants and products are mixed at the indicated initial concentrations, at a temperature of 25°C, and at a pressure of 1.0 atm, and the reaction is allowed to proceed to equilibrium.

As you can see from Eq.(9.3), the actual free energy change depends on both the standard free energy change and the actual initial concentrations of reactants and products.

9.1.2. Biochemical Free Energy Changes

9.1.2A. Biochemical Standard Free Energy Change. While ΔG° adequately describes ordinary chemical reactions, it fails to do so for biochemical systems. The reason is that many biochemical reactions involve *protons* as either a reactant or a product. Because ΔG° refers to reaction conditions at which all reactants and products are at an initial 1.0*M* concentration, reactions involving protons would require a proton concentration of $[H^+] = 1.0M$, so that pH = 0.

A pH of zero constitutes a highly "unbiological" condition since most biochemical systems operate in the vicinity of neutrality (pH 7.0). Because ΔG° is not very meaningful under physiological conditions, we use a different standard that we term **biochemical standard free energy change** and designate $\Delta G^{\circ\prime}$ ("delta G zero prime"); it is based on a standard state that has a proton concentration of $10^{-7}M$ (pH 7.0).

We define the biochemical standard free energy change *mathematically* in analogous fashion to ΔG°. Specifically,

$$\Delta G^{\circ\prime} = -RT \ln K'_{bio} \qquad (9.4)$$

where K'_{bio} represents the equilibrium constant calculated for the biochemical reaction conditions of pH 7.0. (A detailed treatment of K'_{bio} and its relationship to K'_{eq} is beyond the scope of this book.)

The *conceptual definition* of $\Delta G^{\circ\prime}$ is also analogous to that for ΔG°:

> The biochemical standard free energy change of a reaction is the free energy change associated with the reaction when all reactants and all products are at an initial concentration of 1.0*M* each, except for protons, the initial concentration of which, unless otherwise specified, is taken as $[H^+] = 10^{-7}M$ (pH 7.0); the temperature is 25°C; the pressure is 1.0 atm; and the reaction is allowed to proceed to equilibrium.

Table 9.2 lists biochemical standard free energy changes ($\Delta G^{\circ\prime}$) for a number of reactions.

Let us return to the ionization of acetic acid. $\Delta G^{\circ\prime}$ for this reaction has a value of -12.8 kJ mol^{-1}. This means that

Table 9.2. Biochemical Standard Free Energy Changes ($\Delta G^{\circ\prime}$) of Some Reactions of Biochemical Relevance (pH 7.0, 25°C)

Reaction[a]	$\Delta G^{\circ\prime}$ (kJ mol^{-1})
Hydrolysis	
$Fumarate^{2-} + H_2O \rightleftharpoons malate^{2-}$	−3.3
$Glycylglycine + H_2O \rightleftharpoons 2\ glycine$	−9.2
$Glucose\ 6\text{-}P^{2-} + H_2O \rightleftharpoons glucose + HPO_4^{2-}$	−13.8
$Fructose\ 1,6\text{-}bis\text{-}P^{4-} + H_2O \rightleftharpoons fructose\ 6\text{-}P^{2-} + HPO_4^{2-}$	−14.2
$Glutamine + H_2O \rightleftharpoons glutamate^{-} + NH_4^{+}$	−14.2
$AMP^{2-} + H_2O \rightleftharpoons adenosine + HPO_4^{2-}$	−14.6
$Maltose + H_2O \rightleftharpoons 2\ glucose$	−15.5
$Ethyl\ acetate + H_2O \rightleftharpoons acetate^{-} + ethanol + H^{+}$	−20.1
$Acetylcholine^{+} + H_2O \rightleftharpoons acetate^{-} + choline^{+} + H^{+}$	−25.1
$Sucrose + H_2O \rightleftharpoons fructose + glucose$	−27.6
$ATP^{4-} + H_2O \rightleftharpoons ADP^{3-} + HPO_4^{2-} + H^{+}$	−30.5
$ADP^{3-} + H_2O \rightleftharpoons AMP^{2-} + HPO_4^{2-} + H^{+}$	−30.5
$ATP^{4-} + H_2O \rightleftharpoons AMP^{2-} + HP_2O_7^{3-} + H^{+}$	−31.8
$HP_2O_7^{3-} + H_2O \rightleftharpoons 2HPO_4^{2-} + H^{+}$	−33.1
$Phosphoarginine^{-} + H_2O \rightleftharpoons arginine^{+} + HPO_4^{2-}$	−38.1
$Phosphocreatine^{2-} + H_2O \rightleftharpoons creatine + HPO_4^{2-}$	−42.7
$Acetylphosphate^{2-} + H_2O \rightleftharpoons acetate^{-} + HPO_4^{2-} + H^{+}$	−43.1
$1,3\text{-}bis\text{-}P\text{-}glycerate^{4-} + H_2O \rightleftharpoons 3\text{-}P\text{-}glycerate^{2-} + HPO_4^{2-}$	−49.4
$Carbamoyl\text{-}P^{2-} + H_2O \rightleftharpoons carbamate^{-} + HPO_4^{2-} + H^{+}$	−51.5
$Phosphoenolpyruvate^{3-} + H_2O \rightleftharpoons pyruvate^{-} + HPO_4^{2-}$	−61.9
Isomerization	
$Dihydroxyacetone\ P^{2-} \rightleftharpoons glyceraldehyde\ 3\text{-}P^{2-}$	+7.7
$Citrate^{3-} \rightleftharpoons isocitrate^{3-}$	+6.7
$Glycerate\ 3\text{-}phosphate^{3-} \rightleftharpoons glycerate\ 2\text{-}phosphate^{3-}$	+4.6
$Fructose\ 6\text{-}phosphate^{2-} \rightleftharpoons glucose\ 6\text{-}phosphate^{2-}$	−1.7
$Glucose\ 1\text{-}phosphate^{2-} \rightleftharpoons glucose\ 6\text{-}phosphate^{2-}$	−7.3
Oxidation (dehydrogenases)	
$Lactate^{-} + NAD^{+} \rightleftharpoons pyruvate^{-} + NADH + H^{+}$	+25.1
$Ethanol + NAD^{+} \rightleftharpoons acetaldehyde + NADH + H^{+}$	+22.6
$Glyceraldehyde\ 3\text{-}P^{2-} + NAD^{+} + HPO_4^{2-} \rightleftharpoons 1,3\text{-}bis\text{-}P\text{-}glycerate^{4-} + NADH + H^{+}$	+6.3
Oxidation (molecular oxygen)	
$Glucose + 6O_2 \rightleftharpoons 6CO_2 + 6H_2O$	−2870
$Palmitic\ acid + 23O_2 \rightleftharpoons 16CO_2 + 16H_2O$	−9782

[a] P = "Phosphate" or "phospho-."

if you were to mix 1.0*M* CH_3COOH, 1.0*M* CH_3COO^{-}, and $10^{-7}M$ H^{+} (pH 7.0), the reaction is exergonic and would proceed spontaneously. The low proton concentration favors the ionization by "pulling" the reaction from left to right. By contrast, we saw above that a high proton concentration ($[H^{+}] = 1.0M$; pH 0) depresses the ionization, so that the standard free energy change (ΔG°) is positive.

9.1.2B. Biochemical Actual Free Energy Change. Free energy changes of nonstandard reaction mixtures—those having initial concentrations of reactants and products that are not 1.0*M* each, but in which the pH is still 7.0—are called **biochemical actual free energy changes** and are designated **$\Delta G'$.** The quantity $\Delta G'$ is related to $\Delta G^{\circ\prime}$ much as ΔG is related to ΔG°:

$$\Delta G' = \Delta G^{\circ\prime} + RT \ln \frac{[\text{products}]}{[\text{reactants}]} \qquad (9.5)$$

Thus, $\Delta G'$ is a function of both $\Delta G^{\circ\prime}$ and the actual initial concentrations of reactants and products.

You may wonder how $\Delta G'$ differs from ΔG. The an-

Table 9.3. Effect of Temperature on Free Energy Changes[a,b]

Temperature (°C)	ΔH (J mol^{-1})	ΔS (J deg^{-1} mol^{-1})	$-T\Delta S$[c] (J mol^{-1})	ΔG[d] (J mol^{-1})
−10	−5619	−20.54	5406	−213
0	−6008	−21.99	6008	0
+10	−6397	−23.39	6623	+226

[a]Reprinted, with permission, from I. M. Klotz, *Energy Changes in Biochemical Reactions*, Academic Press, New York (1967).
[b]For the reaction $H_2O(l) \rightleftharpoons H_2O(s)$.
[c]Calculated for T in degrees Kelvin (see Eq. 9.1).
[d]$\Delta G = \Delta H - T\Delta S$ (Eq. 9.1).

swer is that the two quantities are identical ($\Delta G = \Delta G'$). There can be only *one actual free energy change for any given reaction,* regardless of how we designate it. Whether we base this actual free energy change on $\Delta G°$ (and designate it as ΔG) or base it on $\Delta G°'$ (and designate it as $\Delta G'$) is irrelevant. There can only be one set of actual conditions defined by the initial concentrations of reactants and products and by the pH.

Accordingly, the *mathematical* and *conceptual* definitions of ΔG apply to $\Delta G'$ as well, and you can calculate the actual free energy change by means of either Eq.(9.3) or Eq.(9.5). Keep in mind, though, that while the two actual free energy changes are identical, the corresponding standard free energy changes ($\Delta G°$ and $\Delta G°'$) are not. Standard free energy changes differ because we base them on *different standard states.* In order to be consistent, and in order to stress that free energy changes are computed for pH 7.0, we will henceforth use exclusively $\Delta G°'$ and $\Delta G'$.

9.1.3. The Effects of Variables on Free Energy Changes

Temperature and concentration represent two important variables that affect free energy changes of chemical reactions.

9.1.3A. The Effect of Temperature.

Temperature influences free energy changes because of its effect on equilibrium constants. Equilibrium constants vary with temperature as described by the **van't Hoff equation:**

$$\log \frac{K'_{eq_1}}{K'_{eq_2}} = \frac{\Delta H°}{2.303R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \qquad (9.6)$$

where K'_{eq_1} and K'_{eq_2} are the equilibrium constants at the absolute temperatures T_1 and T_2, respectively, and $\Delta H°$ is the standard enthalpy change, assumed to be constant over the temperature range T_1–T_2. We can illustrate the effect of temperature by considering the conversion of liquid water to ice at −10, 0, and 10°C (Table 9.3). At all three temperatures, the reaction has a negative entropy change; the system becomes less random as liquid water changes to solid ice. The reaction also has a negative enthalpy change at all three temperatures; it is *exothermic* and gives off heat. The free energy change, however, is negative, zero, or positive, and it determines whether the reaction proceeds as written.

At −10°C, the reaction has a negative free energy change and proceeds spontaneously. Water cooled to −10°C freezes to form ice. At 0°C, the reaction has a free energy change of zero, and the system is at equilibrium. A mixture of water and ice will keep as such, provided you maintain the temperature at 0°C. Lastly, at 10°C, the reaction has a positive free energy change and does not proceed as written. At this temperature, water cannot freeze spontaneously to form ice.

9.1.3B. The Effect of Concentration.

Table 9.4 illustrates the effect of concentration by listing free energy changes for the isomerization of glyceraldehyde3-phosphate (G3P) and dihydroxyacetone phosphate (DHAP) at pH 7.0. This represents a key reaction in *glycolysis,* the major catabolic pathway of carbohydrates.

When reactant and product concentrations are 1.0*M* each (line 1 in Table 9.4), the free energy change consti-

Table 9.4 Effect of Concentration on Free Energy Changes[a]

Dihydroxyacetone phosphate (*M*)	Glyceraldehyde 3-phosphate (*M*)	Free energy change at pH 7.0 and 25°C (kJ mol^{-1})
1.0	1.0	+7.7 ($\Delta G°'$)
2.0×10^{-1}	9.0×10^{-3}	0 ($\Delta G'$)
1.0×10^{-1}	1.0×10^{-4}	−9.5 ($\Delta G'$)
1.0×10^{-4}	1.0×10^{-1}	+24.8 ($\Delta G'$)

[a]For the reaction
Dihydroxyacetone phosphate (DHAP) ⇌ glyceraldehyde 3-phosphate (G3P).

tutes $\Delta G^{\circ\prime}$. Its positive value indicates that the reaction is endergonic and does not proceed. The remaining free energy changes of Table 9.4 represent $\Delta G'$ values. Initial concentrations chosen for line 2 lead to a free energy change of zero; the system is *at equilibrium,* and no further net change takes place. You can calculate the ratio of [G3P]/[DHAP] at equilibrium from the data in Table 9.4:

$$(9.0 \times 10^{-3})/(2.0 \times 10^{-1}) = 0.045$$

or from Eq. (9.5) by setting $\Delta G' = 0$ and $\Delta G^{\circ\prime} = +7700$ J mol^{-1}:

$$0 = 7700 + (8.31)(298.2)(2.303)\log([\text{G3P}]/[\text{DHAP}])$$

$$\log([\text{G3P}]/[\text{DHAP}]) = -1.3492$$

$$[\text{G3P}]/[\text{DHAP}] = 0.045$$

Thus, to drive the isomerization toward G3P synthesis, the ratio [G3P]/[DHAP] can have a maximum value of 0.045. Ratios smaller than 0.045 result in exergonic reactions that proceed spontaneously (line 3). Ratios larger than 0.045 lead to endergonic reactions that cannot take place unless driven by added energy (line 4).

Our discussion of the effects of temperature, pH, and concentration on free energy changes leads to a most important conclusion:

> Any chemical reaction may be made to go in one direction or another by suitable changes in temperature and/or pH (if the reaction is pH dependent) and/or the concentrations of reactants and products.

In later chapters, we consider biochemical reactions in terms of their $\Delta G^{\circ\prime}$ values. Remember at all times that a reaction that is feasible on the basis of its $\Delta G^{\circ\prime}$ value *may or may not* be feasible on the basis of its $\Delta G'$ value. Feasibility under intracellular conditions depends on the biochemical *actual* ($\Delta G'$) and not the *standard* ($\Delta G^{\circ\prime}$) free energy change of the reaction. While temperature and pH are constant for both of these free energy changes, concentrations of reactants and products may differ widely. *What ultimately determines whether and in which direction a reaction proceeds inside the cell are the intracellular concentrations of reactants and products.*

9.2. ENERGY-RICH COMPOUNDS

The overall process whereby chemical energy of nutrients is released and used in metabolism involves a series of oxidation–reduction reactions termed **biological oxidation.** In plants and animals, biological oxidation refers specifically to **aerobic cellular respiration** (Chapters 11 and 12). This comprises the oxidation of nutrients in the *citric acid cycle,* transport of the electrons thus released via the *electron transport system,* and synthesis of ATP in a process called *oxidative phosphorylation.*

9.2.1. Biological Oxidation

After scientists realized that they could treat biochemical reactions like ordinary chemical reactions, they concluded that biological oxidation must be accompanied by large decreases in free energy. The complete oxidation of glucose to CO_2 and H_2O, for example, has a $\Delta G^{\circ\prime}$ value of -2870 kJ mol^{-1}. Investigators then began to ask how these sizable amounts of energy are released and used by an organism.

Immediate release of the entire energy in the form of heat could be ruled out. First of all, producing such large quantities of heat would be devastating to most cells, as it would lead to denaturation of enzymes, other proteins, and nucleic acids. Denaturation of these vital biomolecules would most likely result in cell death. Second, living systems generally operate essentially in an *isothermal* fashion—they function at constant temperature. According to the second law of thermodynamics, one cannot derive useful work from heat in an isothermal system. Even if cells could tolerate the deleterious effects of a large release of heat, they could not derive usable energy from it.

The answer to the problem of energy utilization lies in compounds that serve as temporary storage forms for the energy released from nutrients. We call these compounds **energy-rich compounds** and can outline the entire process as follows:

- Catabolism of foodstuffs releases the chemical energy stored in nutrients in the form of free energy.
- Before the released free energy can be used in anabolism, it must be stored in energy-rich compounds.
- *ATP (5′*-adenosine triphosphate) represents the most prevalent energy-rich compound. Its synthesis from ADP (5′-adenosine diphosphate) and P_i requires the input of 30.5 kJ mol^{-1} and constitutes a strongly endergonic reaction.
- The free energy released from nutrients drives the synthesis of ATP and other energy-rich compounds. These reactions reshuffle the stored chemical energy of nutrients and convert it to stored chemical energy of energy-rich compounds. The process is characteristic of the catabolic phase of metabolism (Figure 9.1).

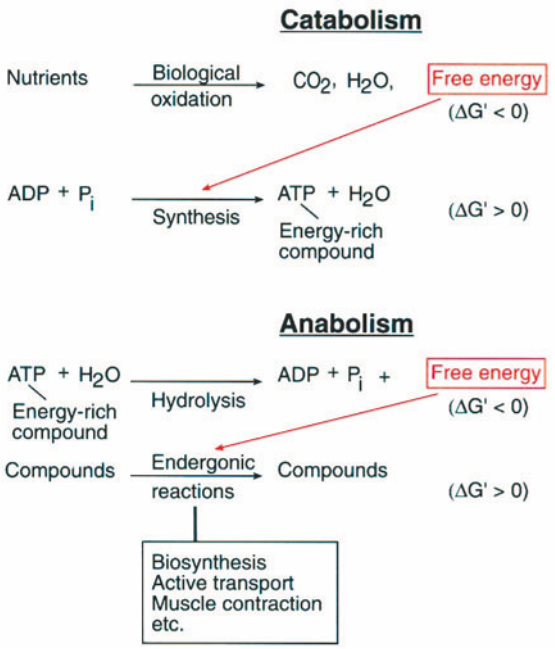


Figure 9.1. The link between free energy changes in metabolism and the role of ATP as an energy-rich compound. Oxidation of nutrients in catabolism releases energy, which drives ATP synthesis. Hydrolysis of ATP provides energy to drive endergonic reactions of anabolism.

- The cell or organism subsequently uses the energy stored in energy-rich compounds by a reversal of the above process. Hydrolysis or other cleavage of an energy-rich compound constitutes an exergonic reaction. The stored free energy released drives an endergonic reaction of metabolism. The process is characteristic of the anabolic phase of metabolism (Figure 9.1).
- The linking of exergonic and endergonic reactions that occurs in both catabolism and anabolism comprises *energetically coupled reactions,* discussed in Section 9.3. We often refer to the entire network of energy metabolism, based on ATP, as the **ATP cycle** (Figure 9.2).

Instead of simple hydrolysis, energy-rich compounds typically undergo an exergonic cleavage reaction, in which part of the molecule is *transferred* to another compound. For this reason, biochemists also describe energy-rich compounds as having **high chemical transfer potentials.** In the case of ATP, cleavage reactions typically involve transfer of either its phosphate group or its AMP moiety to another compound; ATP has a *high phosphate transfer potential.* In older terminology, energy-rich compounds were called *high-energy compounds.* In this book, we will use the term "energy-rich" exclusively.

9.2.2. Definition of Energy-Rich Compounds

Energy-rich compounds occupy a central position in the operation of metabolism. To understand their function clearly, we begin with a definition. *An energy-rich compound is a compound the hydrolysis of which has a highly negative biochemical standard free energy change.* This definition contains several parts.

First, note that we define energy-rich compounds by reference to their *hydrolysis* reaction. The free energy changes involved in other reactions in which these compounds participate are irrelevant. It is also immaterial whether an actual hydrolysis or some other cleavage takes place as the compound functions in metabolism. If the hydrolysis of a compound is highly exergonic, we deem the compound energy-rich. Some researchers define an energy-rich compounds more broadly as any compound that undergoes a reaction having a highly negative free energy change. We limit our discussion here to compounds undergoing hydrolysis, since that constitutes the commonly accepted definition.

Second, we need to clarify what we mean by a *highly negative biochemical standard free energy change.* The free energy change of interest is $\Delta G^{\circ\prime}$. Recall that other measures of free energy (ΔG°, $\Delta G'$) may differ significantly from $\Delta G^{\circ\prime}$. If hydrolysis of a compound has a highly negative $\Delta G^{\circ\prime}$, we consider the compound energy-rich.

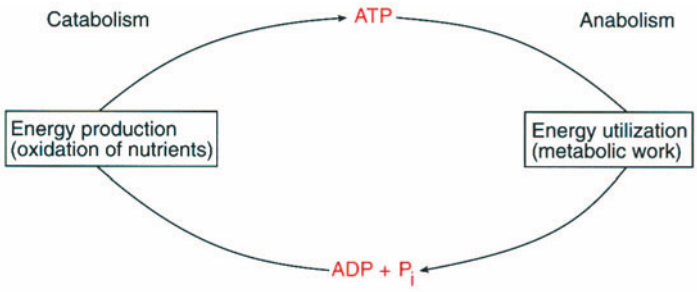


Figure 9.2. The ATP cycle.

The dividing line between moderately and highly negative free energy changes is arbitrary and depends on the values actually determined for various compounds in metabolism. Generally speaking, hydrolysis of an energy-rich compound has a biochemical standard free energy change equal to, or more negative than, about -30.5 kJ mol^{-1} (see Table 9.2).

Third, we have to stress that the definition of an energy-rich compound refers to the *free energy change of the reaction.* As you saw at the beginning of this chapter, a reaction's free energy change constitutes the difference between the sum of the free energies of *all of the products* and the sum of the free energies of *all of the reactants.* As an example, consider the hydrolysis of ATP:

$$ATP^{4-} + H_2O \rightleftarrows ADP^{3-} + P_i^{2-} + H^+$$

The free energy change of this reaction equals the difference in free energy between (ADP + P_i + H^+) and (ATP + H_2O). The biochemical standard free energy change between (ADP + P_i + H^+) and (ATP + H_2O) is -30.5 kJ mol^{-1} (see Table 9.2).

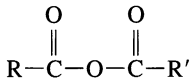
We must distinguish clearly between the energy change of a *reaction* and the energy change for breakage of an individual bond. Biochemists refer to the bond involved in hydrolysis of an energy-rich compound, especially when dealing with a phosphorylated compound (like ADP or ATP), as an **energy-rich bond** (previously called a *high-energy bond*). We frequently designate such energy-rich bonds by a *squiggle* (~).

To a chemist, an energy-rich bond is one that requires a large amount of energy in order to be broken; the bond has a large *bond energy.* If breaking the bond between atoms A and B requires a large amount of energy, chemists consider the A—B bond to be energy-rich. *A biochemist understands something quite different by the term energy-rich bond!* To a biochemist, an energy-rich bond is a bond that, when broken by *hydrolysis,* leads to a highly negative free energy change *for the reaction.* The energy relates to hydrolysis and not to any other type of reaction that might lead to bond cleavage. Moreover, *the energy involved is that of the entire reaction and not that involved in bonding two atoms.* The biochemist considers the free energy content of all products and all reactants in a hydrolysis reaction; the chemist deals with the energy of the bond between two atoms.

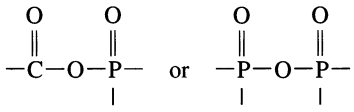
9.2.3. Types of Energy-Rich Compounds

Having defined energy-rich compounds, let us see what types of compounds fall into this category and why their hydrolysis reactions are so exergonic. Three major types of energy-rich compounds occur—**acid anhydrides, special esters,** and **derivatives of phosphamic acid** (Figure 9.3).

An ordinary acid anhydride has the following structure:



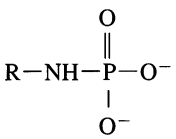
Acid anhydrides of importance in biochemistry generally have one or both of the carbon atoms replaced by phosphorus atoms, so that the compound has the structure



I **II**

We refer to compounds having structure **I** as *mixed anhydrides.* Acid anhydrides in which both carbons have been replaced by phosphorus atoms (**II**) include nucleoside di- and triphosphates like ADP, GDP, ATP, and GTP.

Special esters constitute a second type of energy-rich compounds. examples of such esters are sulfur esters (*acyl thioesters*) and *esters of enols.* The third type of energy-rich compounds consists of derivatives of *phosphamic acid,* such as *phosphocreatine.* These compounds have the general formula



9.2.4. Reasons for Negative $\Delta G^{\circ\prime}$ of Hydrolysis

Now we have to ask what makes the compounds shown in Figure 9.3 energy-rich. Why do their hydrolysis reactions proceed with highly negative free energy changes? The hydrolysis of ATP to ADP and P_i illustrates some of the factors that may be involved.

Consider the hydrolysis of ATP at pH 7–8. At that pH, adenine is uncharged (Figure 7.3). However, free phosphoric acid and the phosphate groups in ADP and ATP have each lost both their primary and their secondary proton by dissociation (Figure 9.4). Inspecting the structures of the reactants and products reveals three reasons for the highly negative free energy change of the hydrolysis reaction: resonance stabilization, electrical repulsion, and free energy of ionization.

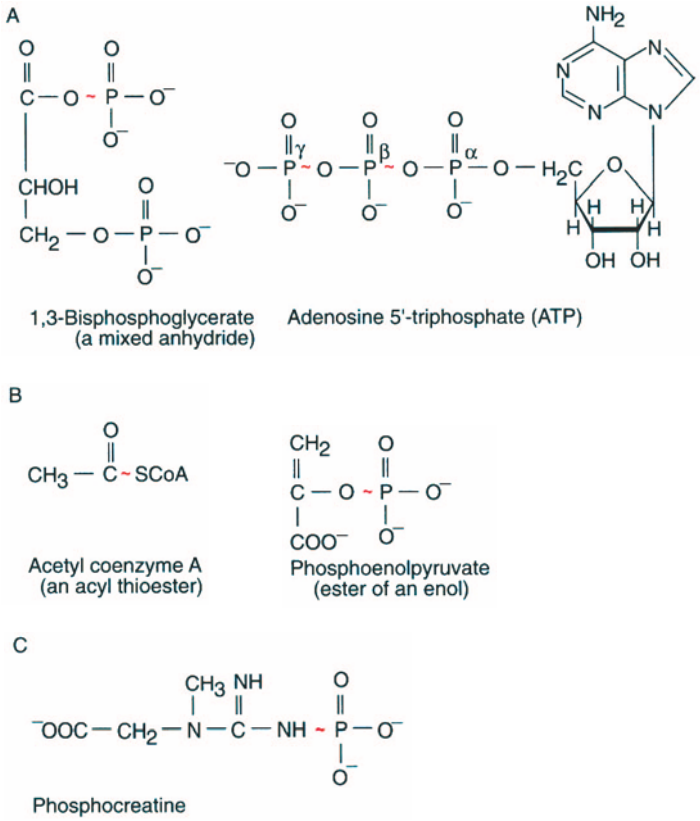
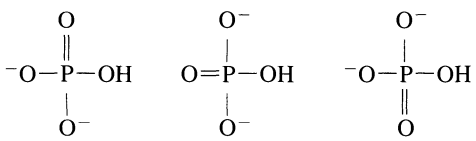


Figure 9.3. Examples of energy-rich compounds. (A) Acid anhydrides; (B) special esters; (C) derivatives of phosphamic acid; ~ designates an energy-rich bond.

9.2.4A. Resonance Stabilization. ATP, ADP, and P_i occur in the form of many resonance structures that differ in the location of single and double bonds between the phosphorus atoms and the oxygen atoms to which they are linked. For example, HPO_4^{2-} has the following possible resonance structures:



When you count all possible combinations of resonance structures, you find a larger number for the products than for the reactants. Because the stability of a compound increases with the number of resonance structures that it can assume, the products of ATP hydrolysis have greater stability than the reactants. In other words, ATP hydrolysis leads to **resonance stabilization.** A great tendency exists for the reaction to proceed as written, converting less stable reactants to more stable products. Thus, resonance stabilization contributes to making $\Delta G^{\circ\prime}$ highly negative.

9.2.4B. Electrical Repulsion. ATP possesses four negative charges in close proximity, resulting in large internal **electrical repulsion.** The intramolecular strain is so great that the molecule tends toward breakage of some of its covalent bonds. In ADP, there occur only three such charges for a molecule of almost the same size as ATP. That still amounts to significant electrical repulsion (in fact, ADP is also an energy-rich compound) but also represents a sizable decrease in repulsion from that in ATP. Effectively, one-quarter of the charges have been eliminated. Thus, ADP, although still unstable, has greater stability than ATP. As in the case of resonance stabilization, electrical repulsion leads to the products having greater stability than the reactants. Therefore, electrical repulsion also enhances the tendency of the reactants to convert to products and contributes to making $\Delta G^{\circ\prime}$ highly negative.

9.2.4C. Free Energy of Ionization. When you add up the charges for ATP and (ADP + P_i) in the hydrolysis of ATP, you find that ATP has a net charge of −4 while the net charge of (ADP + P_i) is −5. You conclude

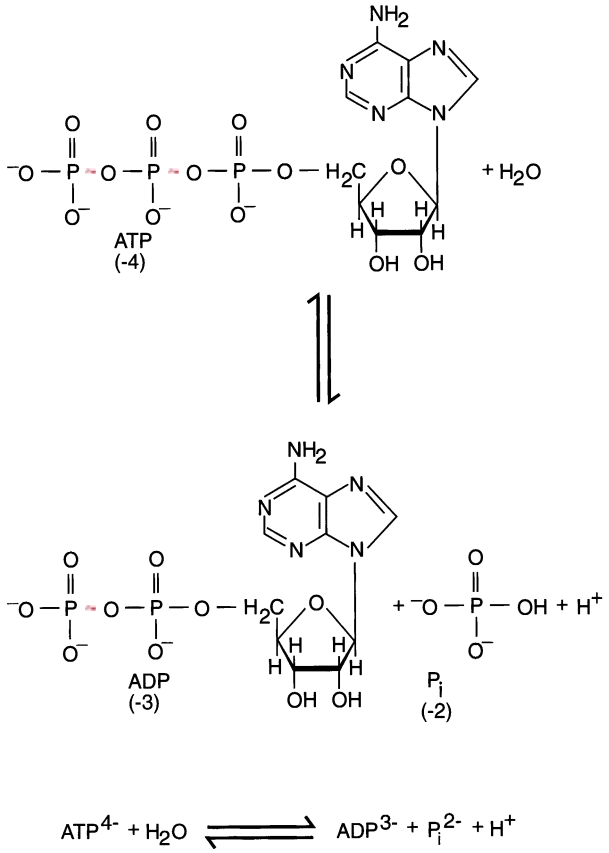


Figure 9.4. The hydrolysis of ATP to ADP and inorganic phosphate.

that another reaction, namely, an *ionization* step, must be hidden in the hydrolysis reaction. A proton must dissociate as ATP undergoes hydrolysis to ADP and P_i, and we have to write the reaction appropriately to indicate preservation of electrical neutrality. In actuality, a hydroxyl anion is removed from water, leaving an excess proton in solution. The OH^- from the water combines with the *phosphoryl group* (PO_3^{2-}), cleaved from ATP, to form HPO_4^{2-}.

Because of the ionization, the total free energy change of the hydrolysis reaction includes a contribution of **free energy of ionization.** Recall that we are evaluating ATP hydrolysis under *biochemical standard conditions,* that is, at pH 7.0. At that pH, the ionization step is exergonic and helps to make the overall free energy change highly negative.

9.3. COUPLED REACTIONS

9.3.1. Linking of Reactions

The essence of energy metabolism lies in linking endergonic and exergonic reactions (see Figure 9.1). When an exergonic reaction *drives* an endergonic one, we refer to the two reactions as **coupled** or **energetically coupled.** Energy-rich compounds function via energetically coupled reactions. Coupling applies both to the synthesis of energy-rich compounds in catabolism and to their use in anabolism. To examine the nature of this coupling, consider the following two reactions:

$$\text{Glutamate}^- + NH_3 + H^+ \rightleftarrows \text{glutamine} + H_2O$$

$$\Delta G_1^{\circ\prime} = +15.0 \text{ kJ mol}^{-1} \quad (9.7)$$

$$ATP^{4-} + H_2O \rightleftarrows ADP^{3-} + P_i^{2-} + H^+$$

$$\Delta G_2^{\circ\prime} = -30.5 \text{ kJ mol}^{-1} \quad (9.8)$$

Here we have one reaction that, under biochemical standard conditions, is endergonic ($\Delta G^{\circ\prime} > 0$) and thermodynamically not feasible. The second reaction is exergonic ($\Delta G^{\circ\prime} < 0$) and can proceed readily. It appears that we should be able to accomplish the reaction having a positive $\Delta G^{\circ\prime}$ value by energetically coupling it to the one having a negative $\Delta G^{\circ\prime}$ value. Using the free energy released by the hydrolysis of ATP, it should be possible to "drive" the energy-requiring reaction of glutamine synthesis. We might attempt to do this by simply adding the two reactions together:

$$\text{Glutamate}^- + NH_3 + H^+ \rightleftarrows \text{glutamine} + H_2O \quad (9.7)$$

$$ATP^{4-} + H_2O \rightleftarrows ADP^{3-} + P_i^{2-} + H^+ \quad (9.8)$$

Overall reaction:

$$\text{Glutamate}^- + NH_3 + ATP^{4-} \rightleftarrows \text{glutamine} + ADP^{3-} + P_i^{2-} \quad (9.9)$$

When we add two reactions, we obtain the overall free energy change by adding those of the component reactions. Hence, for the reaction given by Eq. (9.9),

$$\Delta G_{\text{overall}}^{\circ\prime} = \Delta G_1^{\circ\prime} + \Delta G_2^{\circ\prime} = 15.0 - 30.5 = -15.5 \text{ kJ mol}^{-1}$$

Because $\Delta G_{\text{overall}}^{\circ\prime}$ is negative, the reaction appears to be feasible. However, this reaction is merely the result of some "paper chemistry" that we engaged in and lacks physical reality. There exists no true chemical linkage between the two original reactions. For all we know, the two reactions may occur at different locations within the cell or at different times. How then can the energy released by one reaction actually be channeled into the other?

There exists only one way to bring about a proper linkage between the two reactions: we must *connect them chemically* by having a *product* of one reaction serve as a *reactant* for the other. Then, and only then, will the two

reactions be obligatorily linked both spacewise and timewise. To illustrate this, consider coupling the following known reactions:

$$\text{Glutamate}^- + \text{ATP}^{4-} \rightleftarrows \gamma\text{-glutamyl phosphate}^{2-} + \text{ADP}^{3-}$$

$$\Delta G_1^{\circ\prime} = 0 \qquad (9.10)$$

$$\gamma\text{-Glutamyl phosphate}^{2-} + \text{NH}_3 \rightleftarrows \text{glutamine} + \text{P}_i^{2-}$$

$$\Delta G_2^{\circ\prime} = -15.5 \text{ kJ mol}^{-1} \qquad (9.11)$$

Overall reaction:

$$\text{Glutamate}^- + \text{NH}_3 + \text{ATP}^{4-} \rightleftarrows \text{glutamine} + \text{ADP}^{3-} + \text{P}_i^{2-}$$

$$\Delta G_{\text{overall}}^{\circ\prime} = -15.5 \text{ kJ mol}^{-1} \qquad (9.12)$$

The overall reaction is identical to that written earlier, but the two schemes differ radically. Simply adding the reactions makes no chemical sense, but linking them via a shared compound—γ-glutamyl phosphate—constitutes a plausible mechanism (Figure 9.5).

Note that both schemes have the same overall $\Delta G^{\circ\prime}$.

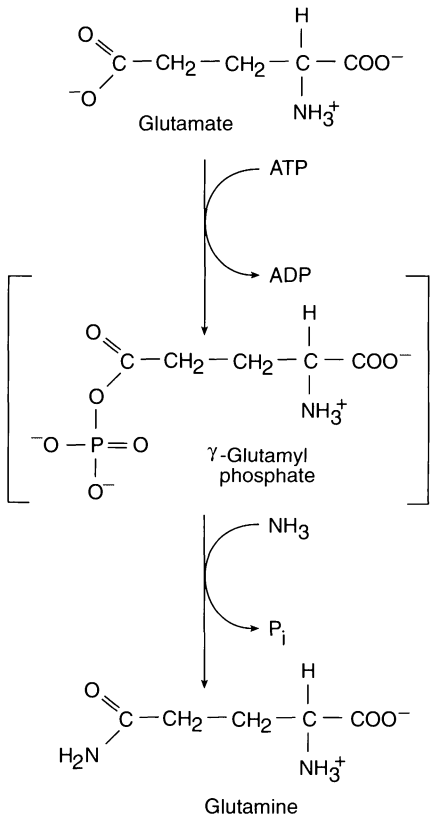


Figure 9.5. Synthesis of glutamine. The same enzyme, glutamine synthase, catalyzes both steps. The intermediate, γ-glutamyl phosphate, is unstable in aqueous solution but is protected from hydrolysis when bound to the enzyme's active site.

This is as it should be, since $\Delta G^{\circ\prime}$, like all thermodynamic functions, depends only on the final and initial states of a system and not on the path whereby the system proceeds from the initial to the final state. *γ-Glutamyl phosphate* constitutes a genuine link between the two reactions; it represents the product of one reaction and the reactant of the other reaction; it serves as a **common intermediate** of the two reactions. Both of the coupled reactions [Eqs. (9.10) and (9.11)] are catalyzed by the same enzyme, *glutamine synthase,* and the true common intermediate is actually not γ-glutamyl phosphate by itself, but rather the *enzyme-bound γ-glutamyl phosphate* complex.

Our coupling of reactions (9.10) and (9.11), under biochemical standard conditions, resulted in the first reaction having a free energy change of zero and the second reaction being strongly exergonic. The first reaction is at equilibrium, the second constitutes a spontaneous reaction, and the overall reaction proceeds readily. These specific aspects are *not* general coupling requirements; they just happen to apply in this particular example.

9.3.2. Parameters of Coupled Reactions

For two or more coupled reactions, we compute the equilibrium constant of the overall reaction by *multiplying* the equilibrium constants of the component reactions:

$$K'_{\text{eq, overall}} = K'_{\text{eq}_1} \times K'_{\text{eq}_2} \times \ldots \times K'_{\text{eq}_n} \qquad (9.13)$$

Because the free energy is a logarithmic function of the equilibrium constant [and $\log(ab) = \log a + \log b$], it follows that

$$\begin{aligned} \Delta G^{\circ}_{\text{overall}} &= -RT \ln K'_{\text{eq, overall}} \\ &= -RT \ln K'_{\text{eq}_1} - RT \ln K'_{\text{eq}_2} - \ldots - RT \ln K'_{\text{eq}_n} \\ &= \Delta G^{\circ}_1 + \Delta G^{\circ}_2 + \ldots + \Delta G^{\circ}_n \end{aligned} \qquad (9.14)$$

In other words, we obtain the free energy change for the overall reaction by *adding* the free energy changes for the individual reactions. For reactions at pH 7.0:

$$\Delta G_{\text{overall}}^{\circ\prime} = \Delta G_1^{\circ\prime} + \Delta G_2^{\circ\prime} + \ldots + \Delta_n^{\circ\prime} \qquad (9.14a)$$

or

$$\Delta G_{\text{overall}}^{\circ\prime} = \Delta G'_1 + \Delta G'_2 + \ldots + \Delta G'_n \qquad (9.14b)$$

and

$$K'_{\text{bio, overall}} = K'_{\text{bio}_1} \times K'_{\text{bio}_2} \times \ldots \times K'_{\text{bio}_n} \qquad (9.13a)$$

9.3.3. Common Intermediate Principle

The preceding discussion leads to the following generalization:

> Any two energetically coupled reactions must always proceed via a common intermediate; there exists no other way of coupling reactions.

We refer to this concept as the **common intermediate principle.** The common intermediate may be a shared component, directly apparent by inspecting the two coupled reactions:

$$A + B \rightleftarrows C + D$$
$$C \rightleftarrows X + Y$$

Alternatively, the common intermediate may participate in the mechanism without its nature being immediately apparent. Enzyme-bound γ-glutamyl phosphate represents a common intermediate of this type.

Reactions involving energy-rich compounds constitute typical examples of energetically coupled reactions. Whenever an energy-rich compound is synthesized, the reaction must be coupled to some exergonic reaction by means of a common intermediate. Likewise, whenever an energy-rich compound is hydrolyzed or otherwise cleaved to drive some endergonic reaction, the two reactions must be coupled by means of a common intermediate.

In some energetically coupled reactions, ATP cleavage produces ADP and P_i; in others, it yields AMP and PP_i. Still other energetically coupled reactions proceed without ATP being involved at all. In some cases, a single enzyme catalyzes both energetically coupled reactions; in other cases, a different enzyme catalyzes each of the coupled reactions. Some coupled reactions require an energy-rich compound of some type; others proceed without the participation of a true energy-rich compound. You can see that many variations of energetically coupled reactions occur in biochemical systems.

However, regardless of the specifics of energy-rich compound involvement, or the specifics of enzyme involvement, *all energetically coupled reactions must proceed via a common intermediate.*

9.3.4. Coupling Requirements

In addition to the requirement for a common intermediate and an overall $\Delta G' \leq 0$, two other aspects of coupled reactions need to be addressed. First, note that it is customary to discuss coupled reactions in biochemistry textbooks in terms of $\Delta G^{\circ\prime}$ values, an approach that we have used as well. However, having done so, we must stress that this constitutes *an acceptable illustration* of the principle of coupled reactions but *not an acceptable description* of actual coupling in biochemical systems.

By using $\Delta G^{\circ\prime}$ values, *we have coupled reactions under, and only under, biochemical standard conditions.* In real biochemical systems, we have to deal with biochemical actual, not standard, free energy changes ($\Delta G'$ rather than $\Delta G^{\circ\prime}$). As you saw earlier, $\Delta G'$ values can differ significantly from corresponding $\Delta G^{\circ\prime}$ values. Only after obtaining $\Delta G'$ values can you proceed to determine whether, and how, any two reactions may be coupled. Moreover, recall that $\Delta G'$can vary greatly, both in magnitude and in sign, as a result of changes in the concentrations of reactants and products. Hence, coupling of two reactions on the basis of their $\Delta G^{\circ\prime}$ values *is neither a necessary nor a sufficient requirement* for coupling these reactions under intracellular conditions.

The other aspect of coupled reactions that we have not yet mentioned we may call the "ultimate coupling requirement." Even after coupling two reactions on the basis of their $\Delta G'$ values, an additional condition has to be met. In order for the proposed mechanism to be thermodynamically feasible, it is necessary that, under biochemical actual conditions, *the free energy change ($\Delta G'$) be zero or less than zero (negative) for each step of the mechanism.* This requirement exists because any step having a positive free energy change is thermodynamically not feasible. For any step to proceed, it must have either $\Delta G' = 0$ (be at equilibrium) or $\Delta G' < 0$ (occur spontaneously). We can now summarize all of the requirements for the coupling of reactions:

1. Any two coupled reactions must proceed via a *common intermediate.*
2. Coupling must be based on *biochemical actual free energy changes* ($\Delta G'$ values).
3. The mechanism must be such that the *overall $\Delta G'$ is equal to or less than zero.*
4. Each step in the mechanism must be *thermodynamically feasible;* for each step in the mechanism, it is necessary that $\Delta G'$ *be equal to or less than zero.*

SUMMARY

Exergonic reactions have a negative free energy change and proceed spontaneously. Endergonic reactions have a positive free energy change and

do not proceed spontaneously; they require an input of energy in order to proceed. The standard free energy change depends on the equilibrium constant of a reaction and on an arbitrarily selected reference state. The actual free energy change is a function of both the standard free energy change and the actual initial concentrations of reactants and products.

In biochemistry, we compute free energy changes for physiological conditions (pH 7) and refer to them as biochemical standard free energy changes ($\Delta G^{\circ\prime}$) and biochemical actual free energy changes ($\Delta G'$). The latter quantity determines whether or not a reaction will proceed as written under intracellular conditions. Free energy changes of reactions vary with temperature and concentration. Any chemical reaction may be made to go in one direction or another by suitable changes in pH (if the reaction is pH dependent), temperature, and the concentrations of reactants and products.

Energy-rich compounds are defined as compounds the hydrolysis of which has a high negative free energy change. We refer to the bond of the energy-rich compound undergoing hydrolysis as an energy-rich bond. Acid anhydrides, special esters, and derivatives of phosphamic acid represent typical energy-rich compounds. Hydrolysis of energy-rich compounds is strongly exergonic because of factors like resonance stabilization, changes in electrical repulsion, and free energy of ionization.

Free energy is produced in catabolism and stored in energy-rich compounds. The stored free energy is subsequently used to drive endergonic reactions characteristic of anabolism. Both synthesis and utilization of energy-rich compounds proceed via energetically coupled reactions in which an exergonic reaction drives an endergonic one. Coupling requires chemical linkage of the two reactions via a common intermediate, and the mechanism must be such that both the overall $\Delta G'$ and the $\Delta G'$ for each step are either equal to or less than zero.

SELECTED READINGS

Cantor, C. R., and Schimmel, P. R., *Biophysical Chemistry,* W. H. Freeman, San Francisco (1980).

Eisenberg, D., and Crothers, D., *Physical Chemistry with Applications to the Life Sciences,* Benjamin-Cummings (1979).

Fox, R. F., *Energy and the Evolution of Life,* W. H. Freeman, New York (1988).

Freifelder, D., *Principles of Physical Chemistry,* Jones & Bartlett, Boston (1985).

Garby, L., and Larsen, P. S., *Bioenergetics: Its Thermodynamic Foundations,* Cambridge University Press, Cambridge (1995).

Harris, D. A., *Bioenergetics at a Glance: An Illustrated Introduction,* Blackwell Scientific, Oxford (1994).

Kim, C. H., and Ozawa, T. (eds.), *Bioenergetics: Molecular Biology, Biochemistry and Pathology,* Plenum, New York (1990).

Makhatadze, G. I., and Privalov, P. L., Energetics of protein structure, *Adv. Protein Chem.* 47:308–425 (1995).

Nicholls, D. G., and Ferguson, S. J., *Bioenergetics,* 2nd ed., Academic Press, London (1992).

Stenesh, J., *Core Topics in Biochemistry,* Cogno Press, Kalamazoo, Michigan (1993).

REVIEW QUESTIONS

A. Define each of the following terms:

Gibbs free energy	Coupled reactions
Energy-rich bond	Resonance stabilization
K'_{bio}	Common intermediate principle

B. Differentiate between the two terms in each of the following pairs:

$\Delta G^{\circ}/\Delta G^{\circ\prime}$	$\Delta G^{\circ}/\Delta G$
$\Delta G^{\circ\prime}/\Delta G'$	$\Delta G/\Delta G'$
Energy-rich bond/energy-rich compound	Enthalpy (heat content)/entropy

C. (1) How is the ATP cycle linked to the reactions of catabolism and anabolism?
(2) Define energy-rich compounds. What types of compounds are energy-rich? Explain the factors that contribute toward making ATP an energy-rich compound.
(3) Explain what is meant by coupled reactions. List all of the coupling requirements.

PROBLEMS

9.1. Why is there only one actual free energy change for any given reaction, regardless of whether you base the calculations on $\Delta G°$ or $\Delta G°'$? In other words, explain why $\Delta G = \Delta G'$.

9.2. Free energy, like all other thermodynamic functions, always depends on the difference (designated delta, Δ) between a final and an initial state. Of what significance is this when comparing (a) *in vivo* and *in vitro* reactions and (b) enzyme-catalyzed and uncatalyzed reactions?

9.3. The binding of three noncompetitive inhibitors (A, B, C) at 25°C to an enzyme is described by the equation $E + I \rightleftarrows EI$ and is characterized by the following standard free energy changes ($\Delta G°$): -800, -4000, and $+2000$ J mol^{-1} for A, B, and C, respectively. Which of the three compounds is the strongest inhibitor under standard conditions? Why?

9.4.* Calculate $\Delta G'$ for the hydrolysis of ATP to ADP and P_i in an actively respiring cell at pH 7.0. Under those conditions, the intracellular $[P_i] = 1.00 \times 10^{-2} M$, the steady-state [ATP]/[ADP] ratio is 10.0, and the biochemical standard free energy change, $\Delta G°'$, for the hydrolysis of ATP is -30.5 kJ mol^{-1}. (Hint: Use [product]/[reactants] = [ADP][P_i]/[ATP].)

9.5. A substrate binds to an enzyme according to the equation $E + S \rightleftarrows ES + H^+$. At equilibrium, the pH of this reaction is 5.0, the substrate concentration is $1.00 \times 10^{-2} M$, and 80.0% of the enzyme is in the form of the enzyme–substrate complex. What is the standard free energy change of the reaction?

9.6.* Given the following:

(1). Glucose 1-phosphate$^{2-} \rightleftarrows$ glucose 6-phosphate^{2-}
$K'_{bio_1} = 10.0$

(2). $2ADP^{3-} \rightleftarrows ATP^{4-} + AMP^{2-}$
$\Delta G°'_2 = +2.03$ kJ mol^{-1}

(3). Glucose 6-phosphate^{2-} + ADP^{3-} + $H^+ \rightleftarrows$ glucose + ATP^{4-}
$\Delta G°'_3 = +19.1$ kJ mol^{-1}

couple these three reactions such that glucose 1-phosphate would be synthesized from glucose and ADP. What is the overall reaction? What are the values for the overall free energy change ($\Delta G°'$) and the overall equilibrium constant (K'_{bio})?

9.7. Given the following:

$$A \rightleftarrows B \qquad K'_{eq} = 1.0 \times 10^{-5}$$
$$C \rightleftarrows D \qquad K'_{eq} = 1.0 \times 10^{-3}$$

calculate the equilibrium constant for the reaction $B + C \rightleftarrows A + D$.

9.8. Refer to Table 9.2 and indicate which of the following you would classify as energy-rich compounds: (a) phosphocreatine; (b) acetylcholine; (c) phosphoenolpyruvate; (d) ethyl acetate; (e) glucose.

9.9. The equilibrium constant for a reaction at 20°C has a value of 2.70×10^{-3}. What is the value of the equilibrium constant at 25°C if $\Delta H°$ is constant and equal to -6.28 kJ mol^{-1} over the temperature range of 20–25°C?

9.10.* Couple the following two reactions

(1) Glucose 6-phosphate^{2-} + $H_2O \rightleftarrows$ glucose + P_i^{2-}
$K'_{bio_1} = 1.00 \times 10^{-2}$

(2) $ATP^{4-} + H_2O \rightleftarrows AMP^{2-} + PP_i^{3-} + H^+$
$\Delta G°'_2 = -31.8$ kJ mol^{-1}

so that ATP hydrolysis drives the synthesis of glucose 6-phosphate. Write a plausible two-step mechanism and then write the overall reaction. Calculate the overall free energy change ($\Delta G°'$) of the reaction. (Hint: K'_{bio} for ATP hydrolysis is [AMP][PP_i]/[ATP].)

9.11.* What is the biochemical actual free energy change ($\Delta G'$) for the overall reaction of the previous problem under the following steady-state conditions?

[AMP] = $1.00 \times 10^{-2} M$ [ATP] = $1.00 \times 10^{-1} M$
[PP_i] = $1.00 \times 10^{-2} M$ [P_i] = $1.00 \times 10^{-5} M$
[glucose] = $1.00 \times 10^{-4} M$ [glucose 6-phosphate] = $1.00 \times 10^{-1} M$

Hint: Omit [H^+] from the [products]/[reactants] term.

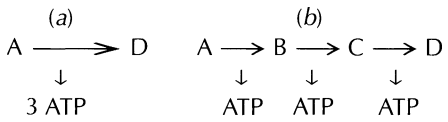
9.12. Refer to Table 9.2 to calculate the equilibrium constant (K'_{bio}) for the reaction

$$\text{Creatine} + P_i^{2-} \rightleftarrows \text{phosphocreatine}^{2-} + H_2O$$

9.13.* What must be the minimum ratio of [creatine] to [phosphocreatine] in order to drive the reaction of the previous problem toward synthesis of phospho-

creatine when $[P_i] = 1.00M$? (Hint: Omit $[H_2O]$ from the [products]/[reactants] term.)

9.14.* The hydrolysis of ATP to ADP and P_i has a free energy change ($\Delta G°'$) of -30.5 kJ mol^{-1}. The catabolic conversion of compound A to compound D is an exergonic reaction ($\Delta G°' = -91.5$ kJ mol^{-1}). Below are listed two theoretical mechanisms for coupling ATP synthesis to the catabolism of compound A. Both mechanisms have a 100% efficiency of energy trapping (three moles of ATP are synthesized per mole of compound A catabolized). Based on your general knowledge of chemistry, which of these two mechanisms represents a more plausible one and why?



9.15. In stage I of catabolism, polymeric molecules undergo degradation to numerous building blocks (see Figure 8.2). Since this amounts to an increase in randomness, the process is characterized by an increase in entropy. Hence, decide whether it is thermodynamically possible for stage I of catabolism to proceed such that $\Delta G = \Delta H$. [Hint: See Eq. (9.1).]

9.16. Consider the following two coupled reactions:

$$A \rightleftarrows B + C \qquad K'_{eq_1}$$
$$C \rightleftarrows X + Y \qquad K'_{eq_2}$$
$$A \rightleftarrows B + X + Y \qquad K'_{eq,overall}$$

Show that $K'_{eq,overall} = K'_{eq^1} \times K'_{eq^2}$ in the following two ways. (a) Write out the expression for $K'_{eq,overall}$. Then write out the expressions for K'_{eq^1} and K'_{eq^2}, multiply them, and compare the result with the expression for $K'_{eq,overall}$. (b) Solve the expressions of K'_{eq^1} and K'_{eq^2} for C. Because C is the common intermediate of the two reactions, its concentration can be expressed by either of the two component reactions. Equate the two expressions for C, and solve the resultant equation for $K'_{eq^1} \times K'_{eq^2}$.

9.17.* What is the lowest pH that would permit the ionization of acetic acid at 25°C when $[CH_3COOH] = 1.00 \times 10^{-3}M$ and $[CH_3COO^-] = 1.00M$? (Hint: Set $\Delta G' = 0$. The ionization of acetic acid has a $\Delta G°'$ value of -12.8 kJ mol^{-1}.)

9.18.* A novel bacterium is suspected of being able to grow on citrate as its sole carbon and energy source. The first step in citrate utilization consists of its conversion to isocitrate (see Table 9.2). Energetically speaking, could this organism grow when the [citrate]/[isocitrate] ratio equals 1000?

9.19. Calculate the number of moles of ATP that, theoretically, could be synthesized from ADP for every 10 moles of maltose hydrolyzed to glucose. Assume a 100% efficency of energy trapping and reactions occurring under biochemical standard conditions. Refer to Table 9.2.

9.20. A 0.10M solution of a weak monoprotic acid, HA, is 10% ionized at 25°C. Calculate (a) the pH of the solution, (b) the equilibrium (ionization) constant, (c) pK'_a, (d) and $\Delta G°$.

9.21. What are the relative molar concentrations of citrate and isocitrate at equilibrium for the isomerization reaction listed in Table 9.2?

9.22.* According to Table 8.6, a normal man aged 19–24 requires a minimum intake of about 7500 kJ per day. Assume that a man who weighs 150 lb (68 kg) consumes only enough food to meet the minimum energy requirement and converts food energy into ATP at an efficiency of 40%. Assume further that $\Delta G'$ for the hydrolysis of ATP under intracellular conditions is -30.5 kJ mol^{-1}. The molecular weight of ATP is 507. Calculate the weight of ATP produced by the man from food per day. What percent of his body weight does this represent?

9.23. What is the difference in $\Delta G°$ at 25°C for two reactions for which the equilibrium constants differ by a factor of 10?

9.24. What is the value of the equilibrium constant at 25°C and pH 7.0 for the isomerization of fructose 6-phosphate to glucose 6-phosphate? (See Table 9.2.)

9.25. What is the biochemical actual free energy change ($\Delta G'$) for the conversion of phosphoarginine to arginine (see Table 9.2) when the following initial concentrations are used? (Hint: Omit $[H_2O]$ from the [products]/[reactants] term.)

$$[\text{Phosphoarginine}] = 3.00 \times 10^{-2}M$$
$$[\text{Arginine}] = 1.00 \times 10^{-3}M$$
$$[HPO_4^{2-}] = 2.00 \times 10^{-5}M$$

9.26.* Use the data in Table 9.2 to calculate $\Delta G°'$ for the following reactions:

(a) Glucose 6-phosphate^{2-} + $6O_2$ ⇄ HPO_4^{2-} + $6CO_2$ + $5H_2O$
(b) Maltose + $2HPO_4^{2-}$ ⇄ 2 glucose 6-phosphate^{2-} + H_2O
(c) Lactate$^-$ + acetaldehyde ⇄ pyruvate$^-$ + ethanol
(d) Glucose 6-phosphate^{2-} + fructose ⇄ sucrose + P_i^{2-}

Carbohydrate Metabolism

10

We begin our study of metabolism with carbohydrates because of their central role in the generation, use, and storage of metabolic energy. Additionally, carbohydrates are of major importance in metabolism because their degradations and interconversions provide the carbon skeletons for the biosynthesis of most other metabolites, from small coenzymes to large structural molecules. After a look at some general aspects of carbohydrate metabolism, we will discuss specific metabolic pathways.

We can divide carbohydrate metabolism into five major parts—*digestion, transport, storage, degradation, and biosynthesis* (Figure 10.1). Starch and glycogen represent the major dietary polysaccharides of animals; their digestion produces **glucose.** As you will see in this chapter, other dietary carbohydrates are either converted to glucose or to intermediates of glucose metabolism. Thus, carbohydrate metabolism becomes effectively glucose metabolism; most major pathways either start or end with glucose (Figure 10.2).

Monosaccharides, the end products of oligo- and polysaccharide digestion, are carried to the liver and distributed from there throughout the body. Blood transports carbohydrates, primarily in the form of glucose, from the liver to muscle tissue and, in the form of *lactate* (formed from glucose), from muscle tissue to the liver.

Carbohydrates and lipids constitute the storage forms of energy in animals. Carbohydrates are stored as glycogen, principally in muscles and the liver. Major degradative pathways of carbohydrates include the breakdown of glycogen to glucose, **glycogenolysis,** and two catabolic pathways for glucose. One, **glycolysis,** leads to production of usable energy; the other, the **pentose phosphate pathway,** generates reducing power and pentoses.

Biosynthesis of carbohydrates includes three major pathways: synthesis of glycogen, **glycogenesis;** synthesis of glucose, **gluconeogenesis;** and fixation of carbon dioxide, **photosynthesis.** In this chapter, we cover all of these catabolic and anabolic pathways except photosynthesis, which we discuss separately in Chapter 15.

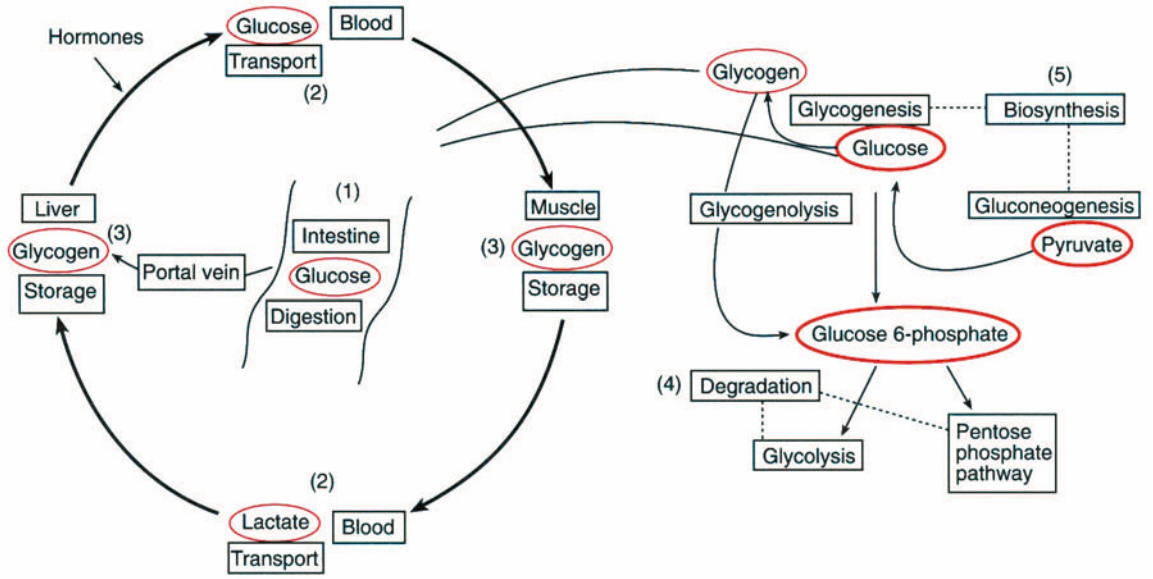


Figure 10.1. Diagram showing the major aspects of carbohydrate metabolism in animals: (1) digestion, (2) transport, (3) storage, (4) degradation, and (5) biosynthesis. Varying portions of the degradative and biosynthetic pathways occur in muscle tissue, the liver, and the blood. An additional biosynthetic pathway, photosynthesis, occurs in plants and other photosynthetic organisms.

10.1. TRANSPORT AND STORAGE OF CARBOHYDRATES

Blood transports carbohydrates primarily in the form of glucose, or *blood sugar.* Normal blood sugar levels fall in the range of 65–100 mg % (mg/100 ml or mg/dl of blood). Glucose transport is linked to that of lactate by reactions of the *Cori cycle* (Section 10.3). Three hormones control blood sugar levels:

- *Insulin,* a polypeptide (Figure 2.7) secreted by the pancreas, *lowers the level of blood sugar* by stimulating glucose transport into cells and by stimulating glycogen synthesis from glucose in both muscles and the liver.
- *Epinephrine (adrenaline),* a derivative of tyrosine (Figure 10.3) secreted by the adrenal gland, *raises the level of blood sugar* by stimulating degradation of glycogen to glucose in muscles and the liver.
- *Glucagon,* a polypeptide of 29 amino acids secreted by the pancreas, *raises the level of blood sugar* by stimulating degradation of glycogen to glucose in the liver.

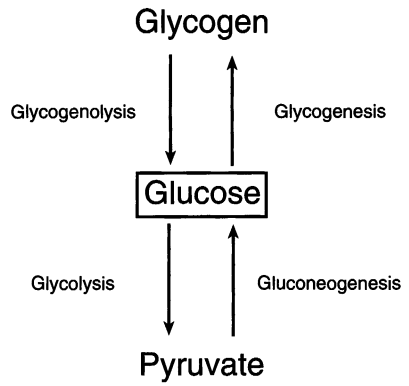


Figure 10.2. Major catabolic and anabolic pathways of glucose metabolism.

Clinicians use measurements of blood sugar to construct *glucose tolerance curves* (Figure 10.4), which serve as screening tests for **diabetes.** For this test, an individual in a fasting state first provides a blood sample to establish the baseline level of blood sugar. The person then drinks a concentrated glucose solution. Samples of blood are withdrawn periodically and analyzed for glucose. In a normal individual, the blood sugar concentration first rises as excess glucose enters the bloodstream, and subsequently drops to its normal level as the glucose is metabolized.

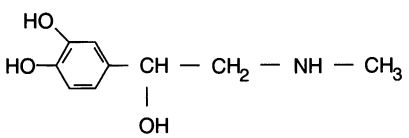


Figure 10.3. Epinephrine.

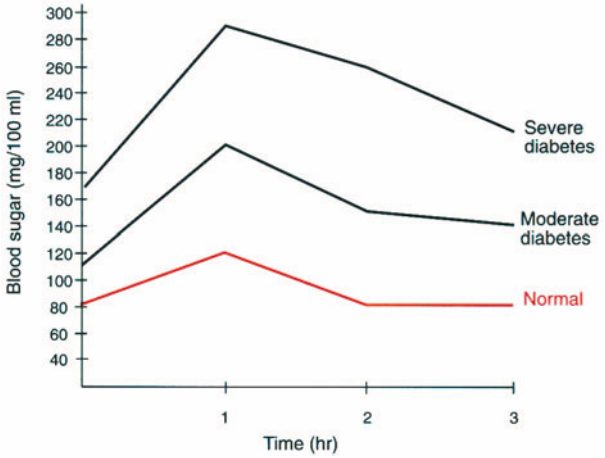


Figure 10.4. Typical glucose tolerance curves.

Typically, this occurs within about two hours. For a diabetic, the curve has a different shape. A diabetic suffers from excessive and uncontrolled levels of blood sugar, so the initial baseline level is higher than that of a healthy person. Ingesting a large amount of glucose exacerbates the problem of impaired glucose metabolism. The blood sugar concentration first rises, as in any individual, but it takes more than two hours to return to the baseline level.

There are two types of diabetes, *juvenile-onset* (insulin-dependent, Type I) and *adult-onset* (insulin-independent, Type II) diabetes. Juvenile-onset diabetes occurs in childhood and is due to an insulin deficiency that may result from insufficient synthesis, accelerated breakdown, or inhibition of insulin. Adult-onset diabetes occurs in middle-aged individuals who have insulin near or even above normal levels. The defect results from a decreased response of cells to insulin due to a scarcity of active insulin receptors in the cell membrane. Glucose tolerance curves are abnormal for both types of diabetics.

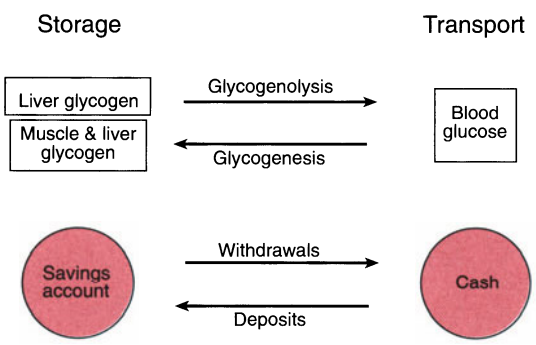


Figure 10.5. The relationship between transport and storage of carbohydrates resembles that between cash and savings. Degradation of muscle glycogen yields lactate rather than glucose (see Figure 10.26).

Carbohydrate transport and storage represent related processes (Figure 10.5). When the blood glucose concentration goes up, as after a meal, some of the glucose is metabolized to provide energy, and some is converted to glycogen and stored in muscles and the liver. Conversely, when blood glucose concentration decreases and energy is needed, some of the glycogen stored in the liver is converted to glucose. Muscle glycogen is catabolized differently (see Figure 10.26).

The interdependence of carbohydrate transport and storage resembles that of cash and savings. When you accumulate a reasonable amount of cash, you may deposit some of it in a savings account. When you need cash to pay for expenses, you withdraw the necessary funds from the savings account.

10.2. GLYCOLYSIS—INDIVIDUAL REACTIONS

The major pathway for generating usable energy from carbohydrates consists of a set of 10 reactions called *glycolysis* ("sugar splitting," from the Greek *glykos,* sweet, and *lysis,* dissolution). The reaction sequence, also known as the *Embden–Meyerhof pathway,* takes place in the cytosol, the fluid portion of the cytoplasm.

It is fitting that we begin our detailed study of metabolism with this pathway. First, glycolysis occurs almost universally in animals, plants, and microorganisms. Second, the reaction sequence plays a central role in generating both energy and metabolic intermediates for other pathways. Third, glycolysis was the first metabolic pathway to be elucidated in detail. Lastly, the regulation of glycolysis is particularly well understood. In fact, glycolysis probably represents the most completely understood metabolic pathway.

The study of glycolysis had its start in 1897 with the discovery by Eduard Buchner (awarded the Nobel Prize in 1907) that a cell-free extract of yeast could carry out fermentation. This groundbreaking work was soon followed by two other important findings. Between 1905 and 1910, Arthur Harden and William Young showed that fermentation requires inorganic phosphate for incorporation into a sugar phosphate, later shown to be fructose 1,6-bisphosphate. Additionally, Harden and Young succeeded in separating a yeast cell-free extract by dialysis into two fractions, both of which are required for fermentation. The nondialyzable fraction is heat-sensitive and was called *zymase;* the dialyzable fraction is heat-stable and was named *cozymase.* The work of other researchers established subsequently that zymase consists of a mixture of enzymes and that cozymase consists of a mixture of cofactors—including ATP, ADP, and NAD^+—and metal ions.

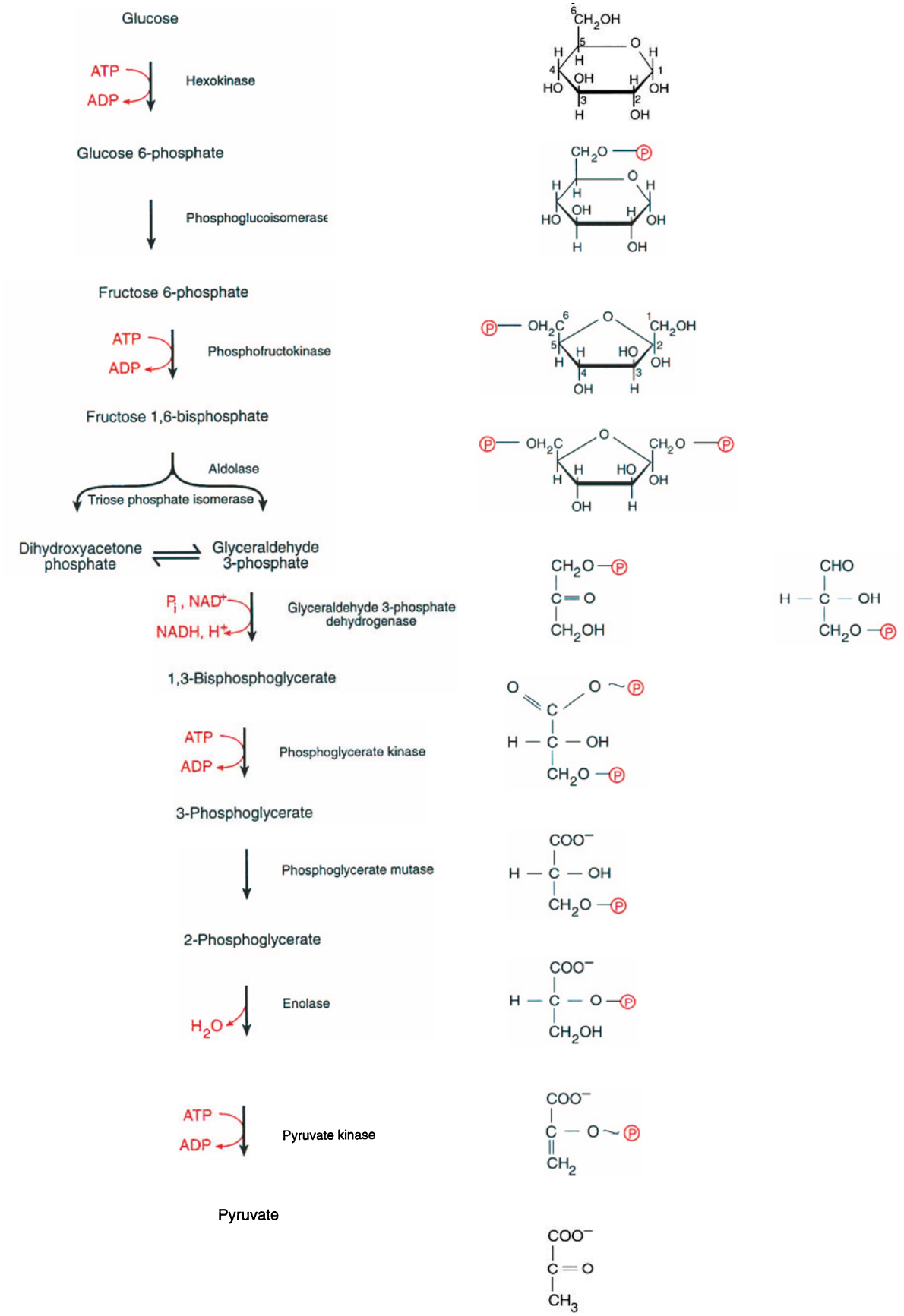


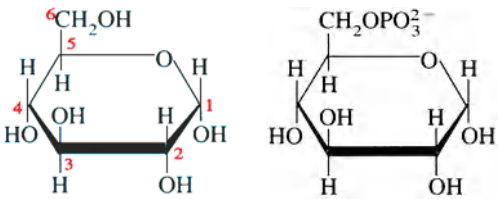
Figure 10.6. The reactions of glycolysis.

These initial investigations were followed by widespread research that dealt both with the alcoholic fermentation of yeast and the related topic of glucose catabolism in muscles. By 1940, the combined efforts of many scientists culminated in elucidation of the complete pathway of glycolysis. The field of biochemistry had been provided with its first major enzymatic pathway. Investigators who made notable contributions in this effort include Gustav Embden, Otto Meyerhof (awarded the Nobel Prize in 1922), Jacob Parnas, Carl and Gerti Cori, Carl Neuberg, Robert Robison, and Otto Warburg.

Glycolysis is an anaerobic degradative process leading to the conversion of *one* molecule of glucose to *two* molecules of pyruvate and capture of a limited amount of energy in the form of ATP. The glycolytic pathway consists of two stages (Figure 10.6). *Stage I* represents an *energy-consuming stage* that accomplishes the conversion of one molecule of glucose (six carbons) to two molecules of glyceraldehyde 3-phosphate (three carbons). This stage includes two *ATP-requiring* reactions; the energy-rich ATP "primes" glucose for glycolysis and also drives a second *phosphorylation* reaction (transfer of a *phosphoryl group*). *Stage II* represents an *energy-generating stage* that accomplishes the conversion of glyceraldehyde 3-phosphate (three carbons) to pyruvate (three carbons). This stage includes two *ATP-yielding* reactions; synthesis of ATP is *coupled* to cleavage of energy-rich bonds in two glycolytic intermediates.

10.2.1. Stage I

10.2.1A. Hexokinase. Glycolysis commences with the **hexokinase** reaction, in which glucose is phosphorylated to *glucose 6-phosphate* by means of ATP:



Glucose + ATP → glucose 6-phosphate + ADP + H^+

A **kinase** catalyzes the transfer of a phosphoryl group from ATP to a metabolite or from an energy-rich compound to ADP. Kinases usually require Mg^{2+}, which is chelated by the phosphoryl groups in ATP (Figure 10.7). The metal ion helps to orient ATP at the enzyme's active site and provides electrostatic shielding for the negative charges of phosphoryl groups. In the absence of such shielding, electron pairs of attacking nucleophiles would be repelled.

In the hexokinase reaction, ATP provides both the energy and the phosphoryl group for producing glucose 6-phosphate. The reaction is highly exergonic (Table 10.1), making it an irreversible step.

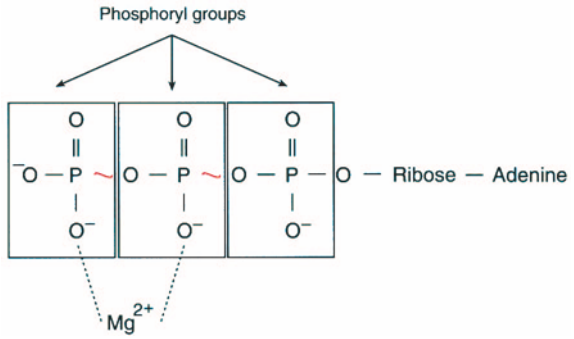


Figure 10.7. Chelation of magnesium ions by phosphoryl groups of ATP.

Glucose phosphorylation has two advantages. First, it converts the initial metabolite to a more reactive form so that subsequent reactions proceed readily; glucose 6-phosphate is chemically more reactive than glucose. Second, phosphorylation traps glucose inside the cell, since free glucose moves across the cell membrane but negatively charged glucose 6-phosphate does not. Phosphorylation ensures that glucose will not be lost from the cell. Instead, it is catabolized for the generation of energy or used to form glycogen.

All of the remaining glycolytic intermediates, except pyruvate, also are phosphorylated compounds that carry a negative charge at physiological pH. These intermediates do not move across the cell membrane since cell membranes are generally impermeable to charged molecules.

Table 10.1. The Energetics of Glycolysis[a]

Enzyme		$\Delta G^{\circ\prime}$ (kJ mol^{-1})	$\Delta G'$ (kJ mol^{-1})
Stage I			
1. Hexokinase		−16.7	−33.5
2. Phosphoglucoisomerase		+1.7	−2.5
3. Phosphofructokinase		−14.2	−22.2
4. Aldolase		+23.8	−1.3
5. Triose-phosphate isomerase		+7.5	+2.5
	Total	+2.1	−57.0
Stage II			
6. Glyceraldehyde 3-phosphate dehydrogenase		+6.3	−1.7
7. Phosphoglycerate kinase		−18.8	+1.3
8. Phosphoglyceromutase		+4.6	+0.8
9. Enolase		+1.7	−3.3
10. Pyruvate kinase		−31.4	−16.7
	Total	−37.6	−19.6

[a] $\Delta G'$ values have been calculated on the basis of the approximate intracellular concentrations of glycolytic intermediates in rabbit skeletal muscle.

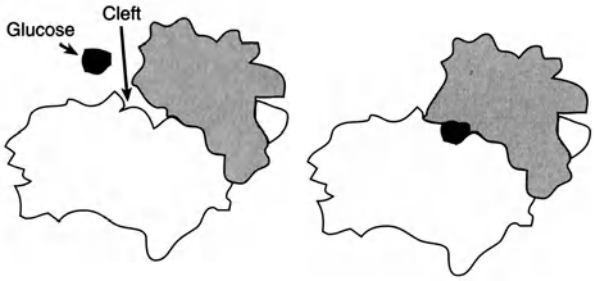


Figure 10.8. Model of hexokinase, showing closure of the cleft upon binding glucose.

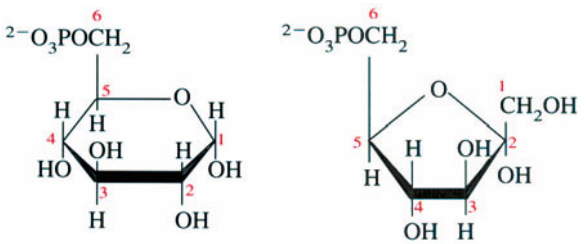
Glucose (uncharged) and lactate and pyruvate (both charged) can move across cell membranes because of the existence of specific transport mechanisms.

Hexokinase was discovered by Otto Meyerhof in 1927. The yeast enzyme is a dimer (MW = 55,000/subunit), but mammalian hexokinase is a monomer (MW = 100,000). Hexokinase catalyzes transfer of a phosphoryl group from ATP to a number of different hexoses. The mammalian enzyme is inhibited by glucose 6-phosphate, the reaction product. X-ray diffraction studies have revealed that all hexokinases have a cleft to which the substrate becomes bound (Figure 10.8). Upon binding of substrate, the enzyme undergoes a large conformational change that leads to closing of the cleft. Glucose phosphorylation takes place with the cleft in its closed state. This type of enzyme–substrate interaction illustrates *induced fit* or a *flexible active site.*

In addition to hexokinase, liver tissue contains a second kinase, *glucokinase,* that acts specifically on glucose but has a large Michaelis constant (K_m) for it. Because of the large K_m, the enzyme becomes effective only at high glucose concentrations, at which time it provides glucose 6-phosphate for glycogen synthesis. Thus glycogen synthesis from glucose occurs only when there exists an abundant supply of glucose.

10.2.1B. Phosphoglucoisomerase.

The second reaction of glycolysis, catalyzed by **phosphoglucoisomerase,** results in isomerization of glucose 6-phosphate, an aldose, to *fructose 6-phosphate,* a ketose:

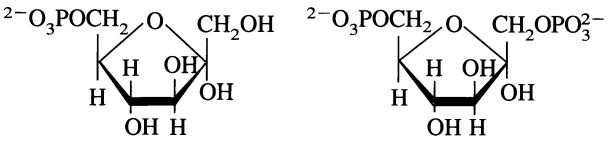


Glucose 6-phosphate ⇄ fructose 6-phosphate

The reaction proceeds via formation of an enediol [−C(OH)=C(OH)−] intermediate. Mammalian phosphoglucoisomerase is a dimer (MW = 61,000/subunit) and requires Mg^{2+} as cofactor.

10.2.1C. Phosphofructokinase.

Phosphofructokinase catalyzes a second phosphorylation by means of ATP:



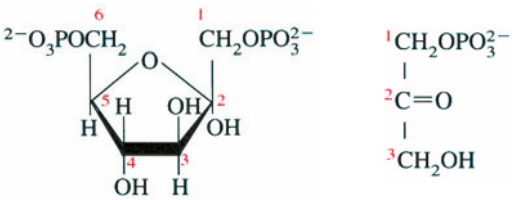
Fructose 6-phosphate + ATP^{4-} → fructose 1, 6-bisphosphate + ADP^{3-} + H^+

As in the hexokinase reaction, ATP provides both the energy to drive the phosphorylation and the phosphoryl group transferred. The phosphofructokinase reaction constitutes the *committed step* (see Section 8.2) of glycolysis. *Fructose 1,6-bisphosphate,* formed in this reaction, has no other metabolic role except serving as an intermediate in glycolysis. This fact and the highly exergonic nature of the reaction ensure that glycolysis proceeds in the direction of glucose → pyruvate.

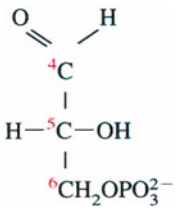
Conversion of fructose 6-phosphate to fructose 1,6-bisphosphate constitutes the most important regulatory site of glycolysis. Phosphofructokinase is an *allosteric enzyme* that requires Mg^{2+} as cofactor and that has various allosteric effectors. Activators of the enzyme include AMP, ADP, fructose 6-phosphate, fructose 1,6-bisphosphate, *fructose 2,6-bisphosphate* (discussed in Section 10.6), and P_i. Inhibitors include ATP, citrate, long-chain fatty acids, and NADH (the reduced form of nicotinamide adenine dinucleotide). The mammalian enzyme is a tetramer (MW = 78,000/subunit).

10.2.1D. Aldolase.

The enzyme catalyzing the fourth reaction of stage I is called **aldolase** because the reverse reaction constitutes an *aldol condensation:*



Fructose 1,6-bisphosphate ⇄ dihydroxyacetone phosphate +



glyceraldehyde 3-phosphate

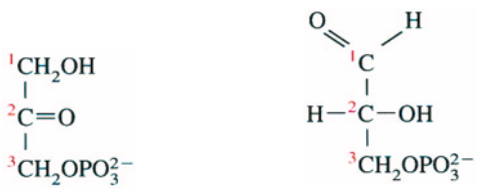
Aldolase brings about cleavage of fructose 1,6-bisphosphate (six carbons) to two three-carbon compounds, *dihydroxyacetone phosphate* and *glyceraldehyde 3-phosphate*. Under biochemical standard conditions, the aldolase reaction is strongly endergonic (see Table 10.1) and will not proceed as written, but under intracellular conditions the reaction is exergonic and proceeds spontaneously. Intracellular concentrations of reactants and products are such that the free energy change ($\Delta G'$) is negative. This illustrates the important point made earlier that free energy changes of reactions may be greatly affected by changes in the concentrations of reactants and products.

Two types of aldolases exist. In **class I aldolases,** found in animals and plants, the mechanism of the reaction involves formation of a **Schiff base** (Figure 10.9). A Schiff base represents a condensation product between a primary amine and either an aldehyde or a ketone, forming a carbon–nitrogen double bond (an *imine* group). In the aldolase reaction, the imine group forms between the ϵ-amino group of a lysine residue in the enzyme and the carbonyl group of the open-chain form of fructose 1,6-bisphosphate. Mammalian aldolase is a tetramer (MW = 40,000/subunit); the predominant form in muscles consists of two types of subunits ($\alpha_2\beta_2$).

Class II aldolases, present in bacteria and protista, are dimers that require Zn^{2+} as cofactor. These enzymes do not form a Schiff base intermediate.

10.2.1E. Triose-Phosphate Isomerase.

The last reaction of stage I comprises an isomerization between dihydroxyacetone phosphate and glyceraldehyde 3-phosphate, produced in the previous step. These two compounds are structural isomers that readily interconvert by means of **triose-phosphate isomerase:**



Dihydroxyacetone phosphate ⇄ glyceraldehyde 3-phosphate

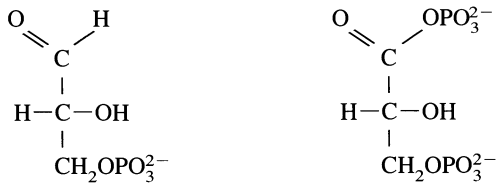
Under intracellular conditions, the reaction is slightly endergonic. It is driven by being coupled to the next reaction, an exergonic step.

10.2.2. Stage II

Of the two products formed in the aldolase reaction, *only glyceraldehyde 3-phosphate enters stage II* of glycolysis. Dihydroxyacetone phosphate isomerizes to glyceraldehyde 3-phosphate, which then enters stage II. Thus, *one molecule of glucose,* when processed through stage I, ultimately yields *two molecules of glyceraldehyde 3-phosphate* that enter the reaction sequence of stage II.

10.2.2A. Glyceraldehyde 3-Phosphate Dehydrogenase.

The first reaction of stage II is an *oxidation–reduction* step catalyzed by **glyceraldehyde 3-phosphate dehydrogenase:**



Glyceraldehyde 3-phosphate + NAD^+ + P_i^{2-} ⇄ 1,3-bisphosphoglycerate + NADH + H^+

Like other biological oxidation–reduction reactions, this particular one requires involvement of a coenzyme. The coenzyme of glyceraldehyde 3-phosphate dehydrogenase is *nicotinamide adenine dinucleotide,* or *NAD*$^+$ (see Figure 11.3). As glyceraldehyde 3-phosphate (an aldehyde) undergoes oxidation to 1,3-bisphosphoglycerate (an acid), the oxidized form of the coenzyme (NAD^+) converts to the reduced form (NADH):

$$NAD^+ + H^- \rightleftarrows NADH$$

Reduction of NAD^+ requires transfer of a hydride ion, H^- (also designated H:), which, in this case, is abstracted from the aldehyde carbon atom of glyceraldehyde 3-phosphate. A hydride ion comprises a proton and two electrons ($H^+ + 2\bar{e}$). Oxidation of a metabolite by a dehydrogenase typically involves removal of two hydrogen atoms ($2H\cdot$, equivalent to $2H^+ + 2\bar{e}$) from the metabolite. Of these, a hydride ion (equivalent to $H^+ + 2\bar{e}$) reduces NAD^+, leaving one proton (H^+) to appear as a product of the reaction. Accordingly, we usually de-

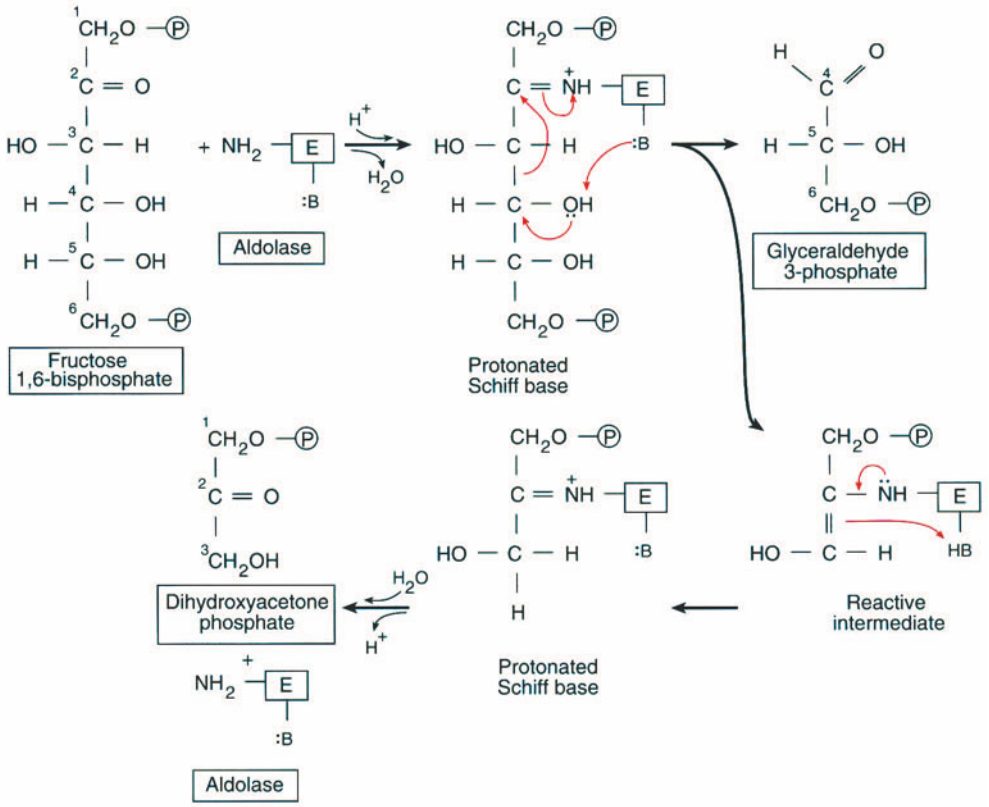


Figure 10.9. The aldolase reaction. Formation of a Schiff base is followed by release of the first product, glyeraldehyde 3-phosphate. A second Schiff base forms, followed by release of the second product, dihydroxyacetone phosphate (E, enzyme; B, a basic group).

pict the interconversion of NAD^+ and NADH as follows:

$$NAD^+ + H^- + H^+ \rightleftarrows NADH + H^+$$

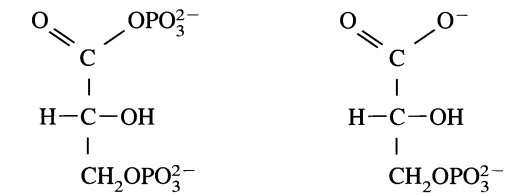
or simply as

$$NAD^+ \rightleftarrows NADH + H^+$$

1,3-Bisphosphoglycerate represents an *energy-rich compound* because of its mixed acid anhydride structure. Only the linkage between the phosphoryl group and C(1) constitutes an *energy-rich bond;* the linkage between the phosphoryl group and C(3) is an ordinary ester bond. In addition to being an intermediate in glycolysis, 1,3-bisphosphoglycerate also serves as the source of *2,3-bisphosphoglycerate* (via the enzyme *bisphosphoglycerate mutase*), an allosteric effector of hemoglobin.

Glyceraldehyde 3-phosphate dehydrogenase contains an active sulfhydryl group that participates in the reaction (Figure 10.10). Mammalian glyceraldehyde 3-phosphate dehydrogenase is a tetramer of four identical subunits (MW = 37,000/subunit), and each subunit has one binding site for NAD^+. Arsenate ($HAsO_4^{2-}$) can substitute for phosphate in the enzyme mechanism. When that occurs, glycolysis can proceed, but ATP synthesis in stage II is inhibited. Decreased ATP synthesis in the presence of arsenate accounts for the toxicity of this compound.

10.2.2B. Phosphoglycerate Kinase. Transfer of the phosphoryl group, linked via an energy-rich bond in 1,3-bisphosphoglycerate, to ADP constitutes the second reaction of stage II. Cleavage of the energy-rich bond is energetically coupled to phosphorylation of ADP. **Phosphoglycerate kinase** catalyzes this reaction, and 1,3-bisphosphoglycerate provides both the energy and the phosphoryl group for phosphorylation of ADP to ATP:



1, 3-Bisphospho- + ADP^{3-} ⇄ 3-phospho- + ATP^{4-}
glycerate glycerate

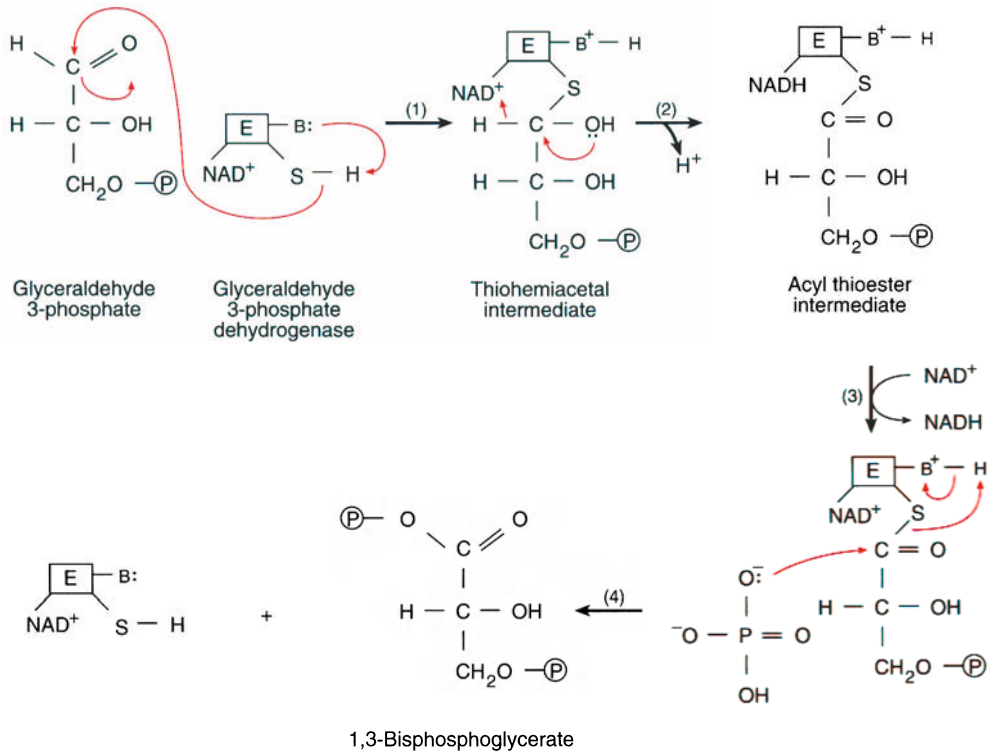
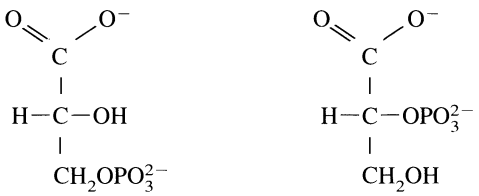


Figure 10.10. The glyceraldehyde 3-phosphate dehydrogenase reaction. (1) A thiohemiacetal forms between glyceraldehyde 3-phosphate and a reactive SH group of the enzyme (B: is a basic group); (2) NAD^+ oxidation yields a thioester plus NADH and H^+; (3) NAD^+ replaces NADH on the enzyme; (4) phosphate attack produces 1,3-bisphosphoglycerate and restores the enzyme to its initial state.

ATP synthesis, catalyzed by phosphoglycerate kinase, illustrates what we term *substrate-level phosphorylation*—phosphorylation of ADP to ATP coupled to cleavage of an energy-rich bond in a metabolite. Substrate-level phosphorylation constitutes one mechanism for ATP synthesis; you will encounter two other mechanisms—*oxidative phosphorylation* and *photosynthetic phosphorylation*—in later chapters. Mammalian phosphoglycerate kinase is a monomer (MW = 64,000) and requires Mg^{2+} as cofactor.

10.2.2C. Phosphoglyceromutase.

An isomerization reaction, catalyzed by **phosphoglyceromutase,** is the next reaction of stage II. This reaction proceeds via the intermediate 2,3-bisphosphoglycerate:

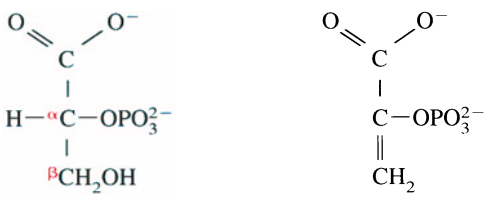


3-Phosphoglycerate ⇄ 2-phosphoglycerate

A **mutase** catalyzes the transfer of a functional group from one position to another on the same molecule. The mammalian enzyme is a dimer (MW = 27,000/subunit) and requires Mg^{2+} as cofactor.

10.2.2D. Enolase.

The next reaction, catalyzed by **enolase,** forms 2-phosphoenolpyruvate (phosphoenolpyruvate for short) by dehydration of 2-phosphoglycerate. Chemically, this constitutes an α,β-elimination reaction in which H_2O forms from substituents of the α and β carbons:



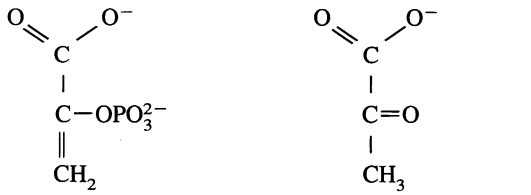
2-Phosphoglycerate ⇄ phosphoenolpyruvate + H_2O

Mammalian enolase is a dimer (MW = 41,000/subunit) and requires Mg^{2+} as cofactor. Fluoride (F^-) is a potent inhibitor of the enzyme, presumably owing to forma-

tion of ionic complexes between enzyme-bound Mg^{2+} and fluorophosphate ions (FPO_3^{3-}).

Phosphoenolpyruvate, aptly abbreviated as *PEP,* is the second energy-rich compound formed in stage II. Its energy-rich character comes from being the ester of an enol rather than of an ordinary alcohol. Cleavage of the energy-rich bond linking the phosphoryl group to C(2) is strongly exergonic because the free energy change includes that due to ester bond hydrolysis and that due to the highly favored shift from the enol to the keto form.

10.2.2E. Pyruvate Kinase. The last reaction of stage II, catalyzed by **pyruvate kinase,** accomplishes the transfer of the phosphoryl group from phosphoenolpyruvate to ADP:



Phosphoenolpyruvate + ADP^{3-} + H^+ → pyruvate + ATP^{4-}

This reaction, like that catalyzed by phosphoglycerate kinase, constitutes a *substrate-level phosphorylation.* Phosphoenolpyruvate provides both the energy and the phosphoryl group for synthesizing ATP.

Mammalian pyruvate kinase is a tetramer of four identical subunits (MW = 57,000/subunit). Several isozymes are known. One type of isozyme, designated L_4, occurs in the liver; another, designated M_4, is found in muscles. The L_4 isozyme is allosterically inhibited by ATP, acetyl coenzyme A, and fatty acids; it is activated by fructose 1,6-bisphosphate. Pyruvate kinase also undergoes covalent modification (under hormonal control), consisting of phosphorylation and dephosphorylation. Dephosphorylated pyruvate kinase represents the active form. Pyruvate kinase requires K^+ and either Mg^{2+} or Mn^{2+} for activity.

Summing up the 10 reactions of stages I and II yields the following overall reaction for glycolysis:

$$\text{Glucose} + 2NAD^+ + 2ADP^{3-} + 2P_i^{2-} \rightarrow 2\ \text{pyruvate}^- + 2ATP^{4-} + 2NADH + 2H^+$$

10.3. GLYCOLYSIS—END PRODUCTS, ENERGETICS, AND CONTROL

10.3.1. Aerobic and Anaerobic Conditions

The glycolytic sequence from glucose to pyruvate constitutes an anaerobic set of reactions; none of the steps requires oxygen. However, glycolysis can occur when the system as a whole is either aerobic or anaerobic (hence the terms aerobic and anaerobic glycolysis). The sequence glucose → pyruvate is unchanged whether oxygen is present or absent, but the metabolic fate of pyruvate depends on whether the system as a whole is aerobic or anaerobic (Figure 10.11).

10.3.1A. Acetyl Coenzyme A and Lactate. Under aerobic conditions, the *pyruvate dehydrogenase complex* (discussed in Section 11.2) catalyzes the conversion of pyruvate to *acetyl coenzyme A* (see Figure 11.7), which enters the *citric acid cycle.* The cycle, in conjunction with the *electron transport system,* leads to complete oxidation of the acetyl group of acetyl coenzyme A to CO_2 and H_2O. Thus, starting with glucose, and using the combined metabolic systems of glycolysis, the citric acid cycle, and the electron transport system, the cell accomplishes *complete oxidation of carbohydrate to CO_2 and H_2O.*

Recall that the first reaction of stage II results in reduction of NAD^+ to NADH. It is essential to regenerate NAD^+ from NADH so that subsequent substrates can be oxidized. If NADH is not reoxidized—or more significantly, if NAD^+ is not regenerated—glycolysis stops owing to lack of NAD^+ for glyceraldehyde 3-phosphate dehydrogenase. Under aerobic conditions, the electron transport system accomplishes the reoxidation of NADH by means of molecular oxygen (more in Section 12.3).

Under anaerobic conditions, pyruvate does not form acetyl coenzyme A, and NADH cannot be reoxidized via the electron transport system. Instead, regeneration of NAD^+ is accomplished by coupling the oxidation of NADH to NAD^+ with the reduction of pyruvate to *lactate* (CH_3-CHOH-COO^-) in a reaction catalyzed by *lactate dehydrogenase* (*LDH;* a tetramer; MW = 33,500/subunit). Therefore, under anaerobic conditions, glucose catabolism produces lactate rather than pyruvate:

$$\text{Pyruvate} + NADH + H^+ \rightleftarrows NAD^+ + \text{lactate}$$
$$\Delta G^{\circ\prime} = -25.1\ \text{kJ mol}^{-1}$$

An aerobic organism cannot survive in a strictly anaerobic environment—the complete absence of oxygen—but it may be able to tolerate a *temporary oxygen shortage.* In humans, strenuous exercise rapidly depletes oxygen by operation of the electron transport system. Under such relatively anaerobic conditions, pyruvate is converted to lactate. The conversion occurs in muscle tissues, and lactate accumulates there first. From the muscles, lactate diffuses into the bloodstream (see Figure 10.1), which transports it to the liver, where it is reconverted to glucose.

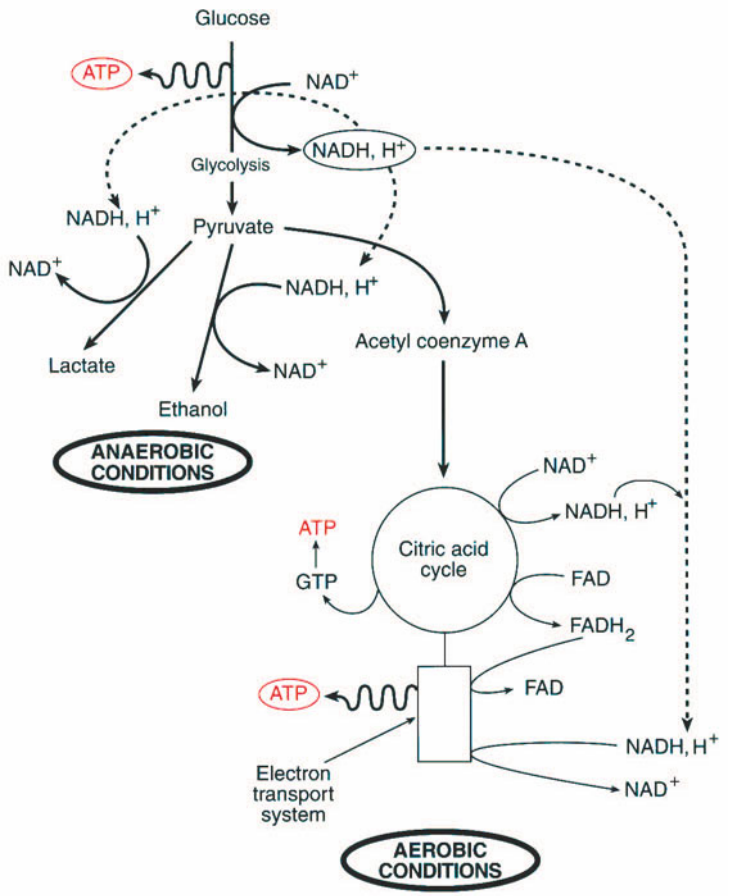


Figure 10.11. Glycolysis under aerobic and anaerobic conditions.

10.3.1B. The Cori Cycle. We refer to the cyclic set of reactions just described as the **Cori cycle** (Figure 10.12), after Carl and Gerti Cori, who first described it. The cycle involves the interplay of muscle, liver, and blood and begins with conversion of glucose to lactate in muscle tissue through glycolysis. The blood then transports lactate to the liver, where it is reoxidized to pyruvate. Pyruvate is converted to glucose via *gluconeogenesis* (discussed in Section 10.6), using liver ATP. Lastly, the blood transports glucose back to muscle tissue.

The Cori cycle has high activity during periods of strenuous exercise. The extra liver ATP, used during gluconeogenesis to process accumulated lactate, must be regenerated by *oxidative phosphorylation* (covered in Section 12.4). Increased oxidative phosphorylation requires extra oxygen, more than that needed for normal operation of the electron transport system. The additional oxygen requirement accounts for the shortness of breath, the panting, you experience with strenuous exercise. We refer to the oxygen required to regenerate used-up liver ATP as *oxygen debt.* It may take a long time before your oxygen consumption rate returns to normal.

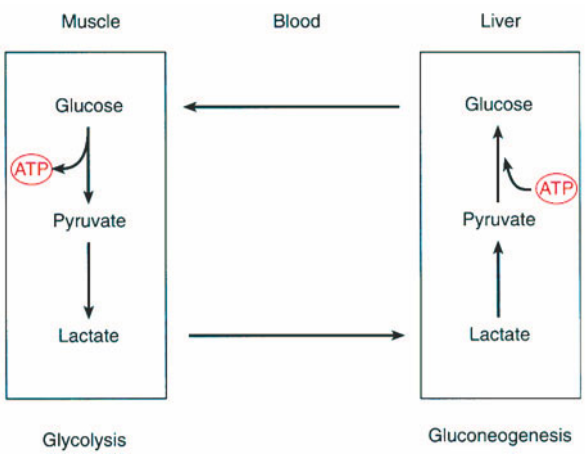


Figure 10.12. The Cori cycle. Muscle glycolysis converts glucose to lactate, which is carried by the blood to the liver, where it is reoxidized to pyruvate. Gluconeogenesis converts pyruvate to glucose, using liver ATP as a source of energy, and the blood transports the glucose back to muscle tissue.

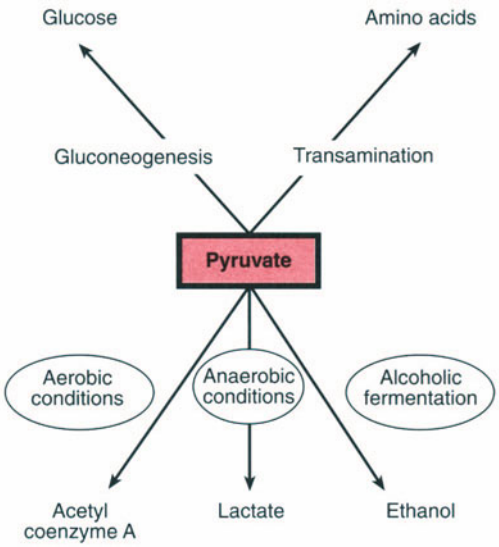


Figure 10.13. Metabolic fates of pyruvate.

The Cori cycle is linked to both glycogen breakdown, *glycogenolysis,* and glycogen synthesis, *glycogenesis.* In muscles, glycogenolysis yields glucose that can be converted to lactate via glycolysis. In the liver, glucose formed by gluconeogenesis can be used for glycogenesis.

10.3.2. Metabolic Fates of Pyruvate

Glycolysis is best defined as the *sequence of reactions from glucose to pyruvate.* Some authors include the pyruvate-to-lactate step in their definition, but the lactate dehydrogenase reaction occurs only under certain conditions. The sequence from glucose to pyruvate, on the other hand, is common to all glycolytic systems.

Pyruvate has several metabolic fates (Figure 10.13). As you saw, conversion of pyruvate to acetyl coenzyme A or to lactate depends on whether the system as a whole is relatively aerobic or anaerobic. However, it also depends on the type of tissue involved and its content of lactate dehydrogenase isozymes.

Consider cardiac muscle and skeletal muscle, which contain, respectively, the H_4 and M_4 isozymes of lactate dehydrogenase (Table 10.2). Kinetic properties of these isozymes are such that cardiac muscle tissue favors conversion of pyruvate to acetyl coenzyme A whereas skeletal muscle tissue favors conversion of pyruvate to lactate. The heart, the most critical organ of animals, is programmed to derive maximum energy from glucose by linking glycolysis to the aerobic citric acid cycle/electron transport system, which yields a large amount of ATP. Skeletal muscle, on the other hand, can function adequately by deriving a smaller amount of energy from glucose by linking glycolysis to the anaerobic production of lactate.

Pyruvate also serves as the initial substrate for *gluconeogenesis.* In yeast, pyruvate is a substrate for *alcoholic fermentation,* the reaction sequence that converts glucose to ethanol. During alcoholic fermentation, *pyruvate decarboxylase* catalyzes the conversion of pyruvate to acetaldehyde (CH_3-CHO). The acetaldehyde is then reduced to ethanol by means of NADH produced in the glyceraldehyde 3-phosphate dehydrogenase reaction. In the process, NADH is oxidized to NAD^+:

$$\text{Pyruvate} + H^+ \xrightarrow{\text{pyruvate decarboxylase}} CO_2 + \text{acetaldehyde}$$

$$\text{Acetaldehyde} + \text{NADH} + H^+ \xrightarrow{\text{alcohol dehydrogenase}} \text{ethanol} + NAD^+$$

The *alcohol dehydrogenase* reaction also constitutes the first step of alcohol metabolism in the liver, except that under those conditions it proceeds in the reverse direction, from ethanol to acetaldehyde. Some of the effects associated with alcohol intake are directly related to aspects of this reaction, namely, the lowering of pH (production of H^+) and the formation of an aldehyde (acetaldehyde).

10.3.3. Energetics

Table 10.1 lists the free energy changes for the 10 reactions of glycolysis. You can see that in all cases there exist significant differences between biochemical standard

Table 10.2. Properties of Lactate Dehydrogenase Isozymes

Property	H_4 isozyme	M_4 isozyme
K_m for pyruvate	High	Low
k_{cat} (for pyruvate → lactate)	Low	High
Inhibition by pyruvate (of lactate → pyruvate)	Strong	Weak
Fate of pyruvate (Pyr)	Pyr → acetyl coenzyme A	Pyr → Lactate
Metabolism	Aerobic	Anaerobic

and actual free energy changes. Remember that only biochemical actual free energy changes truly describe the tendency of reactions to proceed under intracellular conditions. On that basis, three reactions (those catalyzed by hexokinase, phosphofructokinase, and pyruvate kinase) are strongly exergonic and, therefore, irreversible. These reactions make the entire glycolytic sequence from glucose to pyruvate a "downhill" (exergonic) process, a unidirectional pathway with an overall $\Delta G'$ of -76.6 kJ mol^{-1}.

Stage I of glycolysis requires the input of two molecules of ATP (Figure 10.14). Stage II yields two molecules of ATP. At first glance, it may appear that the *net* production of ATP equals zero, but this is incorrect. Every molecule of glucose ultimately gives rise to *two* molecules of glyceraldehyde 3-phosphate. Processing of these in stage II results in formation of *two* molecules of pyruvate and a total production of *four* molecules of ATP. Thus, there occurs a *net gain* of *two* molecules of ATP per molecule of glucose converted to pyruvate.

Under anaerobic conditions, lactate dehydrogenase catalyzes the oxidation of the NADH formed in glycolysis to NAD^+. This reaction does not yield any additional energy. Hence, *the maximum energy yield from glycolysis under anaerobic conditions is two molecules of ATP per molecule of glucose* (for the sequence glucose → 2 lactate).

The energy yield of glycolysis under anaerobic conditions is low because glucose undergoes only *limited oxidation* in forming pyruvate. As you will see later (Section 12.6), when glycolysis becomes linked to the citric acid cycle under aerobic conditions, glucose undergoes *complete oxidation* to CO_2 and H_2O. The more extensive the oxidation of a metabolite is, the larger the number of hydrogens abstracted from it and the greater the energy yield. You can compare the oxidation states of pyruvate and glucose by computing their C:H:O ratios—1:2:1 for glucose ($C_6H_{12}O_6$) and 1:1.33:1 for pyruvic acid ($C_3H_4O_3$). Thus, pyruvate is only slightly more oxidized (less reduced) than glucose. Because pyruvate production from glucose captures only a small amount of energy in the form of ATP, you might ask whether glycolysis constitutes a wasteful pathway. The answer is that, on the contrary, glycolysis represents a well-engineered process that succeeds in extracting a limited amount of energy from glucose without oxidizing it appreciably.

You can calculate the *efficiency of energy conservation* attained by anaerobic glycolysis as follows. Conversion of glucose → 2 lactate has a theoretical $\Delta G^{\circ\prime}$ of -196.6 kJ mol^{-1} and results in formation of 2 ATP. Since synthesis of ATP requires 30.5 kJ mol^{-1}, a total of 61.0 kJ mol^{-1} of glucose catabolism is conserved in the form of ATP synthesis. This represents an efficiency of

$$\frac{61.0 \text{ kJ mol}^{-1}}{196.6 \text{ kJ mol}^{-1}} \times 100 = 31\%$$

Keep in mind, though, that this represents an approximate calculation, based on $\Delta G^{\circ\prime}$ values. Calculating actual efficiencies must be based on pertinent intracellular concentrations of reactants and products and use $\Delta G'$ values. Thus, the true efficiency of energy conservation in any given cell may be higher or lower than 31%.

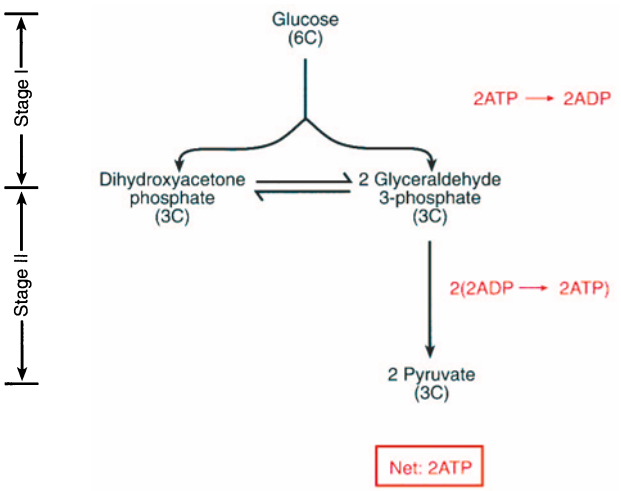


Figure 10.14. The energetics of glycolysis. Stage I requires the input of 2 ATP per molecule of glucose; stage II generates 2 ATP per molecule of glyceraldehyde 3-phosphate. Since one molecule of glucose yields two molecules of glyceraldehyde 3-phosphate, the net yield is 2 ATP per molecule of glucose.

10.3.4. Pasteur and Crabtree Effects

Scientists have recognized two general control mechanisms of glycolysis and related metabolic pathways for many years. The mechanisms are known, after their discoverers, as the *Pasteur effect* and the *Crabtree effect.* We can understand both effects qualitatively by considering the complete catabolism of glucose, which can be divided into two parts (Figure 10.15). The first part comprises the conversion of glucose to pyruvate via glycolysis. These reactions require no oxygen and yield only a *small amount of ATP.* The second part consists of the conversion of pyruvate to acetyl coenzyme A, which enters the citric acid cycle. Operation of the citric acid cycle in conjunction with the electron transport system requires aerobic conditions and yields a *large amount of ATP.*

Now consider the **Pasteur effect.** Louis Pasteur observed more than 100 years ago (1861) that when yeast

cells were exposed to aerobic conditions, their glucose consumption and ethanol production decreased sharply. The more aerobic the system was, the greater the decrease in glycolysis. The Pasteur effect refers to the *inhibition of glycolysis by oxygen.*

Under aerobic conditions, the citric acid cycle/electron transport system part of glucose catabolism functions well and produces large amounts of energy in the form of ATP. Accordingly, no need exists to process every available glucose molecule through glycolysis in order to derive the maximum amount of energy possible from glucose catabolism. The greater is the activity of the citric acid cycle/electron transport system, the fewer glucose molecules need to be processed via glycolysis. Hence, the more aerobic the system, the greater is the inhibition of glycolysis. We now know the molecular basis of the Pasteur effect. Increased aerobic conditions result in increased production of ATP, and ATP strongly inhibits phosphofructokinase, thereby slowing down glycolysis.

In some ways, the Crabtree effect constitutes the opposite of the Pasteur effect. The **Crabtree effect** refers to the *inhibition of oxygen consumption produced by increasing concentrations of glucose.* When the glucose concentration is high, large amounts of it are catabolized via glycolysis. This results in appreciable production of energy in the form of ATP despite the low energy yield of the pathway. Under these conditions, no need exists to process every pyruvate molecule formed in glycolysis by means of the citric acid cycle/electron transport system to produce still much larger amounts of ATP. The greater the glucose concentration, the lower is the activity of the citric acid cycle/electron transport system part of the combined pathway. Hence, increasing glucose concentrations inhibit oxygen consumption. We do not yet know the molecular basis of the Crabtree effect.

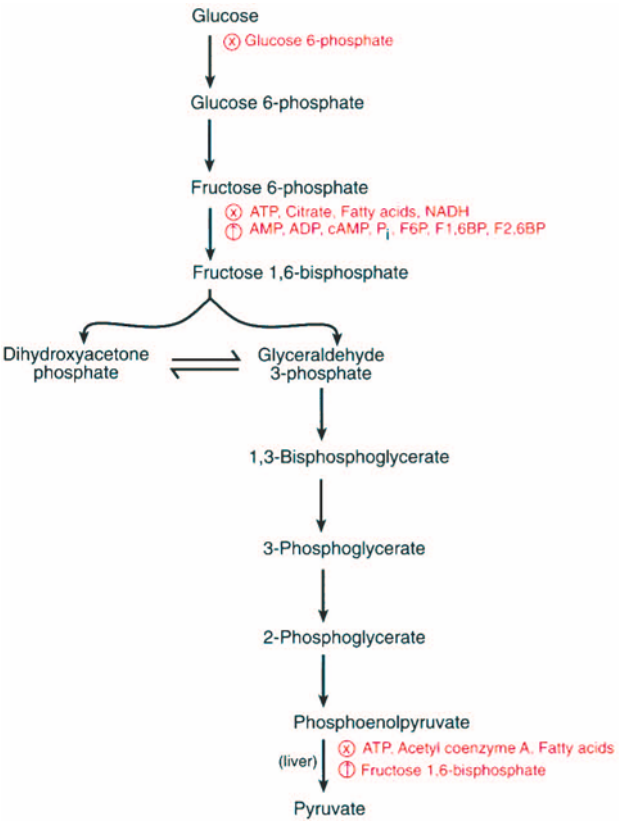


Figure 10.16. Regulation of glycolysis. ⊗ Inhibitor, ⊕ activator. F6P, F1,6BP, and F2,6BP designate fructose 6-phosphate, fructose 1,6-bisphosphate, and fructose 2,6-bisphosphate, respectively.

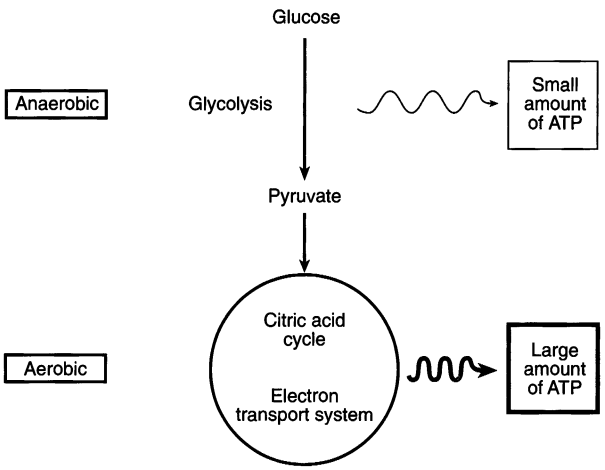


Figure 10.15. Basis of the Pasteur and Crabtree effects. Carbohydrate metabolism consists of an anaerobic segment having a low energy yield (glycolysis) and an aerobic segment having a high energy yield (citric acid cycle/electron transport system). Extensive activity in one segment depresses activity in the other.

Control of glycolysis itself occurs at the three highly exergonic, irreversible steps involving hexokinase, phosphofructokinase, and pyruvate kinase (Figure 10.16). Control occurs largely via allosteric enzyme effectors, including adenine nucleotides, NADH, and phosphorylated monosaccharides. Low cellular ATP concentration and high AMP and ADP concentrations stimulate glycolysis and increase production of needed ATP. By contrast, high cellular levels of ATP and NADH depress glycolysis and prevent synthesis of unnecessary ATP.

10.3.5. Catabolism of Other Carbohydrates

Glycolysis provides a catabolic pathway not just for glucose but also for other carbohydrates. Important

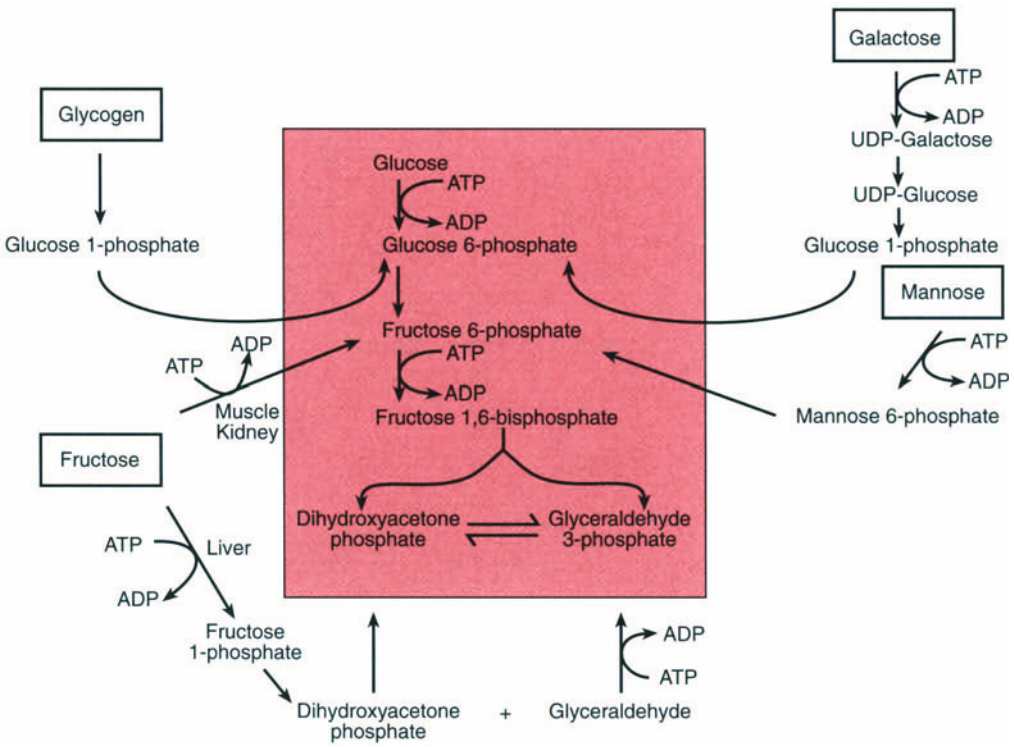


Figure 10.17. Catabolism of various carbohydrates. Many form glycolytic intermediates and enter the sequence at those points.

nonglucose carbohydrates include glycogen, disaccharides, and monosaccharides. We will discuss glycogen catabolism in Section 10.5. Disaccharides are hydrolyzed to their component monosaccharides in the intestinal tract. Many monosaccharides form glycolytic intermediates (Figure 10.17). After entering glycolysis at specific points in the pathway, these intermediates are catabolized by the remaining portion of the pathway.

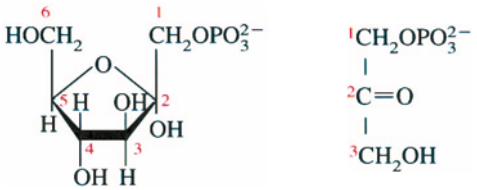
10.3.5A. Fructose. Fructose enters glycolysis via two routes. In muscles and kidneys, it undergoes phosphorylation to fructose 6-phosphate in a reaction catalyzed by hexokinase:

$$\text{Fructose} + \text{ATP}^{4-} \rightarrow \text{fructose 6-phosphate}^{2-} + \text{ADP}^{3-} + \text{H}^{+}$$

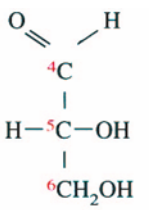
In the liver, fructose is phosphorylated at C(1) to form fructose 1-phosphate in a reaction catalyzed by *fructokinase:*

$$\text{Fructose} + \text{ATP}^{4-} \rightarrow \text{fructose 1-phosphate}^{2-} + \text{ADP}^{3-} + \text{H}^{+}$$

Aldolase catalyzes the cleavage of fructose 1-phosphate to dihydroxyacetone phosphate and glyceraldehyde:



Fructose 1-phosphate ⇄ dihydroxyacetone phosphate +



glyceraldehyde

Dihydroxyacetone phosphate can enter glycolysis directly. Glyceraldehyde must first be phosphorylated to glyceraldehyde 3-phosphate in a reaction catalyzed by *triose kinase:*

$$\text{Glyceraldehyde} + \text{ATP}^{4-} \rightarrow \text{glyceraldehyde 3-phosphate}^{2-} + \text{ADP}^{3-} + \text{H}^{+}$$

Each of the two routes by which fructose enters glycolysis requires the input of 2 ATP in stage I. Hence, con-

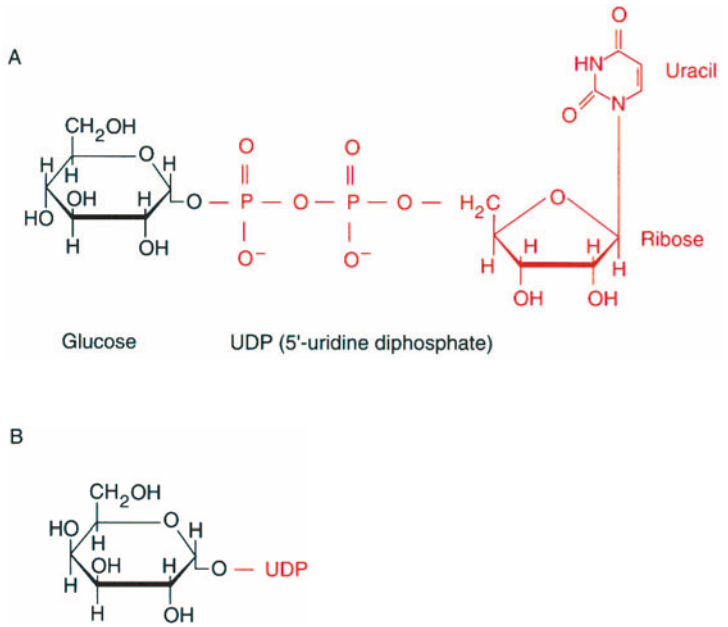


Figure 10.18. (A) UDP-glucose; (B) UDP-galactose.

version of fructose to pyruvate via the combined reactions of stages I and II yields 2 ATP per molecule of fructose for either route. This means that, in terms of energy yield, fructose is equivalent to glucose. However, fructose catabolism in the liver bypasses the phosphofructokinase step with its multiple regulatory factors (see Figure 10.16). Consequently, diets rich in fructose or sucrose can lead to overproduction of pyruvate. Pyruvate, in turn, forms acetyl coenzyme A, a precursor for cholesterol and fatty acid biosynthesis. You can see that in terms of metabolic consequences, fructose is not equivalent to glucose, and high-fructose diets may lead to undesirable nutritional effects for humans.

10.3.5B. Mannose. Hexokinase catalyzes the phosphorylation of mannose to mannose 6-phosphate. *Phosphomannoisomerase* catalyzes the isomerization of mannose 6-phosphate to fructose 6-phosphate:

$$\text{Mannose} + \text{ATP}^{4-} \rightarrow \text{mannose 6-phosphate}^{2-} + \text{ADP}^{3-} + \text{H}^{+}$$

$$\text{Mannose 6-phosphate}^{2-} \rightleftarrows \text{fructose 6-phosphate}^{2-}$$

10.3.5C. Galactose. The first step in galactose catabolism also involves a phosphorylation, catalyzed by *galactokinase:*

$$\text{Galactose} + \text{ATP}^{4-} \rightarrow \text{galactose 1-phosphate}^{2-} + \text{ADP}^{3-} + \text{H}^{+}$$

Subsequent steps involve *nucleoside diphosphates* that function as carriers of hexose groups. In these steps, galactose 1-phosphate first reacts with *uridine diphos-*

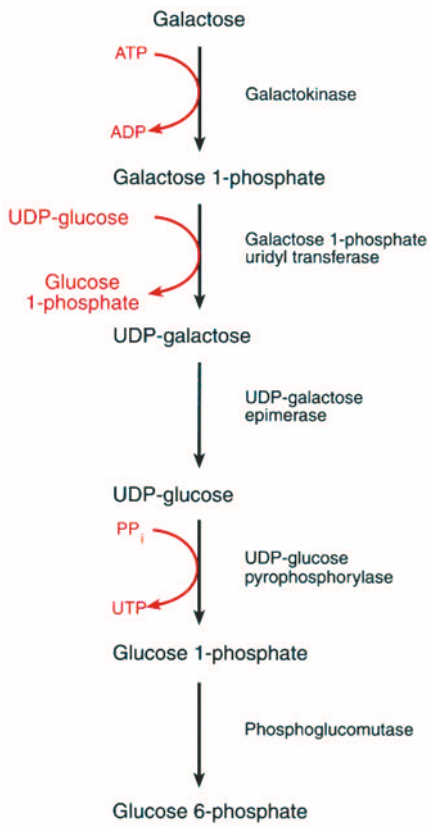


Figure 10.19. Pathway for the utilization of galactose.

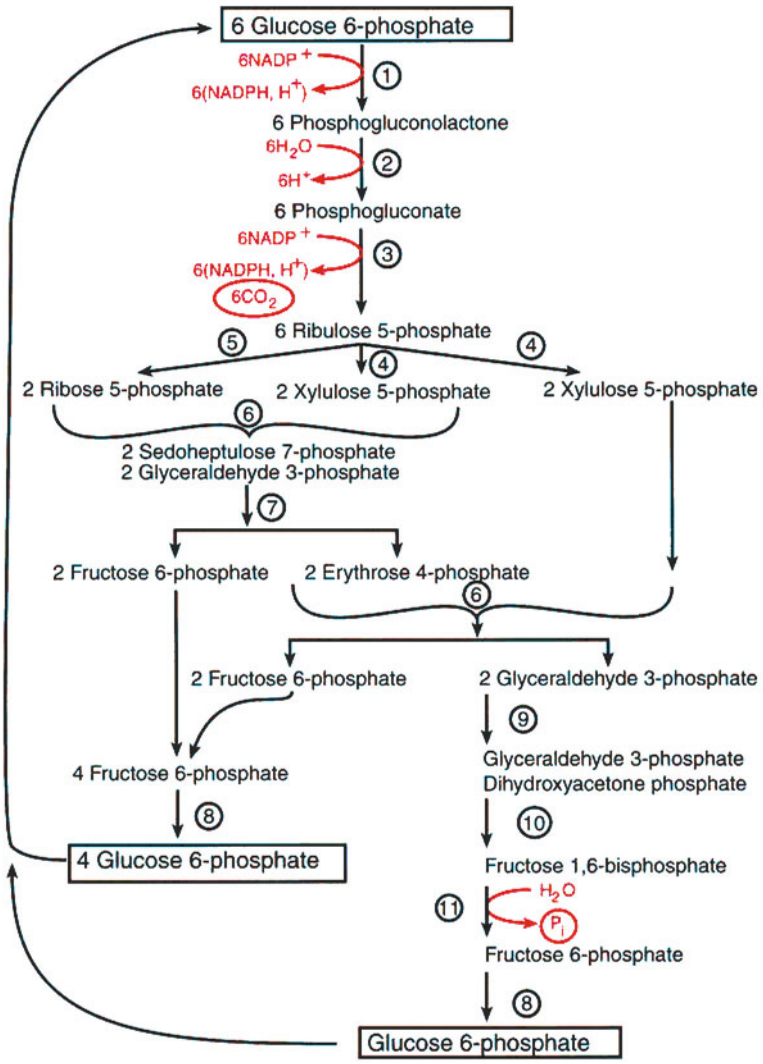


Figure 10.20. The pentose phosphate pathway. Six molecules of glucose 6-phosphate enter the pathway, and five are regenerated. The cycle accomplishes the equivalent of degrading one glucose 6-phosphate to 6 CO_2 and P_i and produces 12 NADPH and 12 H^+. Numbers designate enzymes. ① Glucose 6-phosphate dehydrogenase; ② 6-phosphogluconolactonase; ③ 6-phosphogluconate dehydrogenase; ④ phosphopentose epimerase; ⑤ phosphopentose isomerase; ⑥ transketolase; ⑦ transaldolase; ⑧ phosphoglucoisomerase; ⑨ triose-phosphate isomerase; ⑩ aldolase; ⑪ fructose 1,6-bisphosphatase.

phate glucose or *UDP-glucose* (Figure 10.18) to form *UDP-galactose*. UDP-galactose then forms UDP-glucose and, ultimately, glucose 6-phosphate (Figure 10.19). The pathway is of particular interest in humans because of the occurrence of a genetic disease, **galactosemia,** characterized by inability to convert galactose to glucose metabolites. A deficiency of galactokinase leads to a mild disorder, but a deficiency of *galactose 1-phosphate uridyl transferase* results in severe symptoms that may include lack of growth, liver failure, and mental retardation.

10.4. PENTOSE PHOSPHATE PATHWAY

The *pentose phosphate pathway*—a second major catabolic pathway of carbohydrates—represents an alternate mechanism for glucose degradation. The pathway occurs in the cytoplasm of animal, plant, and bacterial cells. Figure 10.20 shows an overview of the reaction sequence. The pentose phosphate pathway comprises a cyclic set of reactions that effects the oxidation of glucose 6-phosphate to CO_2 and P_i and leads to the production of large amounts of *reduced nicotinamide adenine dinucleotide phosphate* or *NADPH* (see Figure 11.3).

10.4.1. Oxidative Phase

We divide the reactions of the pentose phosphate pathway into two phases. The first of these, the *oxidative phase,* accomplishes the decarboxylation of glucose (six carbons) to *ribulose* (five carbons) and generates reducing power in the form of NADPH. This phase (Figure 10.21) consists of three reactions:

1 Oxidation of glucose 6-phosphate by means of *glucose 6-phosphate dehydrogenase* to *6-phosphogluconolactone,* an intramolecular ester formed between C(1) and C(5).

2 Hydrolysis of the lactone to *6-phosphogluconate* by means of *6-phosphogluconolactonase.*
3 Oxidative decarboxylation of 6-phosphogluconate, catalyzed by *6-phosphogluconate dehydrogenase,* to *ribulose 5-phosphate.*

Note that in order for glucose to be metabolized via the pentose phosphate pathway, it must first be phosphorylated to glucose 6-phosphate. The phosphorylation requires an input of ATP and can be catalyzed by hexokinase or glucokinase.

10.4.2. Nonoxidative Phase

The remaining reactions of the pentose phosphate pathway consist of interconversions of carbohydrates by reshuffling of their carbon skeletons. Two key enzymes of this phase, *transaldolase* and *transketolase,* are similar in their substrate specificities. Both transfer a carbon fragment from a ketose, serving as donor, to an aldose, serving as acceptor. The enzymes differ in the size of fragment transferred (Figure 10.22).

Transaldolase catalyzes the transfer of a *three-carbon* fragment (a *dihydroxyacetone group*) from a 2-keto sugar to C(1) of an aldose, thereby forming a new ketose and a new aldose. The mechanism of action resembles that of aldolase in glycolysis.

Transketolase catalyzes the transfer of a *two-carbon* fragment (a *glycolaldehyde group*) from a 2-keto sugar to C(1) of an aldose, thereby producing a new ketose and a new aldose. Transketolase requires Mg^{2+} and *thiamine pyrophosphate* (*TPP*) as cofactors. (See Figure 11.9 for the structure of TPP.) Figure 10.23 shows the reshuffling of carbon skeletons in the pentose phosphate pathway in schematic fashion.

In the operation of the pentose phosphate pathway, *six* molecules of glucose 6-phosphate enter the cycle. Each glucose 6-phosphate loses one carbon as CO_2. Ultimately, five molecules of glucose 6-phosphate are regenerated:

Overall reaction:

$$6\ \text{Glucose 6-phosphate}^{2-} + 7H_2O + 12NADP^+ \rightarrow 6CO_2 + 5\ \text{glucose 6-phosphate}^{2-} + 12NADPH + 12H^+ + P_i^{2-}$$

Consequently, what the pathway accomplishes is *equivalent* to the complete oxidation of *one* molecule of glucose 6-phosphate to six molecules of CO_2 and P_i:

Net reaction:

$$\text{Glucose 6-phosphate}^{2-} + 7H_2O + 12NADP^+ \rightarrow 6CO_2 + 12NADPH + 12H^+ + P_i^{2-}$$

10.4.3. Functions of the Pathway

The pentose phosphate pathway has three main functions:

1 It generates reducing power in the form of NADPH. Note that NADH and NADPH have different roles and *cannot be interchanged* in metabolism. NADH functions in the electron transport system to generate energy in the form of ATP. NADPH serves as a reducing agent for metabolic systems requiring reduction, such as photosynthesis, fatty acid biosynthesis, and conversion of ribonucleotides to deoxyribonucleotides.
2 It is a source of carbohydrates of different structures (three-, four-, five-, six-, and seven-carbon skeletons).
3 It provides a means for interconversions of different monosaccharides.

Pentose phosphate pathway activity is low in muscle but high in adipose tissue, where the cycle generates re-

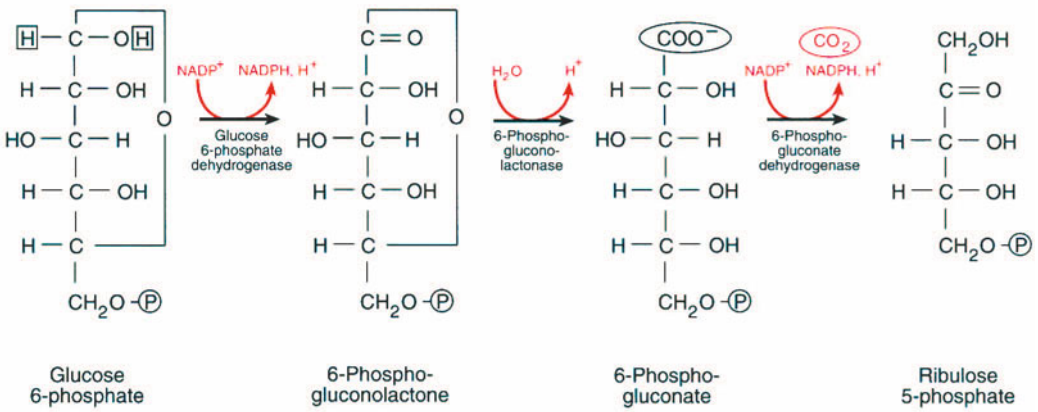


Figure 10.21. Oxidative phase of the pentose phosphate pathway.

ducing power for fatty acid biosynthesis. Pathway activity is also high in vertebrate red blood cells, where the cycle ensures a supply of reduced glutathione (GSH) formed by action of *glutathione reductase:*

$$GSSG + NADPH + H^+ \xrightarrow{\text{glutathione reductase}} 2GSH + NADP^+$$

Red blood cells require reduced glutathione in large amounts to (a) maintain the structural integrity of proteins by shielding their —SH groups; (b) protect membrane lipids against oxidation by peroxides; and (c) maintain the iron of hemoglobin in its divalent state (Fe^{2+}). Decreased levels of NADPH may result in alteration of proteins, peroxidation of lipids, and formation of methemoglobin (Fe^{3+}), all changes that weaken red blood cell membranes and render the cells sensitive to rupture (*hemolysis*).

Some individuals have a genetic defect characterized by a glucose 6-phosphate dehydrogenase deficiency. They have low concentrations of NADPH in their red blood cells, are prone to severe cases of anemia, and are sensitive to certain seemingly harmless drugs, including the antimalarial drug *primaquine.* The drugs oxidize some of the available NADPH, thereby putting added strain on cells that have a low concentration of NADPH to start with. Because of the decreased concentration of NADPH, affected

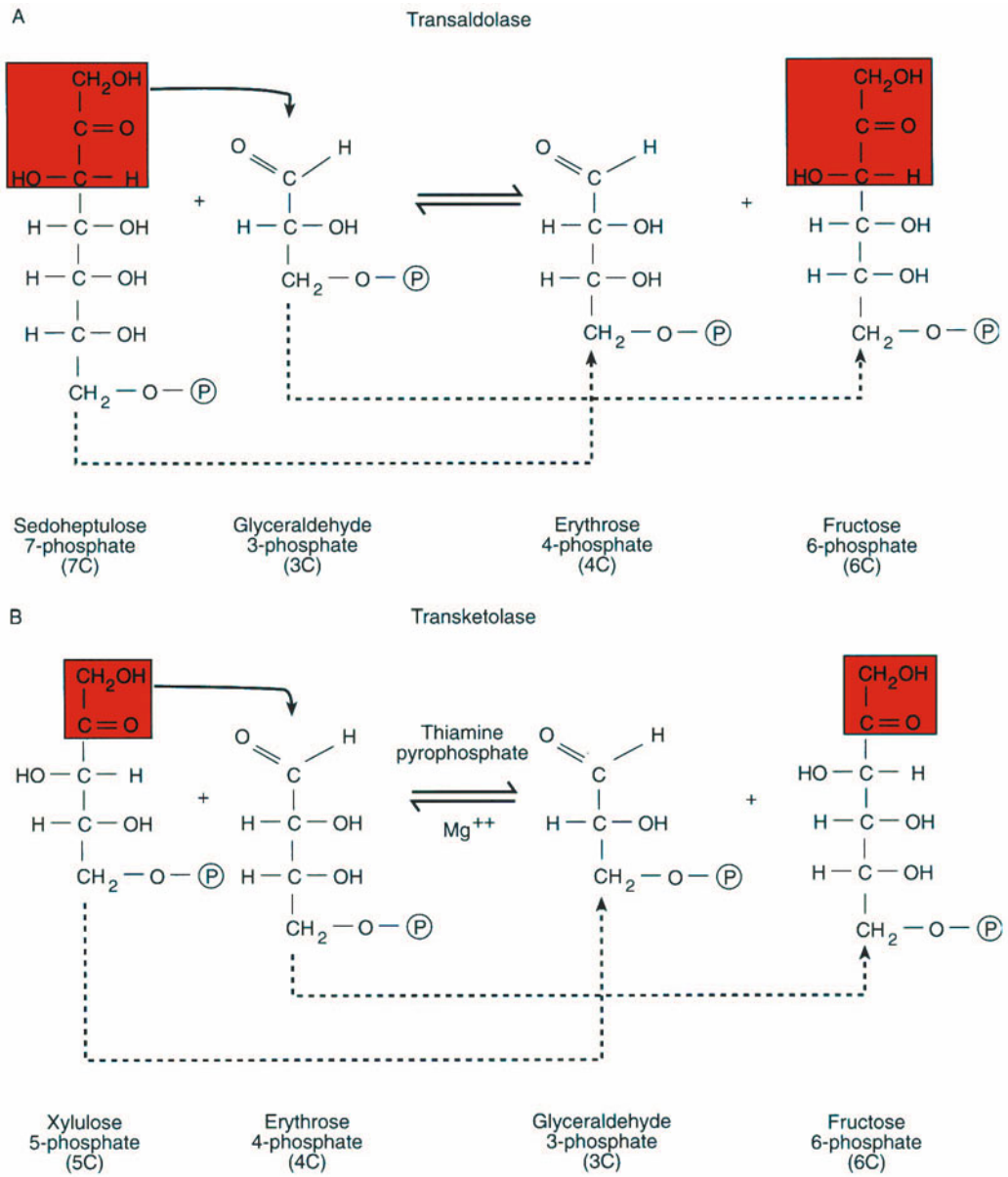


Figure 10.22. Illustrative reactions of transaldolase (A) and transketolase (B).

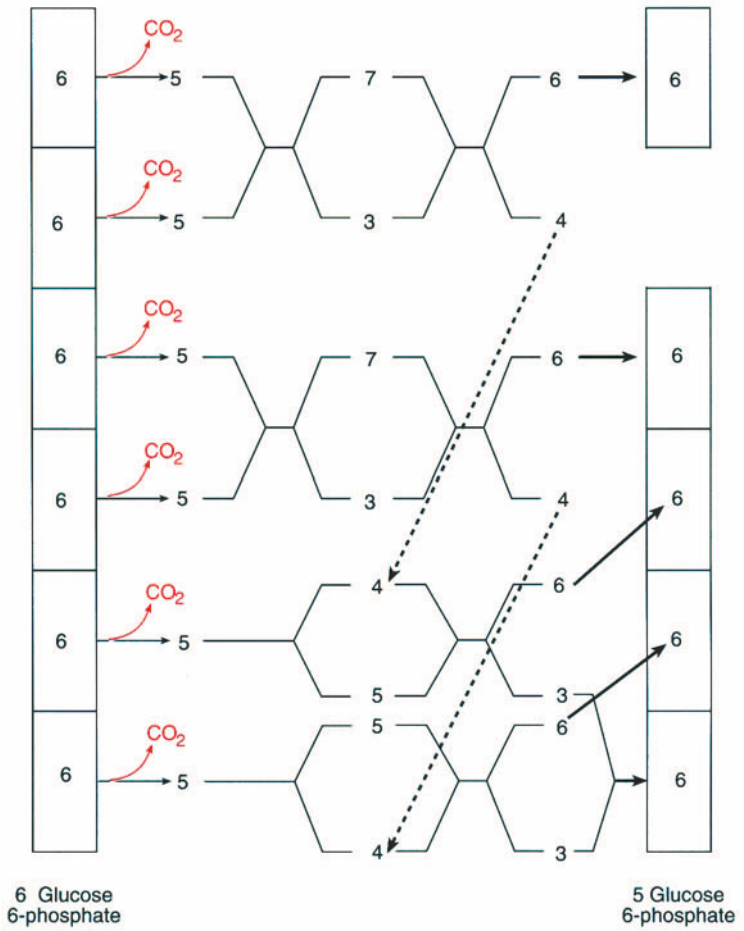


Figure 10.23. Schematic diagram of the rearrangements of carbon skeletons in the pentose phosphate pathway. Six molecules of glucose 6-phosphate enter the pathway on the left and regenerate five molecules on the right. Numbers indicate the number of carbon atoms per molecule.

cells rupture easily, and administration of such "harmless" drugs may result in *hemolytic anemia*—an extensive anemia that includes massive destruction of red blood cells.

10.5. GLYCOGEN DEGRADATION AND SYNTHESIS

10.5.1. Glycogenolysis

Glycogen, the storage form of carbohydrates in animals, must be broken down to glucose before it can be used for generation of energy. The same holds for starch, the storage carbohydrate of plants. Degradation of both polysaccharides involves stepwise reactions in which the chains are shortened by removal of one glucose residue at a time. **Glycogen phosphorylase,** or **phosphorylase** for short, catalyzes the degradation of glycogen (*glycogenolysis*); a similar enzyme, *starch phosphorylase,* catalyzes the degradation of starch.

10.5.1A. Glycogen Phosphorylase. In the phosphorylase reaction, glycogen undergoes cleavage at its *nonreducing end,* and the terminal glucose residue is released as *glucose 1-phosphate* (Figure 10.24). Phosphorylation of glucose occurs without expenditure of ATP. In this respect, glycogenolysis differs from glycolysis and the pentose phosphate pathway, both of which begin with an ATP-dependent phosphorylation of glucose to glucose 6-phosphate.

We refer to the type of reaction catalyzed by phosphorylase as **phosphorolysis.** Generally speaking, phosphorolysis represents a reaction in which a covalent bond is cleaved and the elements of phosphoric acid are added across the bond (see also Figure 10.10). If phosphoric acid is in the form of H_3PO_4, then one of the products combines with H^+ and the other product combines with $H_2PO_4^-$. In the phosphorylase reaction, the reacting form of phosphoric acid is HPO_4^{2-}, and the products combine with H^+ and PO_4^{3-}, respectively. Cleavage by phosphoric acid in phosphorolysis is analogous to cleavage by water in hydrolysis.

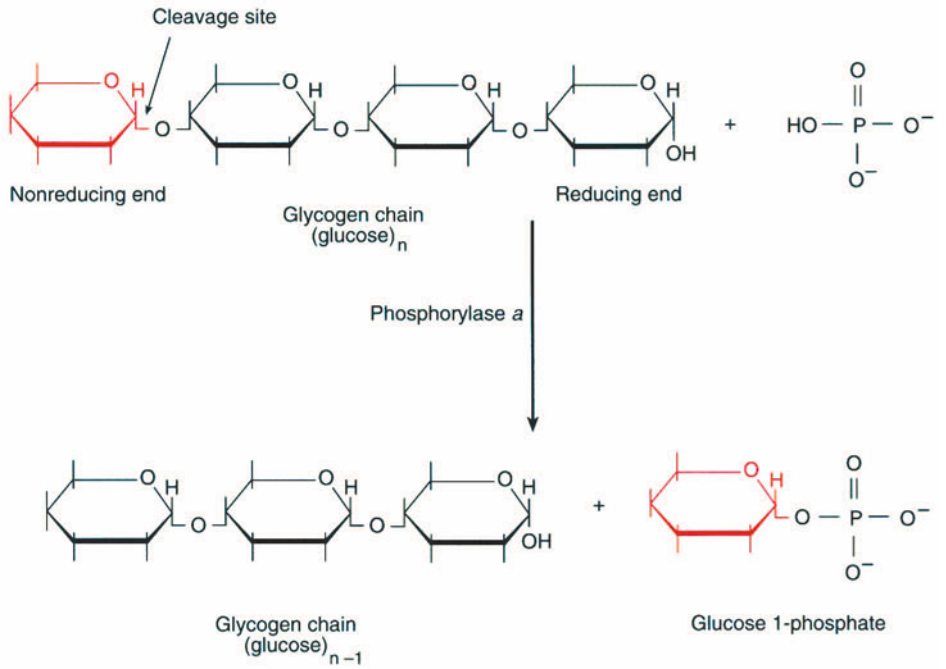


Figure 10.24. Action of glycogen phosphorylase. The enzyme catalyzes a phosphorolysis reaction whereby one glucose residue at a time is removed from the nonreducing end of glycogen in the form of glucose 1-phosphate.

Following the phosphorylase reaction, **phosphoglucomutase** catalyzes the isomerization of glucose 1-phosphate to glucose 6-phosphate, an intermediate in glycolysis:

Glucose 1-phosphate ⇄ glucose 6-phosphate

Phosphorylase causes digestion of glycogen chains by catalyzing the breaking of α(1 → 4) glycosidic bonds between glucose residues. The enzyme cannot catalyze the breaking of α(1 → 6) glycosidic bonds at branch points in the molecule, and its action stops within four residues of a branch point. Complete degradation of glycogen requires the action of an additional enzyme called **debranching enzyme.** Debranching enzyme has two enzymatic activities, a transfer function and a hydrolytic function, making it a *bifunctional enzyme.*

When phosphorylase action stops, debranching enzyme catalyzes a transfer of three residues to the nonreducing end of some other branch or of the core chain (Figure 10.25). Following transfer, debranching enzyme catalyzes the hydrolysis of the glucose residue left at the original branch point. The extended branch, or core chain, can now be attacked further by phosphorylase.

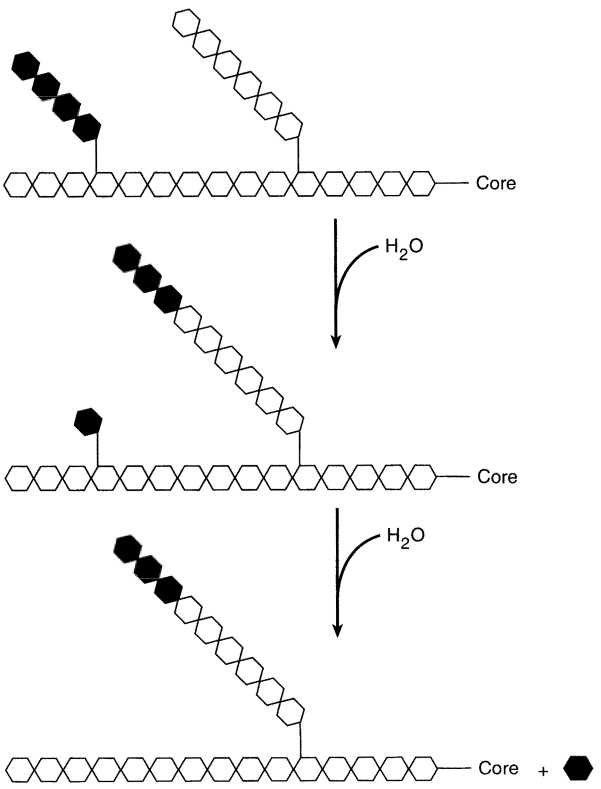


Figure 10.25. Mechanism of action of debranching enzyme. Three glucose residues of a "limit branch" are transferred to some other branch, followed by hydrolysis of the glucose residue at the original branch point.

10.5.1B. Interaction of Hormones with Cells. Glycogenolysis represents a key link between stored energy of carbohydrates (glycogen) and their workable currency (glucose). Not surprisingly, a process of

such pivotal importance is subject to complex control. The control mechanism of glycogenolysis involves the interaction of *hormones* with the cell membrane and an intracellular *enzyme cascade.* The control mechanism illustrates the chain of events that is triggered when a hormone binds to the cell membrane of a target cell and produces a specific response within the cell.

Hormone interactions with target cells involve three important components—*receptors, G proteins,* and *second messengers.* **G proteins** (so called because they bind guanine nucleotides) serve as membrane-bound signal transducers between hormone receptors in the cell membrane and some intracellular signaling system; they bind GDP in their inactive state and GTP in their active state.

When a hormone binds to a receptor on the outer side of the cell membrane, the receptor undergoes a conformational change that allows it to interact with an inactive G protein. The interaction of receptor and G protein leads to activation of the G protein and causes dissociation of its bound GDP. The dissociating GDP is replaced by GTP. The active G protein, with its bound GTP, interacts with and activates an intracellular signaling system. Subsequent to this interaction, a specific substance is released inside the cell and initiates intracellular changes. We term the substance functioning in this capacity a *second messenger,* and we term the hormone binding to the cell membrane a first messenger.

Active G proteins possess *GTPase* (GTP hydrolase) activity. GTPase catalyzes the hydrolysis of GTP, bound to the active G protein, and thereby converts it back to GDP. Because of this enzymatic activity, the activation of a G protein is short-lived. Nevertheless, the receptor–G protein interaction results in *amplification* of the hormonal signal because each hormone–receptor complex interacts with and activates many G proteins. Moreover, each activated G protein activates many intracellular signaling systems before it becomes inactivated.

10.5.1C. Enzyme Cascade of Glycogenolysis. Receptors for the control system of glycogenolysis are located on the outside of the cell membrane and bind the hormones *epinephrine* and *glucagon* (Figure 10.26). When either of these two hormones becomes bound to the receptor, a G protein is activated that in turn activates a molecule of **adenylate cyclase.** Adenylate cyclase is bound to the interior side of the cell membrane and catalyzes the formation of **cyclic AMP (3′,5′-cyclic adenylic acid; cAMP)** from ATP:

$$ATP^{4-} + H_2O \longrightarrow cAMP^{-} + PP_i^{3-}$$

Cyclic AMP constitutes the second messenger of this control system, transmitting the signal from the first messenger, the hormone, to a cellular enzyme and thereby initiating the intracellular changes. Cyclic AMP functions as an allosteric effector of **protein kinase,** the first enzyme in the intracellular enzyme cascade. Protein kinase activates phosphorylase kinase, which in turn activates phosphorylase. Investigators estimate that the total cascade mechanism of glycogenolysis provides an amplification of about 25×10^6 for the incoming signal. In other words, during the time that one hormone molecule acts on a membrane receptor, some 25×10^6 molecules of active glycogen phosphorylase form.

Protein kinase, the enzyme activated by cAMP, is a tetramer of two catalytic (C) and two regulatory (R) subunits. When cAMP binds to R subunits, C subunits dissociate. Catalytic subunits are active only after they dissociate from the tetramer. Protein kinase activates **phosphorylase kinase,** another tetramer but composed of four nonidentical subunits. Two of the larger subunits contain serine residues that become phosphorylated when the enzyme is activated. The smallest subunit consists of a regulatory protein called *calmodulin* (*cal*cium-*modul*ating prote*in*) that binds calcium. Calmodulin, a small protein of 148 amino acids, has four high-affinity binding sites for Ca^{2+}. Calmodulin's amino acid sequence is highly conserved, and the protein is widespread among eukaryotes. Calmodulin participates in the regulation of numerous cellular processes by controlling the level of Ca^{2+}.

Binding of calcium by phosphorylase kinase leads to partial activation of the enzyme. The enzyme has greatest activity when it is both phosphorylated and has calcium bound to it. The fourth subunit of phosphorylase kinase contains the active site that catalyzes the phosphorylation and activation of glycogen phosphorylase, the last enzyme in the cascade. Glycogen phosphorylase exists as a dimer of two identical subunits. Its active form has a phosphate group esterified at a serine residue in each subunit. Both phosphorylase kinase and phosphorylase can be dephosphorylated, and inactivated, by specific *phosphatases.* We refer to enzymes, such as protein kinase or phosphatase, that catalyze interconversions of two forms of another enzyme as **converter enzymes.**

10.5.1D. "Fight or Flight" Response. Glycogen structure, the specificity of glycogen phosphorylase, and hormonal control of glycogenolysis play key roles in an important physiological response called the **"fight or flight" response.** This response is elicited in vertebrate animals when they are subjected to stress. Various stress stimuli lead to secretion of epinephrine, the principal hormone governing the response. Epinephrine helps the animal prepare for the emergency and cope with the imposed stress by triggering a number of changes, including an increase in heart rate and blood pressure and an increase in the generation of ATP.

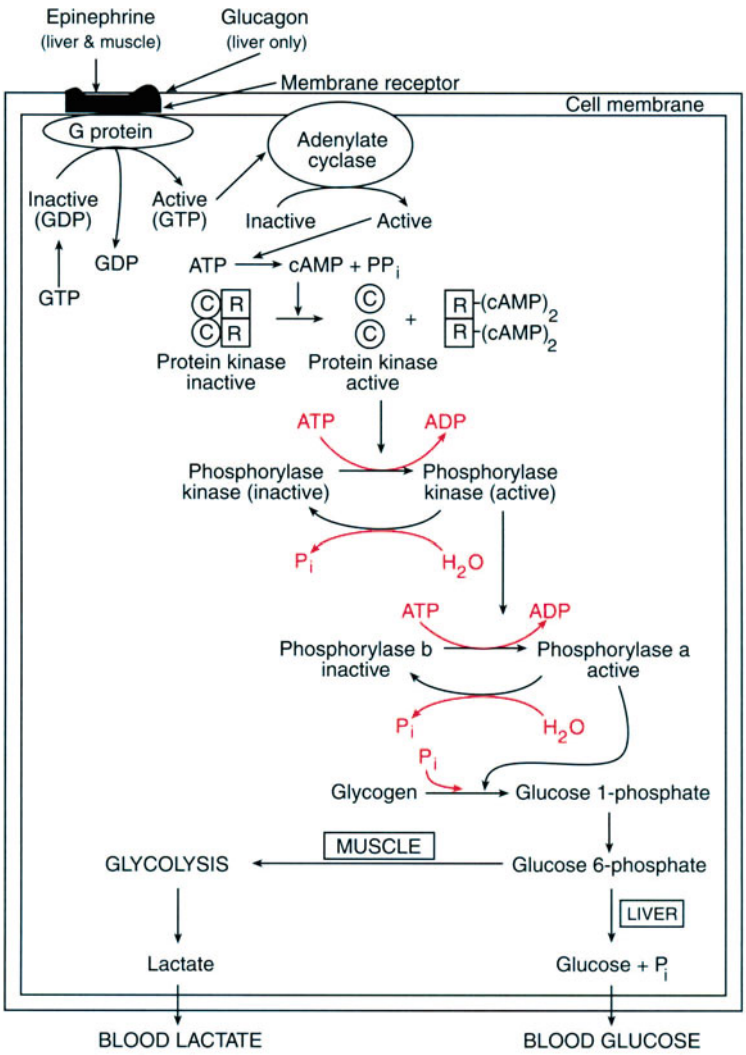


Figure 10.26. The hormonally controlled enzyme cascade regulating the activity of glycogen phosphorylase.

Having $\alpha(1 \rightarrow 6)$ branch points means that glycogen molecules possess a larger number of nonreducing ends than if the same mass were to exist in straight-chain molecules. Because glycogen phosphorylase acts only at nonreducing ends, the branched structure of glycogen leads to production of larger amounts of glucose 1-phosphate *per unit time* than would be produced by phosphorylase from a comparable mass of straight-chain molecules. Increased production of glucose 1-phosphate presents an advantage for the "fight or flight" response because it allows for rapid generation of ATP over and above that resulting from the enzyme cascade.

Formation of glucose 1-phosphate in muscle stimulates glycolysis and results in generation of ATP (Figure 10.26). Simultaneously, production of glucose 1-phosphate in the liver leads to an increase in the level of blood sugar that further fuels muscles with glucose for generation of additional ATP. The ATP formed helps muscles cope with the stress that triggered the release of epinephrine.

10.5.2. Glycogenesis

10.5.2A. Glucose Activation. *Glycogenesis,* or the synthesis of glycogen from glucose, occurs when the concentration of glucose exceeds that needed for generation of energy by glycolysis and when glycogen stores must be replenished. Glycogenesis involves a glycolytic enzyme and several other enzymes that are unique to this anabolic pathway. The process begins with activation of glucose to glucose 6-phosphate, catalyzed by the glycolytic enzyme hexokinase, or by glucokinase when the glucose concentration is high:

$$\text{Glucose} + \text{ATP}^{4-} \rightarrow \text{glucose 6-phosphate}^{2-} + \text{ADP}^{3-} + \text{H}^{+}$$

Next, phosphoglucomutase catalyzes the isomerization of glucose 6-phosphate to glucose 1-phosphate:

$$\text{Glucose 6-phosphate}^{2-} \rightleftarrows \text{glucose 1-phosphate}^{2-}$$

Glucose 1-phosphate, the product of glycogen breakdown, is also the building block for glycogen synthesis. In a reaction catalyzed by *UDP-glucose pyrophosphorylase,* glucose 1-phosphate is converted to uridine diphosphate glucose, or *UDP-glucose* (Figure 10.18):

$$\text{Glucose 1-phosphate}^{2-} + \text{UTP}^{4-} + \text{H}^{+} \rightarrow \text{UDP-glucose}^{2-} + \text{PP}_i^{3-}$$

In this reaction, the phosphoryl oxygen of glucose 1-phosphate attacks the α-phosphorus atom of UTP. UDP-glucose and similar *nucleotide-linked sugars* represent compounds in which the nucleoside diphosphate functions as a carrier of an activated *glycosyl group.* Nucleotide-linked sugars constitute energy-rich compounds that play important roles in carbohydrate metabolism.

The reaction catalyzed by UDP-glucose pyrophosphorylase is freely reversible ($\Delta G^{\circ\prime} \approx 0$). You might expect this since both the reactants and the products have a total of two energy-rich bonds each (two in UTP, one each in UDP-glucose and PP_i). However, whenever pyrophosphate forms in metabolism, it generally undergoes hydrolysis to inorganic phosphate ($PP_i \rightarrow 2P_i$) in a reaction catalyzed by *pyrophosphatase.* This reaction is strongly exergonic because PP_i constitutes an energy-rich compound of the acid anhydride type. Hydrolysis of PP_i in the UDP-glucose pyrophosphorylase reaction ensures that formation of UDP-glucose is irreversible.

10.5.2B. Glycogen Synthase.

Glycogen synthase catalyzes the polymerization of activated glucose residues. The enzyme occurs tightly bound to intracellular glycogen granules. During polymerization, one glucose residue at a time adds to the nonreducing end of the growing polysaccharide chain (Figure 10.27):

$$(\text{Glucose})_n + \text{UDP-glucose}^{2-} \rightarrow (\text{glucose})_{n+1} + \text{UDP}^{3-} + \text{H}^{+}$$

The reaction is exergonic ($\Delta G^{\circ\prime} = -13.4$ kJ mol^{-1}), but the enzyme requires a *primer* for initiation of polymerization. Glycogen molecules present in the cell can serve in this capacity and become extended by the action of glycogen synthase. It appears that the best primers consist of short-chain glycogen molecules. The enzyme has a small K_m for short chains but larger K_m values for longer chains.

Complete glycogen synthesis requires a mechanism for the synthesis of primers. Primer synthesis occurs via two steps and involves a protein named *glycogenin* (MW = 37,000). The first step seems to require a specific glucosyltransferase that catalyzes transfer of a glucose residue from UDP-glucose to glycogenin. The glucose becomes attached via a glycosidic bond involving the 1′-OH of glucose (the

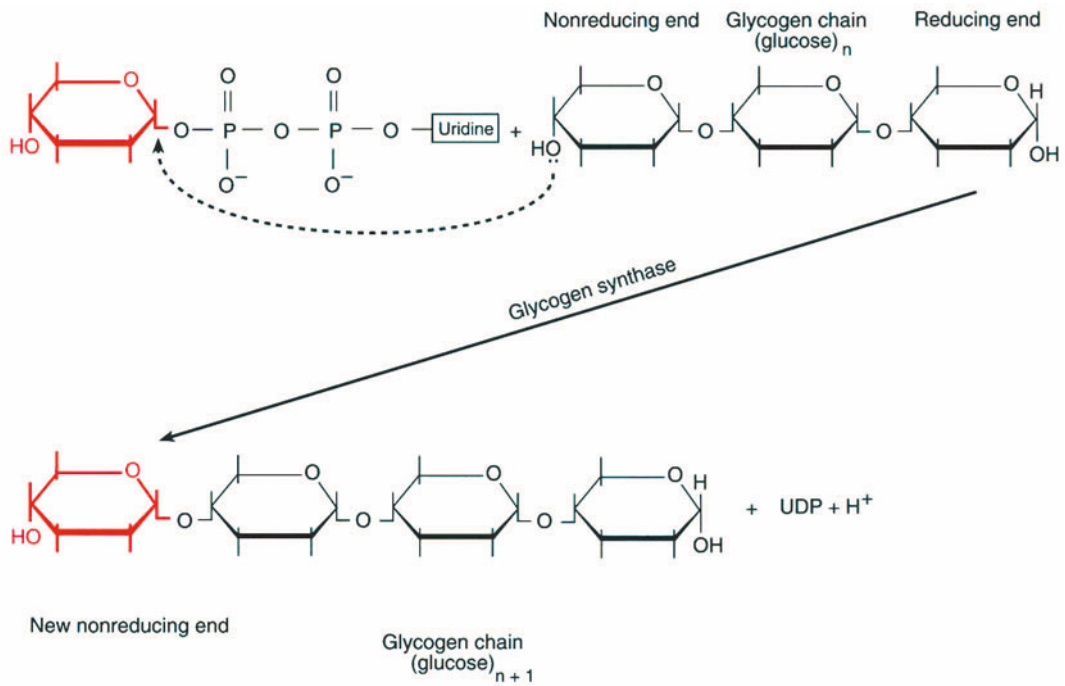


Figure 10.27. The glycogen synthase reaction. One glucose residue at a time is added to the nonreducing end of glycogen.

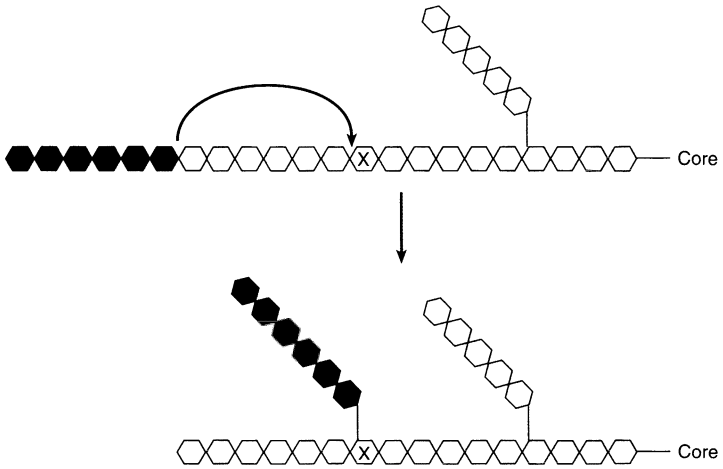


Figure 10.28. Mechanism of action of branching enzyme. A segment of 6 or 7 glucose residues is removed from a chain containing at least 11 residues and is reanchored as a branch at least 4 residues removed from another branch.

reducing end) and the phenolic hydroxyl group of a tyrosine residue in glycogenin. The second step, catalyzed by glycogenin itself, consists of extending the bound glucose with up to seven additional glucose residues, using UDP-glucose as a substrate. Subsequent linear extension of the primer, attached to glycogenin, is catalyzed by glycogen synthase. Thus, glycogenin serves both as an enzyme and as a scaffold for glycogen synthesis. UDP, released during synthesis of the primer, is reconverted to UTP by action of *nucleoside diphosphate kinase:*

$$\text{UDP} + \text{ATP} \rightleftarrows \text{UTP} + \text{ADP}$$

Action of glycogen synthase produces straight chains in which glucose residues are linked by $\alpha(1 \rightarrow 4)$ glycosidic bonds. To complete the synthesis of glycogen and form branches attached via $\alpha(1 \rightarrow 6)$ glycosidic bonds requires the action of **branching enzyme** (*1,4 → 1,6 transglycosylase).* Branching enzyme catalyzes transfer of a segment containing 6 or 7 residues—from a chain containing at least 11 residues—to some place on the same chain or on a different chain. The new branch point must be at least four residues removed from an adjacent branch point (Figure 10.28). Branching proceeds readily and does not involve large free energy changes because the energy required for breaking an $\alpha(1 \rightarrow 4)$ bond is very similar, but opposite in sign, to that required for forming an $\alpha(1 \rightarrow 6)$ bond.

10.5.2C. Regulation of Glycogenesis. Control of glycogenesis is complex. One regulatory effect results from the concentration of glycogen itself. Glycogen exerts feedback inhibition on its own formation by inhibiting glycogen synthase; consequently, the K_m of the enzyme increases with increasing glycogen chain length.

A second regulatory mechanism involves covalent modification of glycogen synthase. Glycogen synthase is an oligomeric protein consisting of four identical subunits (MW = 85,000 each). Each subunit can be phosphorylated and dephosphorylated at the OH of a serine residue (Figure 10.29). Phosphorylation and dephosphorylation of glycogen synthase are catalyzed by, respectively, the same kinase and the same phosphatase that catalyze phosphorylation and dephosphorylation of glycogen phospho-

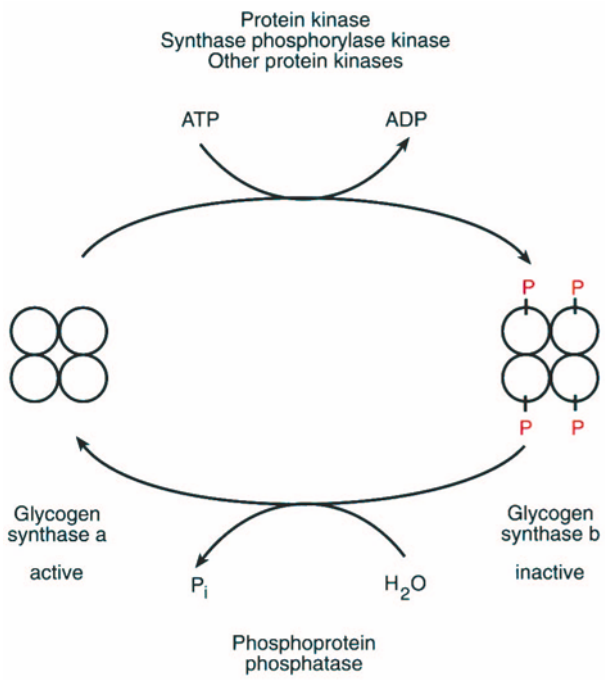


Figure 10.29. Covalent modification of glycogen synthase. The enzyme is a tetramer, composed of four identical subunits.

rylase. Glucagon and epinephrine, which regulate glycogen phosphorylase, also control activation of the kinase and the phosphatase that act on glycogen synthase. In both instances, activation involves a similar enzyme cascade (Figure 10.30); cyclic AMP serves as a second messenger for both glycogen synthesis and glycogen degradation.

Dephosphorylated glycogen synthase, designated "*a*," constitutes the *active* form of the enzyme. Glycogen synthase *a* is independent of the concentration of glucose 6-phosphate. *Phosphorylated* glycogen synthase, designated "*b*," constitutes the *inactive* form of the enzyme. Several phosphorylated varieties of glycogen synthase occur because at least six other protein kinases, in addition to that also acting on glycogen phosphorylase, can deactivate glycogen synthase by phosphorylation. The various phosphorylations involve at least nine different serine residues per subunit. A number of secondary messengers of hormone action, including cAMP, Ca^{2+}, and diacylglycerol, regulate the activity of several kinases that function in glycogenolysis and glycogenesis.

The "*b*" form of glycogen synthase is active only in the presence of glucose 6-phosphate. At high concentrations of glucose 6-phosphate, the inactive enzyme becomes active; glucose 6-phosphate functions as an allosteric effector of glycogen synthase *b*. The allosteric effect of glucose 6-phosphate constitutes yet another regulatory mechanism of the enzyme and makes good metabolic sense. Activation of glycogen synthase by high concentrations of glucose 6-phosphate ensures that excess glucose 6-phosphate is converted to glycogen and stored, rather than degraded via glycolysis to generate unnecessary energy.

Insulin also has an effect on glycogen synthesis. Insulin stimulates phosphatase, thereby increasing the dephosphorylation and activation of glycogen synthase. Accordingly, insulin opposes the action of glucagon and epinephrine, promotes glycogen synthesis from glucose, and leads to a lower level of blood sugar.

From our discussion, you can see that glycogenolysis and glycogenesis are *reciprocally regulated* (Figure 10.31); when one process is activated, the other is inhibit-

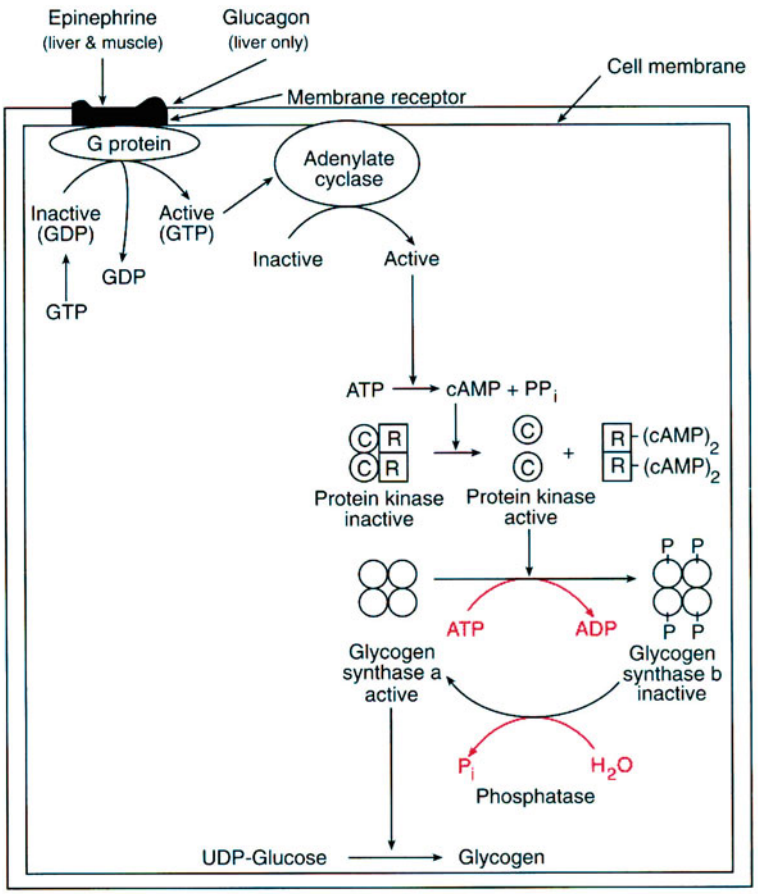


Figure 10.30. The hormonally controlled enzyme cascade regulating the activity of glycogen synthase. Initial reactions are identical to those of the glycogen phosphorylase cascade (Figure 10.26).

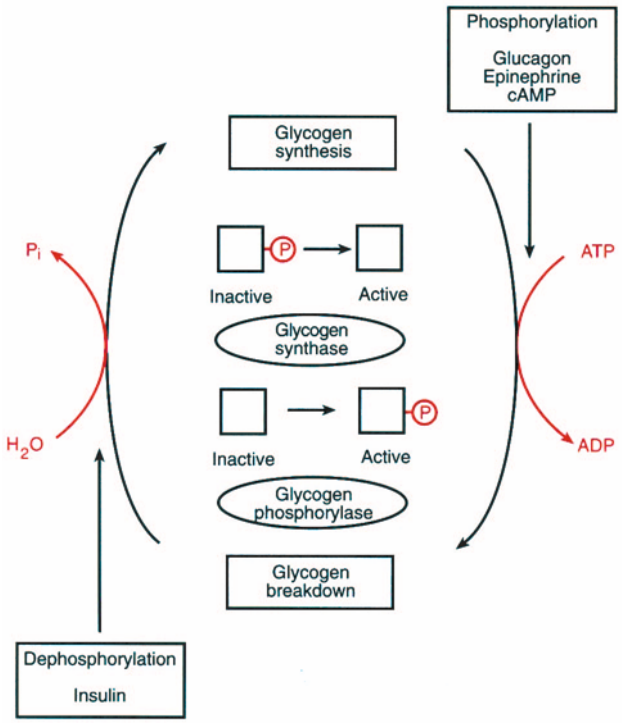


Figure 10.31. Reciprocal regulation of glycogen synthesis and degradation. Activation of one pathway leads to inhibition of the other.

ed. Phosphorylation leads to activation of glycogen phosphorylase but inactivation of glycogen synthase. Dephosphorylation has the opposite effects. Reciprocal regulation allows for effective control of carbohydrate metabolism.

While both glycogen phosphorylase and glycogen synthase are regulated by similar hormonally controlled enzyme cascades, note that the glycogen breakdown cascade has one more cycle than the glycogen synthesis cascade. This means that glycogen degradation has greater sensitivity than glycogen synthesis; degradation can be amplified to a greater extent than synthesis. Researchers have estimated that glycogen degradation can proceed about 300 times faster than glycogen synthesis. The different sensitivity of the two cascades is metabolically useful. It is critical that the cell possess the capacity to rapidly generate glucose from glycogen stores and convert it to usable energy. It is far less critical that the cell be capable of rapidly storing excess glucose in the form of glycogen.

Scientists have identified a number of hereditary disorders of glycogen metabolism, called **glycogen storage diseases** (see Table 8.3). Some of these lead to glycogen that has an abnormal structure, such as glycogen having short outer chains or no outer chains or glycogen having long and unbranched chains. Other glycogen storage diseases, while resulting from specific enzymatic deficiencies, lead to glycogen that has a normal molecular structure. Abnormalities in glycogen metabolism are usually most pronounced in the liver, but heart and muscle glycogen metabolism can also be defective.

10.6. GLUCONEOGENESIS

Gluconeogenesis refers to the synthesis of glucose from noncarbohydrate precursors such as lactate, pyruvate, glycerol, and amino acids (Figure 10.32). Gluconeogenesis is crucial for survival of humans and other animals because these organisms possess a number of tissues, including the central nervous system, kidney medulla, red blood cells, and testes, that use primarily glucose as fuel to generate energy. An adequate level of blood sugar must be maintained so that these tissues receive the necessary amounts of glucose. Generally, dietary intake of carbohydrates and carbohydrate stores are insufficient to meet the organism's need for glucose.

Much of the dietary supply of glucose becomes rapidly exhausted by catabolism. Strenuous exercise, a carbohydrate-deficient diet, and a fast longer than one day all lead to accelerated depletion of dietary glucose and require glucose synthesis at even greater than normal rates.

Stored carbohydrate is also insufficient to meet the organism's need for glucose. In humans, body fluids contain about 20 g of glucose, and carbohydrate stored as glycogen amounts to about 190 g. By contrast, metabolism requires about 160 g of glucose per day, of which approximately 120 g/day is needed for operating the brain. Based on these quantities, the readily available glucose in a human can only meet the energy demands for approximately one day.

Thus, despite dietary intake of glucose and the presence of carbohydrate stores, glucose must be synthesized at all times to provide a continuous supply to those tissues using it as primary energy source.

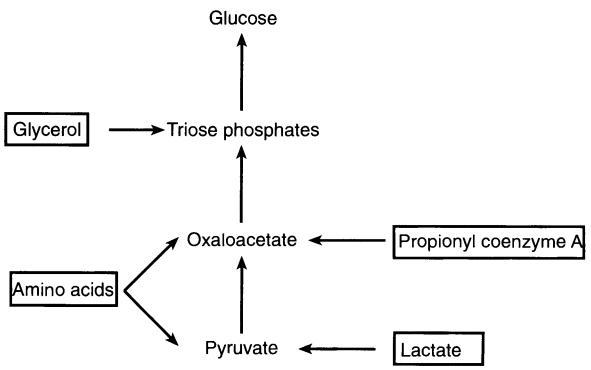


Figure 10.32. Major precursors for the biosynthesis of glucose. Some arrows represent multiple-step reactions.

10.6.1. Reversal of Glycolysis

In a sense, gluconeogenesis constitutes a reversal of glycolysis: glucose is synthesized from pyruvate. Recall, however, that an anabolic pathway *cannot* be the exact reverse of a catabolic sequence (Section 8.2); hence, gluconeogenesis cannot be an exact reversal of glycolysis. Glycolysis proceeds unidirectionally from glucose to pyruvate because of three *strongly exergonic* and irreversible steps that involve the conversions of (a) glucose to glucose 6-phosphate, (b) fructose 6-phosphate to fructose 1,6-bisphosphate, and (c) phosphoenolpyruvate to pyruvate. In order to accomplish a reversal of glycolysis, each of these three steps, which are *strongly endergonic* in the reverse direction, must be replaced by a suitable pathway alteration termed a **bypass.** The three bypasses of gluconeogenesis occur in the following order (Figure 10.33):

Bypass I:	Pyruvate to phosphoenolpyruvate
Bypass II:	Fructose 1,6-bisphosphate to fructose 6-phosphate
Bypass III:	Glucose 6-phosphate to glucose

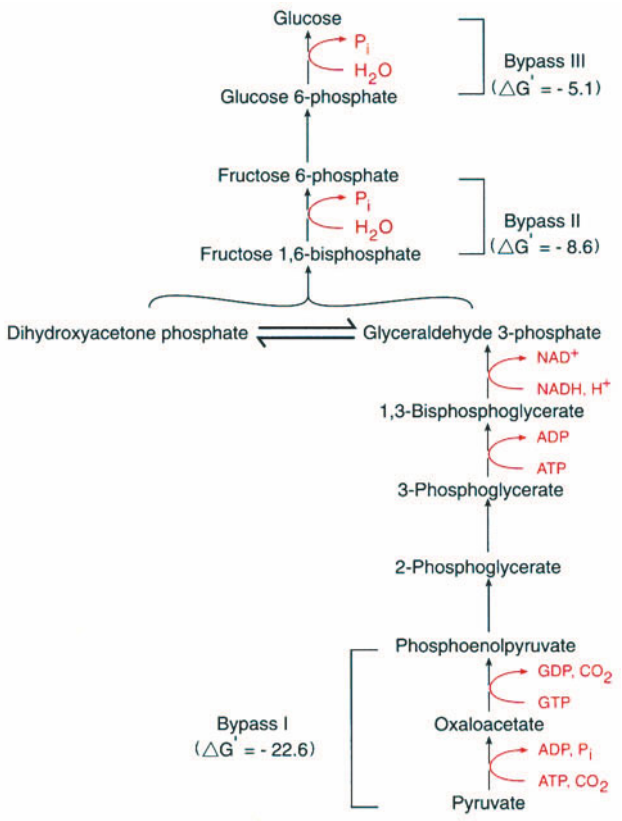


Figure 10.33. The pathway of gluconeogenesis. Free energy changes ($\Delta G'$) of the bypasses are in kilojoules per mole, determined for liver.

The remaining steps of glycolysis are readily reversible reactions ($\Delta G' \approx 0$) that can be traversed in either direction. These steps form part of both glycolysis and gluconeogenesis. Biochemists usually consider oxaloacetate to be the true starting material of gluconeogenesis. Although gluconeogenesis commences with pyruvate, the capacity to form oxaloacetate constitutes the actual step permitting a reversal of glycolysis.

10.6.1A. Bypass I. Bypass I achieves the conversion of pyruvate to phosphoenolpyruvate. The energy barrier is overcome by changing the reaction to a *two-step mechanism, requiring two nonglycolytic enzymes and a nonglycolytic intermediate (oxaloacetate). Both the reactants and the products of the glycolytic step are altered.*

Pyruvate carboxylase catalyzes the first step, involving a carboxylation of pyruvate to form oxaloacetate ($^-OOC\text{-}CH_2\text{-}CO\text{-}COO^-$), an intermediate of the citric acid cycle:

$$\text{Pyruvate}^- + CO_2 + ATP^{4-} + H_2O \rightarrow \text{oxaloacetate}^{2-} + ADP^{3-} + P_i^{2-} + 2H^+$$

Pyruvate carboxylase, a tetramer composed of four identical subunits (MW ≈ 120,000 each), occurs in the mitochondrial matrix. Each subunit binds Mg^{2+} and carries *biotin* (see Section 11.5) as a prosthetic group. Acetyl coenzyme A is a powerful allosteric activator of pyruvate carboxylase; without bound acetyl coenzyme A, the enzyme has essentially no activity.

Phosphoenolpyruvate carboxykinase catalyzes the second step:

$$\text{Oxaloacetate}^{2-} + GTP^{4-} \rightarrow \text{phosphoenolpyruvate}^{3-} + GDP^{3-} + CO_2$$

The overall reaction is:

$$\text{Pyruvate}^- + ATP^{4-} + GTP^{4-} + H_2O \downarrow \text{phosphoenolpyruvate}^{3-} + ADP^{3-} + GDP^{3-} + P_i^{2-} + 2H^+$$

The subcellular location of phosphoenolpyruvate carboxykinase varies among different organisms. In mouse and rat livers, the enzyme occurs almost exclusively in the cytosol; in pigeon and rabbit livers, it occurs inside the mitochondria; and in guinea pigs and humans, it is more or less evenly distributed between cytosolic and mitochondrial compartments. The intracellular location of phosphoenolpyruvate carboxykinase necessitates specific transport systems for either oxaloacetate or phospho-

enolpyruvate. If the enzyme is located in the cytosol, oxaloacetate must leave the mitochondria for conversion to phosphoenolpyruvate in the cytosol. If the enzyme is located in the mitochondria, phosphoenolpyruvate formed inside the mitochondria must move out into the cytosol for participation in gluconeogenesis. All subsequent reactions of gluconeogenesis, beginning with phosphoenolpyruvate, occur in the cytosol. Other sources of oxaloacetate for gluconeogenesis include degradation of amino acids and degradation of fatty acids with an odd number of carbons atoms.

10.6.1B. Bypass II. Bypass II occurs at the conversion of fructose 1,6-bisphosphate to fructose 6-phosphate. Strict reversal of the glycolytic reaction that converts fructose 6-phosphate to fructose 1,6-bisphosphate, with phosphate provided by ATP, requires synthesis of ATP from ADP, a step that is strongly endergonic. The difficult ATP synthesis is avoided by substituting a simple hydrolysis reaction that involves a *different enzyme. Both the reactants and the products of the glycolytic step are altered.* The reaction of bypass II is catalyzed by fructose 1,6-bisphosphatase:

$$\text{Fructose 1,6-bisphosphate}^{4-} + H_2O \rightarrow \text{fructose 6-phosphate}^{2-} + P_i^{2-}$$

10.6.1C. Bypass III. Bypass III occurs at the conversion of glucose 6-phosphate to glucose and is chemically analogous to bypass II. Here, too, reversal of the corresponding glycolytic step requires synthesis of ATP from ADP. Once again, the difficult ATP synthesis is avoided by substituting a simple hydrolysis reaction that involves a *different enzyme. Both the reactants and the products of the glycolytic step are altered.* The reaction is catalyzed by glucose 6-phosphatase, an enzyme bound to the endoplasmic reticulum:

$$\text{Glucose 6-phosphate}^{2-} + H_2O \rightarrow \text{glucose} + P_i^{2-}$$

10.6.1D. Overall Reaction. Bypass I requires energy in the form of both ATP and GTP. Because GTP is equivalent to ATP in terms of energy-rich bonds, the energy requirement of bypass I totals 2 ATP (two energy-rich bonds). Synthesis of 1,3-bisphosphoglycerate from 3-phosphoglycerate requires an additional ATP, making a total requirement of 3 ATP. However, synthesis of one molecule of glucose requires two molecules of pyruvate, forming two molecules of 1,3-bisphosphoglycerate. Accordingly, the total energy cost of gluconeogenesis is 6 ATP per molecule of glucose synthesized and we can write the overall reaction as

$$\begin{gathered} 2\ \text{Pyruvate}^- + 2\text{NADH} + 4\text{ATP}^{4-} + 2\text{GTP}^{4-} + 6H_2O \\ \downarrow \\ \text{glucose} + 2\text{NAD}^+ + 4\text{ADP}^{3-} + 2\text{GDP}^{3-} + 6P_i^{2-} + 2H^+ \end{gathered}$$

Actually, the true energy cost exceeds the six energy-rich bonds indicated by this equation. Conversion of two molecules of 1,3-bisphosphoglycerate to two molecules of glyceraldehyde 3-phosphate requires input of 2 NADH. If not thus expended, NADH could be oxidized to NAD^+ via the electron transport system, resulting in a gain of 4 or 5 ATP depending on the type of shuttle required (see Section 12.6). Loss of 4 or 5 ATP raises the *effective energy cost* of gluconeogensis to 10 or 11 ATP per molecule of glucose synthesized.

10.6.2. Glucose–Alanine Cycle

Two important cycles function to provide the liver, the central organ for gluconeogenesis, with appropriate substrates for glucose synthesis: the *Cori cycle,* discussed in Section 10.3, and the **glucose–alanine cycle.**

In the glucose–alanine cycle, glucose is converted to pyruvate in muscle tissue, and *transamination* (Section 14.2) converts pyruvate to alanine. Blood then transports the alanine to the liver, where transamination reconverts it to pyruvate, which then serves as substrate for gluconeogenesis. Blood returns the glucose produced to the muscles (Figure 10.34).

A major difference between the glucose–alanine and Cori cycles lies in the nature of the three-carbon com-

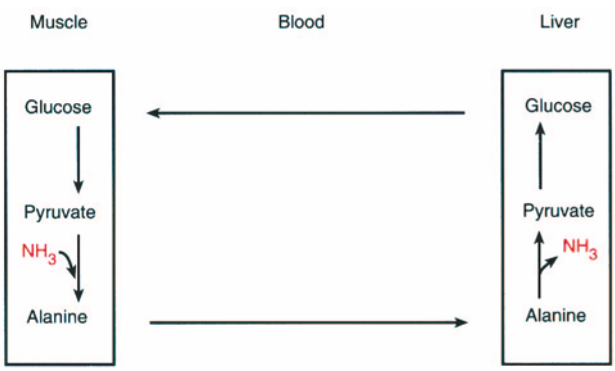


Figure 10.34. The glucose–alanine cycle. Muscle glycolysis converts glucose to pyruvate. Gluconeogenesis converts pyruvate to glucose in the liver. Interconversions of alanine and pyruvate in both tissues involve transaminations.

pound transported by the blood from muscle tissue to the liver. Blood transports lactate in the Cori cycle but transports alanine in the glucose–alanine cycle.

10.6.3. Effect of Alcohol

Ingesting alcohol inhibits gluconeogenesis and leads to lower blood sugar levels. A potentially dangerous low level of blood sugar, termed *hypoglycemia,* can develop. The effect results from the first reaction of alcohol catabolism—the oxidation of ethanol to acetaldehyde, catalyzed by *alcohol dehydrogenase:*

$$\text{Ethanol} + \text{NAD}^+ \rightarrow \text{acetaldehyde} + \text{NADH} + \text{H}^+$$

This reaction produces NADH and leads to an increase in the NADH/NAD^+ ratio of the cell. Increased concentrations of NADH drive the *lactate dehydrogenase* and *malate dehydrogenase* (Section 11.3) reactions toward formation of lactate and malate, respectively:

$$\text{Pyruvate}^- + \text{NADH} + \text{H}^+ \rightarrow \text{lactate}^- + \text{NAD}^+$$

$$\text{Oxaloacetate}^{2-} + \text{NADH} + \text{H}^+ \rightarrow \text{malate}^{2-} + \text{NAD}^+$$

As lactate and malate form, concentrations of pyruvate and oxaloacetate decrease. Low levels of pyruvate and oxaloacetate result in a strong depression of gluconeogenesis and a severe lowering of the blood sugar level.

10.6.4. Regulation of Gluconeogenesis

If neither glycolysis nor gluconeogenesis were regulated, the combined pathways would constitute a giant **futile cycle** that achieves nothing except dissipation of the free energy of hydrolysis of ATP and GTP (Figure 10.35). Fortunately, this is not the case.

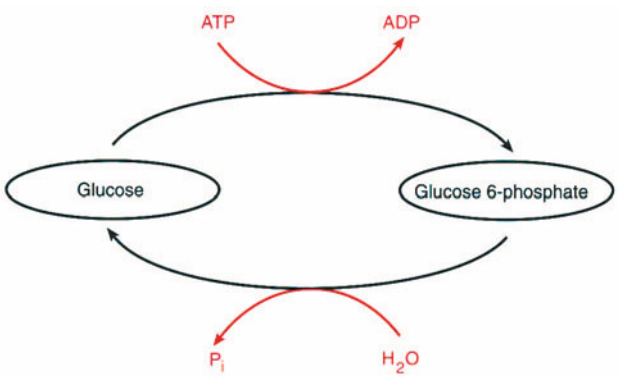


Figure 10.35. Hypothetical example of a futile cycle.

Like glycogenolysis and glycogenesis, glycolysis and gluconeogenesis are reciprocally regulated. Control points occur at the three steps where the forward and reverse reactions can be independently regulated—in other words, where bypasses function in gluconeogenesis. Regulatory effects in animals occur at all three steps for glycolysis, but only at two steps for gluconeogenesis (Figure 10.36). Major control of gluconeogenesis, much as that of glycolysis, involves allosteric enzyme effectors.

10.6.4A. Fructose 2,6-Bisphosphate. Fructose 2,6-bisphosphate, an isomer of fructose 1,6-bisphosphate, is a key allosteric effector in the regulation of glycolysis and gluconeogenesis; it functions as an *activator of phosphofructokinase in glycolysis* and as an *inhibitor of fructose 1,6-bisphosphatase in gluconeogenesis.* The concentration of fructose 2,6-bisphosphate is regulated by controlling both its rate of synthesis and its rate of degradation. *Phosphofructokinase-2* (the glycolytic enzyme is termed phosphofructokinase-1) catalyzes the synthesis of fructose 2,6-bisphosphate, and *fructose bisphosphatase-2* catalyzes its degradation.

The enzymatic activities of both phosphofructokinase-2 and fructose bisphosphatase-2 occur on the same *bifunctional* protein. Both activities are subject to multiple allosteric control and to covalent modification by phosphorylation. Phosphorylation inactivates phosphofructokinase-2 and activates fructose bisphosphatase-2. Phosphorylation/dephosphorylation is controlled by cAMP-dependent protein kinase and a phosphoprotein phosphatase. These multiple factors combine to provide a finely tuned regulation of the concentration of fructose 2,6-bisphosphate.

10.6.4B. Carbohydrate Catabolism and Anabolism. We can summarize the overall regulation of glycolysis, glycogenolysis, glycogenesis, and gluconeogenesis as follows. A well-fed organism has a good dietary supply of glucose and does not require additional production of blood sugar by metabolic processes. Under these conditions, the liver functions to conserve metabolic fuels so that both carbohydrate and lipid stores become augmented. Some glucose serves directly for biosynthesis of glycogen (glycogenesis), and the rest undergoes catabolism (glycolysis) to acetyl coenzyme A, which serves as the starting material for fatty acid synthesis.

An organism that has fasted has a poor dietary supply of glucose and requires continuous production of blood sugar by metabolic processes. Under these conditions, the liver functions to draw on stored metabolic fuels by stimulating glycogen breakdown to glucose (glyco-

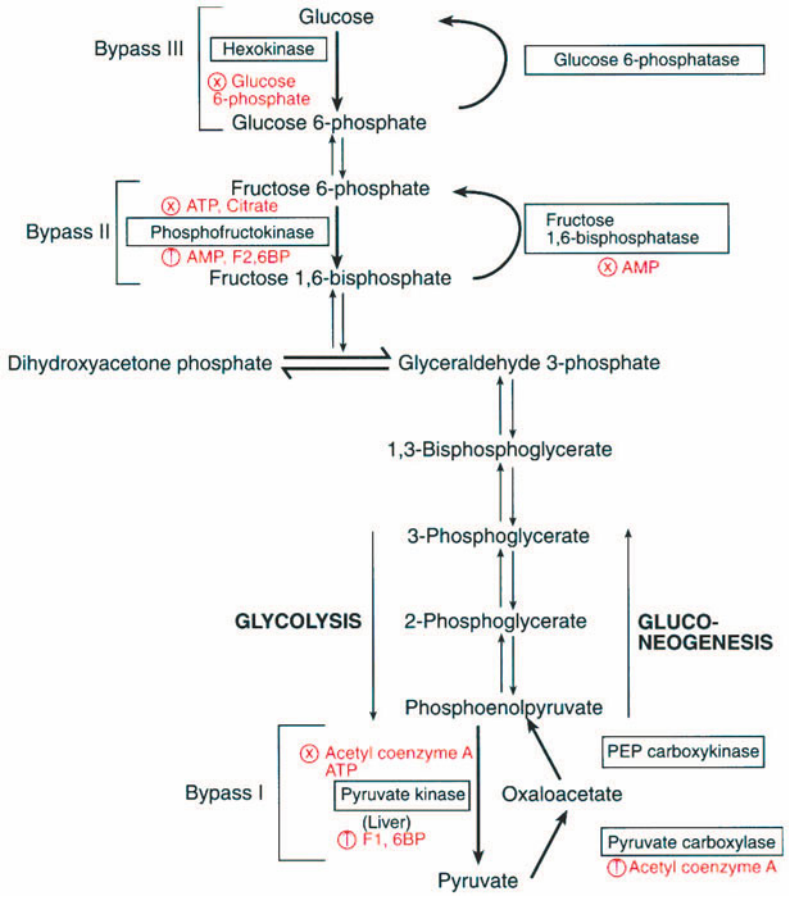


Figure 10.36. Reciprocal regulation of glycolysis and gluconeogenesis. ⊗ Inhibitor, ↑ activator. Only major inhibitors and activators of phosphofructokinase are shown. F2,6BP and F1,6BP designate fructose 2,6-bisphosphate and fructose 1,6-bisphosphate, respectively.

genolysis). The liver also functions to enhance glucose biosynthesis (gluconeogenesis), using primarily protein degradation products as precursors of pyruvate and oxaloacetate.

10.7. BIOSYNTHESIS OF OTHER CARBOHYDRATES

10.7.1. Polysaccharides

Most polysaccharides are synthesized from **nucleotide-linked sugars.** Both muscle and liver glycogen form from UDP-glucose, and ADP-glucose serves as precursor for the biosynthesis of glycogen in bacteria and for the biosynthesis of starch (amylose and amylopectin) in plants. For cellulose synthesis, different plant species use either UDP-glucose, ADP-glucose, or CDP-glucose as the activated form of the glycosyl residue. Chitin, a homopolysaccharide of *N*-acetylglucosamine, forms from UDP-*N*-acetylglucosamine. Biosynthesis of hyaluronic acid, a polymer of alternating units of glucuronic acid and *N*-acetylglucosamine, requires the action of two enzymes whose substrates are UDP-glucuronate and UDP-*N*-acetylglucosamine, respectively.

Synthesis of some bacterial **dextrans** constitutes an exception to the involvement of nucleotide-linked sugars. Dextrans are storage polysaccharides in yeasts and bacteria. They consist of D-glucose residues linked largely by $\alpha(1 \rightarrow 6)$ glycosidic bonds, with occasional branches formed by $\alpha(1 \rightarrow 2)$, $\alpha(1 \rightarrow 3)$, or $\alpha(1 \rightarrow 4)$ glycosidic bonds. Some bacterial species growing in the human mouth synthesize dextran—a major component of dental plaque—by transglycosylation, using sucrose as a substrate:

$$n \text{ Sucrose} \rightarrow (\text{glucose})_n + n \text{ fructose}$$

Dextran

Because dextran formation utilizes sucrose, consumption of execssive amounts of sucrose creates a serious concern for dental hygiene.

10.7.2. Oligosaccharides

Biosynthesis of sucrose in plants also proceeds via nucleotide-linked sugars:

$$\text{UDP-glucose}^{2-} + \text{fructose 6-phosphate}^{2-} \rightarrow$$
$$\text{sucrose 6-phosphate}^{2-} + \text{UDP}^{3-} + \text{H}^+$$

Sucrose

Lactose forms in mammary glands by action of lactose synthase, an enzyme consisting of two subunits. One subunit, called galactosyl transferase, catalyzes the following reaction when present by itself:

$$\text{UDP-galactose}^{2-} + \textit{N}\text{-acetylglucosamine} \rightarrow$$
$$\textit{N}\text{-acetyllactosamine} + \text{UDP}^{3-} + \text{H}^+$$

The second subunit, called α-lactalbumin, is a mammary gland protein that associates with preexisting galactosyl transferase and alters its specificity. After interaction with α-lactalbumin, the enzyme can use glucose as an acceptor of the galactosyl group instead of *N*-acetylglucosamine, thereby leading to lactose synthesis:

$$\text{UDP-galactose}^{2-} + \text{glucose} \rightarrow \text{lactose} + \text{UDP}^{3-} + \text{H}^+$$

In animals, the dimeric form of the enzyme occurs only in the mammary gland. Having the capacity to change galactosyl transferase activity to that of lactose synthase permits the gland to synthesize large quantities of lactose, required for the production of milk.

Biosynthesis of glycoproteins proceeds via two main pathways. O-linked oligosaccharides form by action of specific glycosyl transferases that catalyze transfer of a monosaccharide unit from a nucleotide-linked sugar to the nonreducing end of an oligosaccharide or to a functional group on a protein. These reactions occur in the Golgi apparatus.

N-linked oligosaccharides form as lipid-linked precursors. They are joined to dolichol, a long-chain polyisoprenoid compound that serves as a carbohydrate carrier. Initial synthesis and some processing take place in the endoplasmic reticulum. Glycoproteins are then transported to the Golgi apparatus for further processing.

10.7.3. Monosaccharides

Glucose can serve as a precursor for all other monosaccharides found in biological systems. Conversely, various monosaccharides can be converted to glucose derivatives. Interconversions of monosaccharides frequently involve isomerization, epimerization, and phosphorylation, and many of the reactions require nucleotide-linked sugars.

Amino sugars, building blocks of glycolipids and glycoproteins, derive from glucosamine. Glucosamine forms by transfer of an amino group from glutamine to fructose 6-phosphate:

$$\text{Fructose 6-phosphate}^{2-} + \text{glutamine}$$
$$\downarrow \text{transamidase}$$
$$\text{glucosamine 6-phosphate}^{2-} + \text{glutamate}^{-} + \text{H}^+$$

Glucosamine 6-phosphate leads to UDP-*N*-acetylglucosamine, which is converted to *N*-acetylneuraminic acid (sialic acid).

SUMMARY

Carbohydrate metabolism consists essentially of the metabolism of glucose. Most major pathways begin or end with glucose, and other carbohydrates are converted to glucose or to intermediates of glucose metabolism. Blood transports carbohydrates in the form of glucose (blood sugar) or lactate. The hormones insulin, epinephrine, and glucagon regulate the level of blood sugar.

Glycolysis comprises the catabolism of glucose to pyruvate and occurs in two stages. Stage I results in degradation of one molecule of glucose (six carbons) to two molecules of glyceraldehyde 3-phosphate (three carbons) and requires input of energy in the form of ATP. Stage II leads to conversion of glyceraldehyde 3-phosphate to pyruvate (three carbons) and is accompanied by production of NADH and generation of energy in the

form of ATP. ATP synthesis proceeds via substrate-level phosphorylation. Under aerobic conditions, pyruvate leads to acetyl coenzyme A, which enters the citric acid cycle. Under anaerobic conditions, pyruvate leads to lactate. Conversion of glucose to pyruvate or lactate yields two molecules of ATP per molecule of glucose catabolized. Combination of glycolysis and the citric acid cycle/electron transport system results in complete oxidation of carbohydrates to CO_2 and H_2O.

In the pentose phosphate pathway, glucose 6-phosphate undergoes oxidation to CO_2 and P_i. The cycle produces reducing power in the form of NADPH, serves to synthesize a number of carbohydrates, and allows for interconversions of monosaccharides.

Glycogen catabolism (glycogenolysis) involves removal of one glucose residue at a time from the nonreducing end of glycogen. Glucose is removed in the form of glucose 1-phosphate in a reaction catalyzed by glycogen phosphorylase. Activation of phosphorylase requires a hormonally controlled enzyme cascade. Glycogen biosynthesis (glycogenesis) is catalyzed by glycogen synthase, which uses UDP-glucose as substrate. Activation of glycogen synthase also requires a hormonally controlled enzyme cascade. Glycogenolysis and glycogenesis are reciprocally regulated; activation of one process is coupled to inhibition of the other.

Synthesis of glucose from noncarbohydrate precursors (gluconeogenesis) consists of a series of reactions that proceed from pyruvate to oxaloacetate to glucose. Some readily reversible reactions of glycolysis form part of gluconeogenesis as well. Three strongly exergonic glycolytic reactions cannot be simply reversed and are bypassed by means of different reactions. Glycolysis and gluconeogenesis are reciprocally regulated. Fructose 2,6-bisphosphate serves as an important allosteric effector for both pathways.

SELECTED READINGS

Beylot, M., Soloviev, M. V., David, F., Landau, B. R., and Brunengraber, H., Tracing hepatic gluconeogenesis relative to citric acid cycle activity *in vitro* and *in vivo, J. Biol. Chem.* 270:1509–1514 (1995).

DiDonato, L., Des Rosiers, C., Montgomery, J. A., David, F., Garneau, M., and Brunengraber, H., Rates of gluconeogenesis and citric acid cycle in perfused livers, assessed from the mass spectrometric assay of the carbon-13-labeling pattern of glutamate, *J. Biol. Chem.* 268:4170–4180 (1993).

Hanson, R. W., and Patel, Y. M., Phosphoenolpyruvate carboxykinase (GTP): The gene and the enzyme, *Adv. Enzymol. Relat. Areas Mol. Biol.* 69:203–281 (1994).

Hardie, D. G., *Biochemical Messengers,* Chapman & Hall, London (1991).

Kyriakis, J. M., and Avruch, J., Sounding the alarm: Protein kinase cascades activated by stress and inflammation, *J. Biol. Chem.* 271:24313–24316 (1996).

Lebioda, L., and Stec, B., Crystal structure of enolase indicates that enolase and pyruvate kinase evolved from a common ancestor, *Nature (London)* 333:683–686 (1988).

Leschine, S. B., Cellulose degradation in anaerobic environments, *Annu. Rev. Microbiol.* 49:399–426 (1995).

Lienhard, G. E., Slot, J. W., and Mueckler, M. M., How cells absorb glucose, *Sci. Am.* 266(1):86–91 (1992).

Pilkis, S. J., El-Maghrabi, M. R., and Claus, T. H., Hormonal regulation of hepatic gluconeogenesis and glycolysis, *Annu. Rev. Biochem.* 57:755–784 (1988).

Pilkis, S. J., Claus, T. H., Kurland, I. J., and Lange, A. J., 6-Phosphofructo-2-kinase/fructose-2,6-bisphosphatase: A metabolic signaling enzyme, *Annu. Rev. Biochem.* 64:799–835 (1995).

Smythe, C., and Cohen, P., The discovery of glycogenin and the priming mechanism for glycogen biosynthesis, *Eur. J. Biochem.* 200:625–631 (1991).

Strader, C. D., Fong, T. M., Tota, M. R., and Underwood, D., Structure and function of G protein-coupled receptors, *Annu. Rev. Biochem.* 63:101–132 (1994).

Taylor, S. S., Buechler, J. A., and Yonemoto, W., cAMP-dependent protein kinase: Framework for a diverse family of regulatory enzymes, *Annu. Rev. Biochem.* 59:971–1005 (1990).

Wood, T., *The Pentose Phosphate Pathway,* Academic Press, Orlando (1985).

REVIEW QUESTIONS

A. Define each of the following terms:

Glycolysis	Gluconeogenesis
Schiff base	Phosphorolysis
Converter enzyme	G protein
Galactosemia	Bypass
Mutase	Cyclic AMP
"Fight or flight" response	Glycogen storage disease
Futile cycle	Nucleotide-linked sugar
Adenylate cyclase	Enolase

B. Differentiate between the two terms in each of the following pairs:

Transaldolase/transketolase	Hexokinase/pyruvate kinase
Pasteur effect/Crabtree effect	Glycogenesis/glycogenolysis
Cori cycle/glucose-alanine cycle	Bypass I/bypass III
Phosphofructokinase/phosphoglucoisomerase	Phosphoglyceromutase/phosphoglucomutase
Branching enzyme/debranching enzyme	Class I aldolase/class II aldolase

C. (1) Write balanced equations for the 10 individual reactions of glycolysis and name each enzyme. What is the overall reaction for (a) stage I and (b) stage II?
(2) Discuss the conditions under which carbohydrates are (a) transported as glucose, (b) transported as lactate, (c) subjected to glycogenesis, and (d) subjected to glycogenolysis. What hormones control the level of glucose in the blood and what are their effects?
(3) Outline the operation of the pentose phosphate pathway. Distinguish between the oxidative and nonoxidative phases, and describe the pathway's functions.
(4) Discuss the mechanism of action of the following enzymes: (a) aldolase; (b) glucose 6-phosphate dehydrogenase; (c) phosphorylase; (d) debranching enzyme.
(5) Why can gluconeogenesis not be an exact reversal of glycolysis? How do the two pathways differ and how are they regulated? What are the energy requirements of gluconeogenesis?
(6) What are G proteins and how do they function?

PROBLEMS

10.1. Would strenuous exercise (partially anaerobic conditions) tend to exacerbate or diminish the effect of ethanol on gluconeogensis? Explain your answer.

10.2. Given that

$$ATP^{4-} + H_2O \rightleftarrows ADP^{3-} + P_i^{2-} + H^+ \qquad \Delta G^{\circ\prime} = -30.5 \text{ kJ mol}^{-1}$$

calculate $\Delta G^{\circ\prime}$ for the following reactions (see Table 10.1):

(a) Phosphoenolpyruvate^{3-} + H_2O ⇄ pyruvate$^-$ + P_i^{2-}
(b) Glucose + P_i^{2-} ⇄ glucose 6-phosphate^{2-} + H_2O

10.3.* Calculate the *net* yield of ATP molecules per molecule of metabolite (underlined) for the following cases. The reactions occur under anaerobic conditions.

(a) Glycolysis of <u>fructose</u> to lactate in liver
(b) Glycolysis of <u>fructose</u> to lactate in kidney
(c) Glycolysis of <u>mannose</u> to pyruvate
(d) Glycolysis of <u>3-phosphoglycerate</u> to pyruvate

10.4.* Calculate the *net* yield of ATP molecules per molecule of sucrose for anaerobic glycolysis of sucrose to lactate, given the following:

$$\text{Sucrose} + H_2O \rightarrow \text{glucose} + \text{fructose}$$

$$\text{Fructose} + ATP^{4-} \rightarrow \text{fructose 6-phosphate}^{2-} + ADP^{3-} + H^+$$

10.5. To evaluate the energy savings achieved by phosphorylase, assume that, in the absence of phosphorylase, cleavage of each glucose residue from glycogen required hydrolysis of one molecule of ATP. On that basis, calculate the number of ATP molecules required to degrade 1.0 g of glycogen to glucose. Assume that the molecular weight of any glucose residue is 162.

10.6. Explain why a high K_m for pyruvate and a low k_{cat} for conversion of pyruvate to lactate (Table 10.2) for the H_4 isozyme of lactate dehydrogenase indicate that conversion of pyruvate to acetyl coenzyme A is favored over conversion of pyruvate to lactate.

10.7. What is the ratio of 1,3-bisphosphoglycerate to 3-phosphoglycerate under biochemical standard conditions when the [ATP]/[ADP] ratio is 10.0?

10.8.* A classmate has proposed that the molecular basis

of the Crabtree effect involves an allosteric effect of glucose on a specific enzyme. What effect would you look for in the following enzymes to see if the suggested hypothesis has any merit? (a) Lactate dehydrogenase (LDH); (b) cytochrome oxidase (a key enzyme of the electron transport system); (c) pyruvate carboxylase (converts pyruvate to oxaloacetate, an intermediate of the citric acid cycle)

10.9. The maximum velocity of muscle glycogen phosphorylase is large, whereas that of the liver enzyme is small. How can you explain this difference in maximum velocity?

10.10. The extreme toxicity of methanol arises not so much from methanol itself but rather from its metabolic conversion to formaldehyde by the action of alcohol dehydrogenase. Part of the treatment for methanol poisoning entails administering large doses of ethanol. Why is this treatment effective?

10.11. A student uses the alcoholic fermentation of yeast to prepare a wine that has an alcohol content of 5.0% w/v (5.0 g of ethanol/100 ml of wine). What is the minimum molar concentration of glucose the fermentation mixture must contain? (MW of ethanol = 46)

10.12.* Four samples of glucose are selectively labeled with ^{14}C at one carbon atom: (a) at C(1); (b) at C(3); (c) at C(4); and (d) at C(6). Subjecting the four samples to anaerobic glycolysis results in conversion of the glucose to lactate. Which carbon atoms of lactate will become labeled in each case?

10.13.* The glucose samples from the previous problem are converted to glucose 6-phosphate by action of hexokinase, and the glucose 6-phosphate is passed once through the pentose phosphate pathway. Under these conditions, which samples would yield ^{14}C-labeled CO_2?

10.14. Consider two hypothetical carbohydrates, a 10-carbon 2-ketose and an 8-carbon aldose. What carbohydrates would you obtain when you treat the two compounds with (a) transaldolase and (b) transketolase?

10.15. Energetically speaking, how many molecules of glucose could be synthesized via gluconeogenesis from the energy released by hydrolysis of 90 molecules of ATP to ADP, followed by hydrolysis of ADP to AMP?

10.16. How many molecules of ATP would be required if the starting material for the pentose phosphate pathway consisted of six molecules of glucose rather than of glucose 6-phosphate?

10.17. In the absence of glycogen breakdown, the hexokinase reaction constitutes the rate-determining step of glycolysis. Would this be true in the presence of glycogen breakdown as well? Explain.

10.18. Distinguish between the rate-determining step (hexokinase) and the committed step (phosphofructokinase) of glycolysis.

10.19. When you add phosphoglucomutase to an equilibrium mixture of glucose 1-phosphate and glucose 6-phosphate, no net reaction takes place. If you were to use ^{32}P-labeled glucose 1-phosphate in this experiment, would some of the label exchange into glucose 6-phosphate? If so, why?

10.20. Why must the NADH produced in glycolysis be oxidized to regenerate NAD^+, regardless of whether the system as a whole is aerobic or anaerobic?

The Citric Acid Cycle

11

The citric acid cycle, together with the electron transport system, constitutes stage III of catabolism (see Figure 8.2), also called *cellular respiration.* Because the citric acid cycle functions in both catabolism (Figure 11.1) and anabolism (Figure 11.2), we call it an **amphibolic pathway.** The citric acid cycle plays a pivotal role in cellular respiration, has multiple interconnections with other pathways, and provides for interconversions of numerous metabolites.

We use the term **citric acid cycle** because citrate is the first compound produced in the cyclic set of reactions. The pathway is also known as the **Krebs cycle** in honor of Hans Krebs, the British biochemist who proposed it in 1937. (Krebs was awarded the Nobel Prize in 1953.) Lastly, we refer to the reaction sequence as the **tricarboxylic acid (TCA) cycle** because the first few intermediates are tricarboxylic acids. The citric acid cycle has three major roles:

1. It is the central oxidative pathway by which all nutrients—carbohydrates, lipids, and proteins—are catabolized in aerobic organisms and tissues.
2. It is an important source of intermediates for a large number of anabolic pathways leading to the biosynthesis of a variety of biomolecules.
3. It is the main source of metabolic energy, which derives from the oxidation–reduction steps of the cycle and their linkage to the electron transport system.

Krebs' proposal of the citric acid cycle followed a number of important findings dealing with cellular respiration. In 1935, Albert Szent-Györgyi showed that adding catalytic amounts of succinate, fumarate, malate, or oxaloacetate to minced muscle tissue greatly stimulated cellular respiration. Oxygen uptake and carbon dioxide production increased beyond the quantities required for oxidation of the added compound. In addition, Szent-Györgyi established that the compounds were interconverted according to the sequence succinate → fumarate → malate → oxaloacetate. Shortly thereafter, Carl Martius and Franz Knoop demonstrated the sequence citrate → *cis*-aconitate → isocitrate → α-keto-

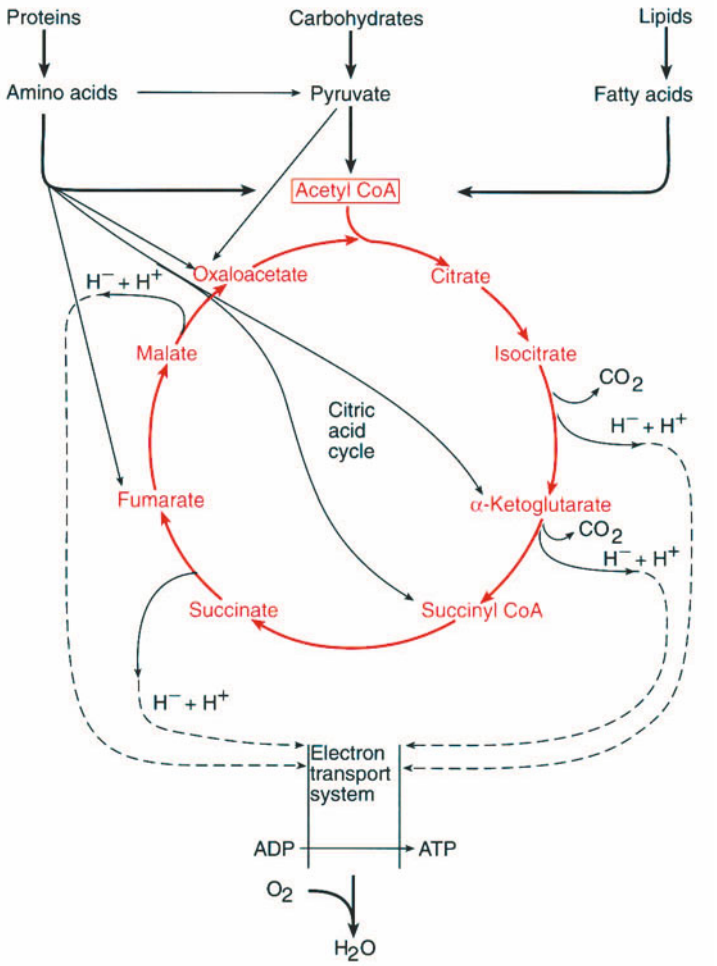


Figure 11.1. The citric acid cycle in catabolism.

glutarate. Since scientists already knew that α-ketoglutarate could be decarboxylated to succinate, they concluded that a pathway extends from citrate to oxaloacetate. Based on this information and on his own extensive studies, Krebs proposed a cyclic set of reactions consisting of the above sequences that was closed by the conversion of oxaloacetate to citrate.

The contribution of Krebs ranks as one of the most important milestones of metabolic biochemistry. Ironically, when he first submitted his manuscript for publication, it was rejected for lack of publishing space.

The citric acid cycle represents one of three catalytic cycles Krebs proposed. In 1932, Krebs and Kurt Henseleit elucidated the *urea cycle* (Section 14.3), the first known metabolic cycle. In 1957, Krebs and H. R. Kornberg proposed a modifed form of the citric acid cycle, called the *glyoxylate cycle* (Section 11.6).

Although Krebs established the existence of the citric acid cycle, the precise mechanism of citrate formation remained unknown for several years. How oxaloacetate leads to citrate became clear only after Nathan Kaplan and Fritz Lipmann discovered *coenzyme A* (1945) and Severo Ochoa and Feodor Lynen showed that acetyl coenzyme A condenses with oxaloacetate to form citrate (1951).

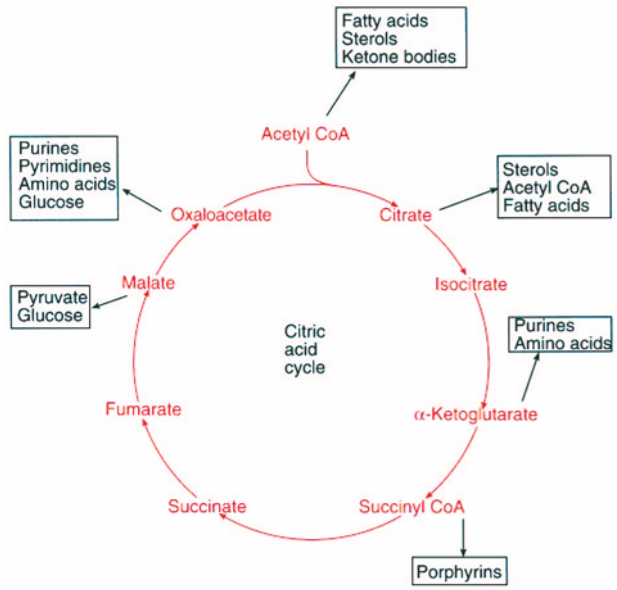


Figure 11.2. The citric acid cycle in anabolism.

11.1. COENZYMES OF THE CYCLE

A key feature of the citric acid cycle consists of four oxidation–reduction reactions, catalyzed by specific dehydrogenases that use either nicotinamide adenine dinucleotide (NAD^+) or flavin adenine dinucleotide (FAD) as coenzyme. As the metabolites undergo oxidation, the coenzymes undergo reduction. In these reactions, the coenzyme functions like a second substrate. Protons and electrons derived from the metabolites reduce NAD^+ and FAD to NADH and $FADH_2$. The reduced coenzymes are subsequently oxidized by the electron transport system with the concomitant synthesis of ATP in a process called oxidative phosphorylation (Section 12.4). In this way, protons and electrons removed from metabolites become channeled into the electron transport system to generate metabolic energy in the form of ATP.

11.1.1. Pyridine-Linked Dehydrogenases

We use the term **pyridine-linked dehydrogenases** to describe enzymes that use either *nicotinamide adenine dinucleotide (NAD^+)* or the related compound *nicotinamide adenine dinucleotide phosphate ($NADP^+$)* as coenzyme. Some pyridine-linked dehydrogenases use NAD^+, and others use $NADP^+$. The term derives from the fact that a pyridine nucleus (a nitrogen-containing six-membered ring) forms part of the coenzyme molecule (Figure 11.3). You have already encountered a number of such enzymes in Chapter 10.

Because the structures of NAD^+ and $NADP^+$ bear some similarity to nucleic acid building blocks, the coenzymes' names include the word "nucleotide." NAD^+ and $NADP^+$ are loosely bound to their respective dehydrogenases and easily dissociate from them. They do *not* constitute prosthetic groups.

We call NAD^+ and $NADP^+$ *pyridine nucleotide coenzymes.* The pyridine ring forms part of the coenzyme's **nicotinamide** moiety. Nicotinamide is a derivative of *niacin (nicotinic acid),* a water-soluble vitamin and member of the vitamin B complex. Like other water-soluble vitamins, niacin is required for assembly of a specific coenzyme molecule, in this case NAD^+ and $NADP^+$.

Niacin is unique in that it can be synthesized in many mammals from the amino acid tryptophan. The extent to which the dietary requirement for niacin can be decreased by dietary tryptophan varies according to the species. The deficiency disease associated with niacin is *pellagra.*

Oxidation–reduction reactions catalyzed by dehydrogenases typically involve removal of two hydrogen atoms from a substrate, as in the oxidation of ethanol to acetaldehyde catalyzed by alcohol dehydrogenase:

$$CH_3\text{-}CH_2OH + NAD^+ \longrightarrow CH_3\text{-}CHO + NADH + H^+$$

Two hydrogen atoms (2H·) are formally equivalent to two protons and two electrons ($2H^+ + 2e^-$). However, oxidation–reduction reactions catalyzed by dehydrogenases re-

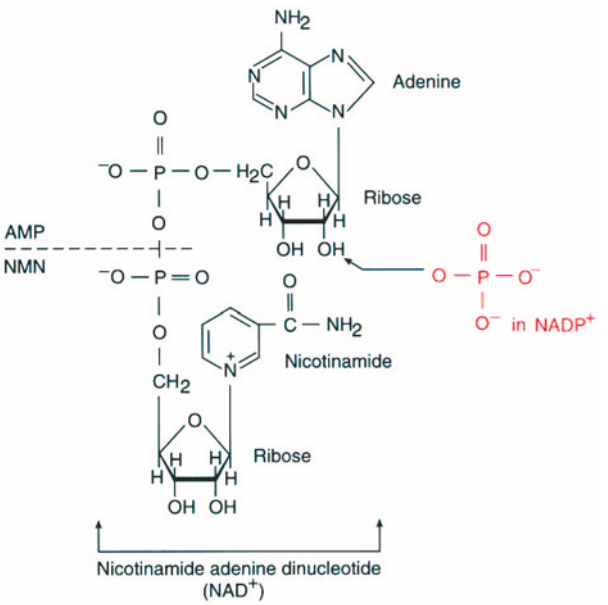
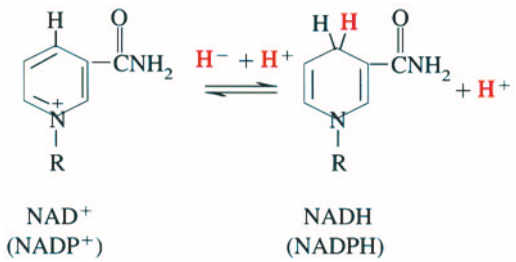


Figure 11.3. Pyridine nucleotide coenzymes. Nicotinamide adenine dinucleotide (NAD^+) consists of adenosine 5′-monophosphate (AMP) linked to nicotinamide mononucleotide (NMN). The structure of $NADP^+$ is identical to that of NAD^+ except for an additional phosphate group at C(2′) of the ribose in AMP.

quire transferring these protons and electrons in the form of hydride ions (H^- or H:) and protons. Because a hydride ion consists of a proton and two electrons ($H^+ + 2e^-$), we have the following equivalence:

$$2H\cdot = H^- + H^+ = 2H^+ + 2e^-$$

As the substrate of a dehydrogenase is oxidized by removal of hydrogen, the enzyme's coenzyme is reduced. In pyridine-linked dehydrogenases, niacin forms the "business end" of the coenzyme; the nicotinamide portion of NAD^+ and $NADP^+$ undergoes reduction (or oxidation) as a metabolite undergoes oxidation (or reduction). In this process, the pyridine nucleus gains (or loses) a hydride ion, and a proton appears in (or disappears from) the medium:



We can follow the progress of pyridine-linked dehydrogenase reactions by means of spectrophotometric measurements, since NADH and NADPH, but not NAD^+ and $NADP^+$, absorb strongly at 340 nm (Figure 11.4). X-ray diffraction studies have shown that the overall molecular structures of pyridine-linked dehydrogenases differ, but their NAD^+ binding sites have similar architectures. Each binding site consists of two parts, one specific for the nicotinamide portion, and the other for the adenine moiety.

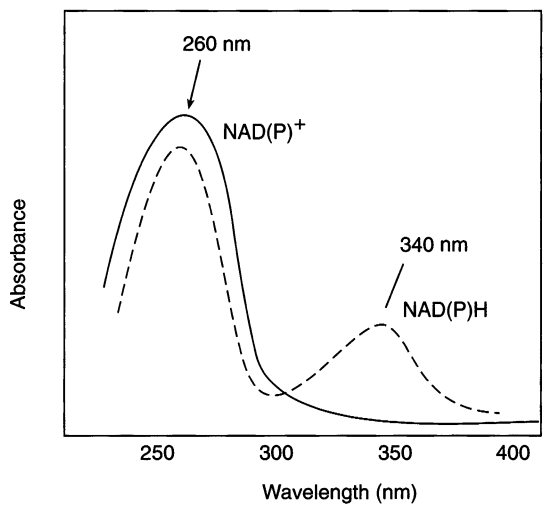


Figure 11.4. Absorption spectra of $NAD(P)^+$ and NAD(P)H.

11.1.2. Flavoproteins

Pyridine-linked dehydrogenases, using NAD^+ as coenzyme, catalyze three of the oxidation–reduction reactions of the citric acid cycle. The fourth is catalyzed by a dehydrogenase using FAD as coenzyme. We refer to *flavin adenine dinucleotide (FAD)* and the related compound *flavin mononucleotide (FMN)* as *flavin coenzymes* (Figure 11.5). Unlike NAD^+ and $NADP^+$, FAD and FMN are tightly bound to their specific dehydrogenases and constitute *prosthetic groups*. Accordingly, we call the complete enzymes **flavoproteins.** Some dehydrogenases use FAD, and others use FMN. Like NAD^+ and $NADP^+$, FAD and FMN have some structural similiarity to nucleic acid building blocks.

Flavin coenzymes contain *vitamin B_2 (riboflavin)* as a structural component. Riboflavin, like niacin, constitutes the "business end" of the coenzyme. Complete reduction of the *isoalloxazine* ring system of FAD and FMN, as distinct from the reduction of NAD^+ or $NADP^+$, requires both a hydride ion and a proton. Hence, oxidation of a metabolite (MH_2) by a flavoprotein does *not* yield a proton as a reaction product:

$$\underset{}{MH_2} + \underset{(FAD)}{FMN} \rightleftharpoons M + \underset{(FADH_2)}{FMNH_2}$$

Reduction of flavins differs from that of pyridine nucleotide coenzymes in yet another aspect. Whereas NAD^+ and $NADP^+$ undergo only two-electron reductions, FMN and FAD can undergo either one- or two-electron reductions. Both coenzymes can participate in two sequential one-electron transfers or in a simultaneous two-electron transfer (Figure 11.6). When flavins undergo a one-electron reduction, a relatively stable free radical forms. The redox versatility of flavins comes into play in the electron transport system, in which FMN serves as a link between electron carriers requiring two- and one-electron transfers.

In addition to their roles in the citric acid cycle, pyridine nucleotide coenzymes and flavin coenzymes also mediate a variety of other biological oxidation–reduction reactions.

11.1.3. Coenzyme A

In addition to NAD^+ and FAD, three other coenzymes—coenzyme A, lipoic acid, and thiamine pyrophosphate—participate in the citric acid cycle.

Next to ATP, **coenzyme A** and its derivative **acetyl coenzyme A** (Figure 11.7) are probably the most important low-molecular-weight biomolecules of metabolism. Coenzyme A (the *A* stands for *acetylation*) activates and carries acetyl and other acyl groups. It functions in reactions that require transfers of these groups.

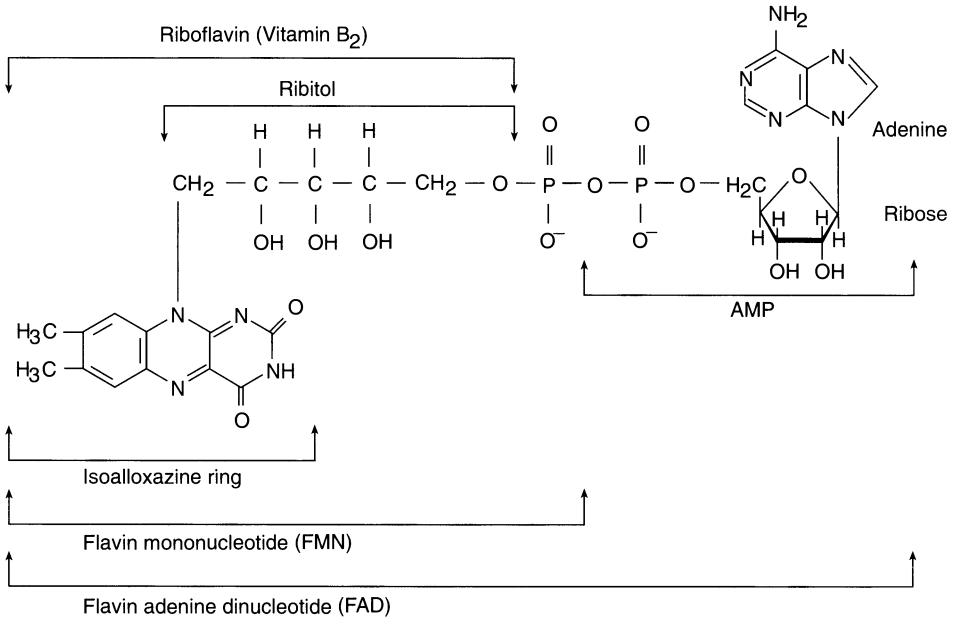


Figure 11.5. Flavin coenzymes.

In some ways, coenzyme A has structural features similar to those of NAD^+ and FAD. It, too, has a nucleotide-like component, contains a pyrophosphate group, and includes a vitamin. Of the entire molecule, however, the comparatively small sulfhydryl group constitutes the "business end" of the coenzyme.

We abbreviate coenzyme A as *CoA* or as *CoA-SH* to emphasize the functional importance of the SH group. Likewise, we abbreviate acetyl coenzyme A as *acetyl CoA* or as CH_3CO-*S*-*CoA*. The vitamin component of coenzyme A is *pantothenic acid,* a water-soluble B vitamin. Deficiencies of pantothenic acid in humans occur only rarely, probably because of the widespread distribution of the vitamin in natural foods.

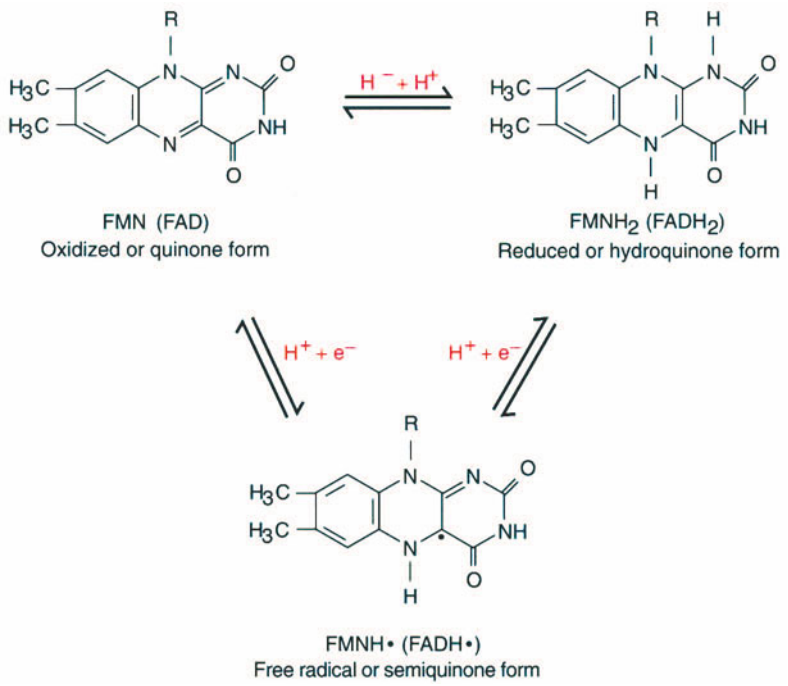


Figure 11.6. Oxidation and reduction of flavin coenzymes.

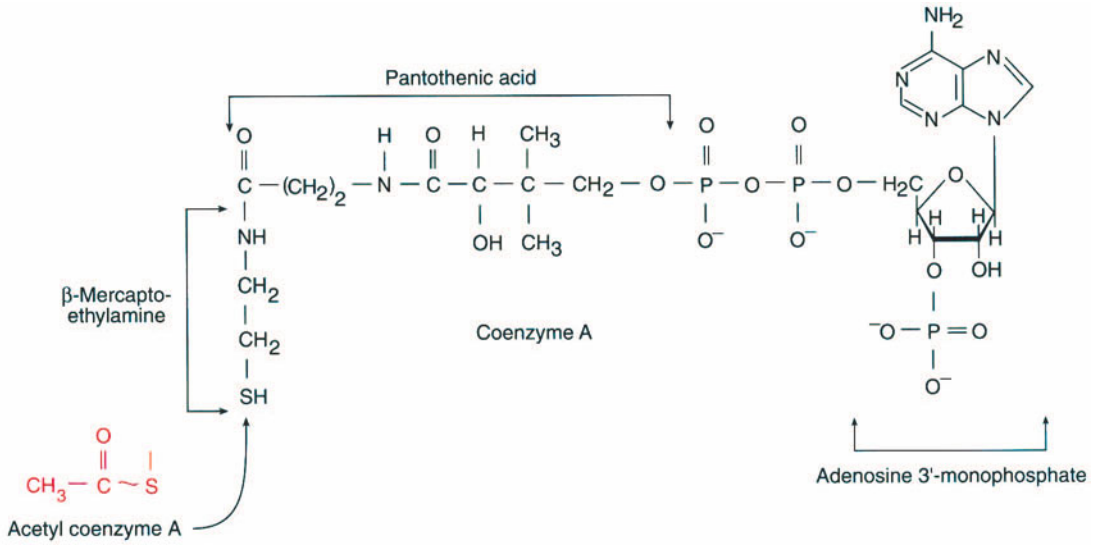


Figure 11.7. Coenzyme A and acetyl coenzyme A. The thioester bond in acetyl coenzyme A (~) is an energy-rich bond.

Acetyl CoA constitutes not only a key coenzyme but also an *energy-rich compound.* It belongs to the class of special esters, specifically that of *thioesters,* in which sulfur replaces an oxygen atom in the ester bond. Any *acyl CoA* compound, such as *succinyl CoA,* one of the intermediates in the citric acid cycle, constitutes an energy-rich compound. Hydrolysis of the thioester bond (~) has a highly negative free energy change associated with it:

$$CH_3CO \sim S\text{-}CoA^{4-} + H_2O \rightarrow CH_3COO^- + CoA\text{-}SH^{4-} + H^+ \qquad \Delta G^{\circ\prime} = -31.4 \text{ kJ mol}^{-1}$$

11.1.4. Lipoic Acid

Lipoic acid functions in acyl transfer reactions. The coenzyme has a heterocyclic, sulfur-containing ring that carries an aliphatic side chain (Figure 11.8A). Lipoic acid occurs in tissues in extraordinarily small amounts. It was first isolated in 1949 (Eli Lilly Co.). By using some *10 tons* of beef liver as starting material, investigators were able to obtain approximately *30 mg* of lipoic acid!

Nutritionists generally classify lipoic acid with the B vitamins because it is a growth factor for some microorganisms. The coenzyme form of the vitamin is **lipoamide (lipoyllysine),** a prosthetic group formed by an amide linkage between the carboxyl group of lipoic acid and the ϵ-amino group of a lysine residue in the enzyme (Figure 11.8B).

11.1.5. Thiamine Pyrophosphate

Thiamine, or *vitamin* B_1, contains two heterocyclic ring systems—a pyrimidine nucleus and a sulfur- and nitrogen-containing ring called a *thiazole* (Figure 11.9). The coenzyme form of thiamine, **thiamine pyrophosphate,** forms by attachment of a pyrophosphate group to the thiazole nucleus. We abbreviate thiamine pyrophosphate as *ThPP* or *TPP,* not to be confused with TTP (thymidine triphosphate). Thiamine pyrophosphate functions in aldehyde transfer reactions and as coenzyme for enzymes catalyzing the decarboxylation of α-keto acids. TPP also serves as coenzyme of transketolases. The deficiency disease associated with vitamin B_1 is *beriberi.*

11.2. SYNTHESIS OF ACETYL COENZYME A

The citric acid cycle consists of a series of eight reactions (Figure 11.10). Initiation of each round of the cycle involves the entrance of a molecule of *acetyl CoA,* which condenses with oxaloacetate to form citrate. Acetyl CoA forms via two major routes from either lipids or carbohydrates. Lipid catabolism yields acetyl CoA directly (Sec-

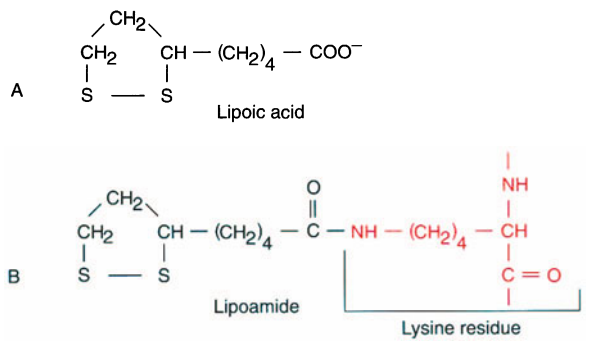


Figure 11.8. (A) Lipoic acid. (B) Lipoamide (lipoyllysine). In lipoamide, lipoic acid is linked covalently to the ϵ-NH_2 group of a lysine residue in an enzyme.

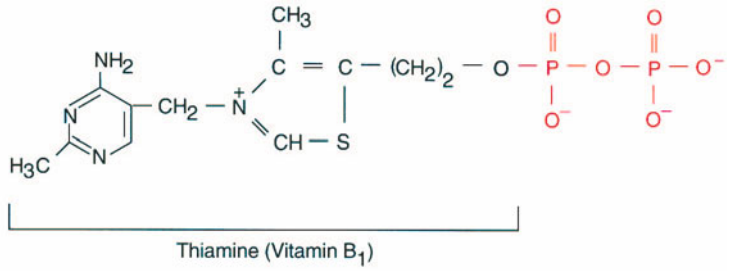


Figure 11.9. Thiamine pyrophosphate. The pyrophosphate group is esterified to a primary alcohol group in thiamine.

tion 13.3). Carbohydrate catabolism produces acetyl CoA by way of pyruvate. Recall that pyruvate constitutes the end product of glycolysis. Following its formation in the cytoplasm, pyruvate moves across the mitochondrial membrane into the matrix via an antiport transport system. In this system, pyruvate transport is coupled to hydroxide-ion transport in the opposite direction. A multienzyme system called the **pyruvate dehydrogenase complex** and located in the mitochondrial matrix then catalyzes the conversion of pyruvate to acetyl CoA.

The pyruvate dehydrogenase complex consists of multiple copies of three enzymes and five coenzymes. The three enzymes are *pyruvate dehydrogenase* (E_1), *dihydrolipoyl transacetylase* (E_2), and *dihydrolipoyl dehydrogenase,* a flavoprotein (E_3). A cubic cluster of E_2 forms the "core" of the complex, with E_1 and E_3 being assembled on the core's surface.

The enzymes form a very large multienzyme system. The *E. coli* pyruvate dehydrogenase complex has a molecular weight of 4.6×10^6 and consists of 24 polypeptide chains of E_1, 24 chains of E_2, and 12 chains of E_3. The mammalian complex has a molecular weight of 8.4×10^6 and contains 120 polypeptide chains of E_1, 60 chains of E_2, and 12 chains of E_3.

Three of the five coenzymes of the pyruvate dehydrogenase complex function as catalysts, and each is associated with a component enzyme: TPP is linked to E_1; lipoic acid is covalently linked to E_2; and FAD forms an integral part of E_3. The remaining two coenzymes, coenzyme A and NAD^+, function as substrates and undergo chemical alteration in the course of the reaction. The overall reaction catalyzed by the pyruvate dehydrogenase complex comprises both an oxidation (dehydrogenation) and a decarboxylation (loss of CO_2). Accordingly, we refer to it as an **oxidative decarboxylation.** The reaction has the following stoichiometry:

$$\text{Pyruvate}^- + \text{CoA}{-}\text{SH}^{4-} + \text{NAD}^+ \xrightarrow[\text{lipoic acid}]{\text{FAD, TPP}} \text{acetyl CoA}^{4-} + \text{CO}_2 + \text{NADH}$$

Note that we place three coenzymes above or below the arrow to indicate their catalytic nature; they are regenerated in their original form at the end of the reaction. The overall reaction is strongly exergonic ($\Delta G^{\circ\prime} = -33.5$ kJ mol^{-1}).

Lipoic acid, attached via an amide bond to a lysine residue of E_2, represents a key feature of the reaction mechanism (Figure 11.11). The resultant **lipoyllysine arm** has a fully extended length of about 14 Å. Researchers believe that this arm functions as a tether, swinging between participating sites on E_1 and E_3. They postulate that at least two arms bind to each E_2 such that the acetyl group can be transferred from one arm to the other.

ATP inhibits the pyruvate dehydrogenase complex, and AMP activates it. In addition, the enzyme system is subject to product inhibition by both NADH and acetyl CoA. These regulatory effects make good metabolic sense. Acetyl CoA initiates the reactions of the citric acid cycle. Oxidative steps of the cycle yield the reduced coenzymes NADH and $FADH_2$; these are oxidized via the electron transport system with the concomitant synthesis of ATP. Because high concentrations of ATP or NADH result in inhibition of the pyruvate dehydrogenase complex, these conditions prevent synthesis of additional and unneeded ATP by the electron transport system. Conversely, low ATP levels and high AMP levels activate the enzyme system, resulting in enhanced ATP synthesis. High concentrations of acetyl CoA inhibit the enzyme system and block synthesis of additional acetyl CoA. In eukaryotes, additional regulation involves covalent modification of E_1; phosphorylated E_1 is inactive whereas the dephosphorylated form is active.

11.3. REACTIONS OF THE CYCLE PROPER

Each of the eight reactions of the citric acid cycle is enzyme-catalyzed. In eukaryotes, all of the reactions take place in the mitochondria.

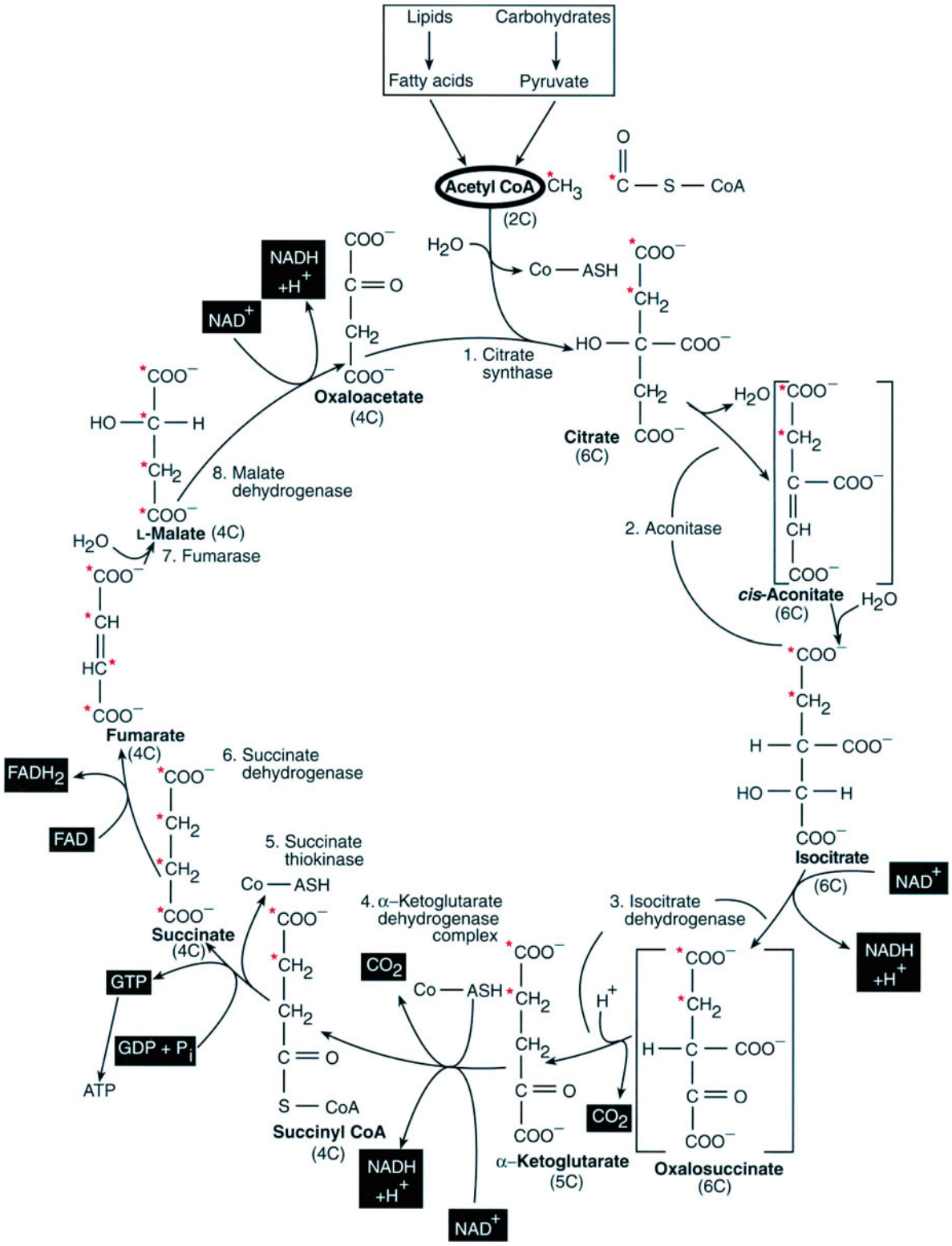


Figure 11.10. The citric acid cycle. Bracketed compounds designate enzyme-bound intermediates. When carbon atoms entering as acetyl CoA are labeled (asterisk), the label becomes scrambled between the two symmetrical halves of succinate. Fate of the label is shown up to malate. The two molecules of CO_2 that are lost do not come from the acetyl CoA that has just entered the cycle but rather from the oxaloacetate formed during the preceding turn.

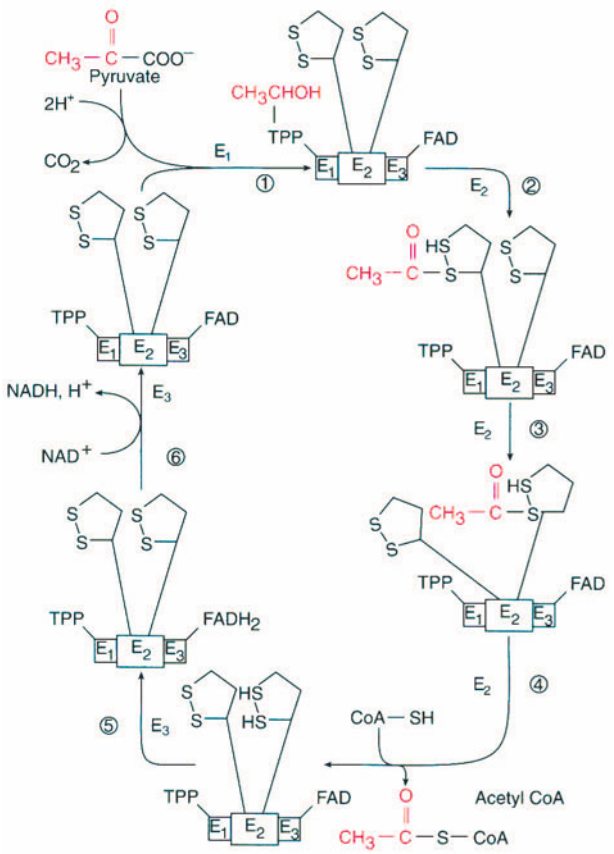
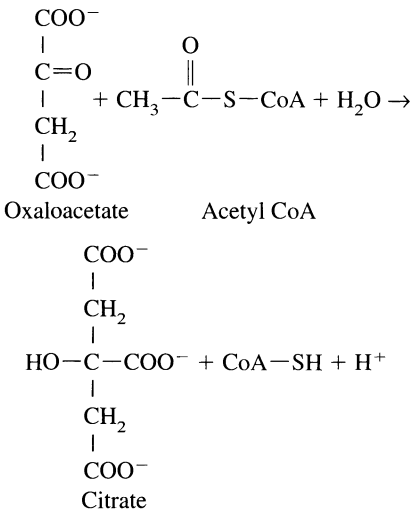


Figure 11.11. Postulated mechanism of the pyruvate dehydrogenase complex. (1) A two-carbon fragment is transferred from pyruvate to E_1; (2) the fragment is transferred to one lipoyllysine arm and simultaneously oxidized to an acetyl group; (3) the lipoyllysine arms swings to transfer the acetyl group to a second arm; (4) the second arm positions the acetyl group for transfer to coenzyme A; (5) E_3 catalyzes the oxidation of the reduced lipoyllysine arm; (6) the reduced flavin is oxidized with NAD^+, forming NADH and H^+.

11.3.1. Citrate Synthase

The first reaction of the cycle proper involves a *condensation* between the acetyl group of acetyl CoA (a two-carbon fragment) and **oxaloacetate** (a four-carbon compound) to form **citrate** (a six-carbon compound). **Citrate synthase** catalyzes this reaction:



The citrate synthase reaction is strongly exergonic (Table 11.1) owing to cleavage of the *energy-rich bond* in acetyl CoA. Citrate synthase is an allosteric enzyme for which ATP, NADH, and succinyl CoA serve as negative effectors. The mammalian enzyme is a dimer, consisting of two identical subunits (MW = 49,000 each).

The citrate synthase reaction represents one of three major regulatory sites of the citric acid cycle. When concentrations of the enzyme's negative allosteric effectors build up, the cycle sequence becomes inhibited. Because

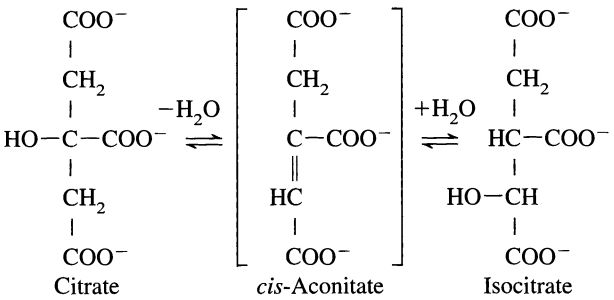
Table 11.1. Reactions of the Citric Acid Cycle

Enzyme	Mitochondrial location	Type of reaction	$\Delta G^{\circ\prime}$ (kJ mol^{-1})
1. Citrate synthase	Matrix	Condensation	−32.2
2. Aconitase	Matrix	Isomerization	+5.0
3. Isocitrate dehydrogenase	Matrix	Oxidative decarboxylation	−20.9
4. α-Ketoglutarate dehydrogenase complex	Matrix	Oxidative decarboxylation	−33.5
5. Succinate thiokinase (succinyl CoA synthase)	Matrix	Hydrolysis/ phosphorylation	−3.3
6. Succinate dehydrogenase	Inner membrane	Dehydrogenation	~0
7. Fumarase	Matrix	Hydration	−3.8
8. Malate dehydrogenase	Matrix	Dehydrogenation	+29.7
		Overall	−59.0

of the cycle's link to the electron transport system, inhibition of citrate synthase results in decreased ATP synthesis. Thus, as in the case of the pyruvate dehydrogenase complex, high concentrations of ATP and NADH prevent synthesis of unneeded ATP.

11.3.2. Aconitase

The **aconitase** reaction has two steps, dehydration of citrate followed by hydration. The intermediate, called ***cis*-aconitate,** remains enzyme-bound. Both steps are stereospecific and lead to formation of isocitrate. The overall reaction constitutes an isomerization of citrate to **isocitrate:**



Isomerization results in conversion of a hydroxyl group of a tertiary alcohol at C(3) to a hydroxyl group of a secondary alcohol at C(4), thereby setting the stage for the subsequent decarboxylation step. Aconitase is an iron–sulfur protein and contains a [4Fe-4S] cluster. The mammalian enzyme consists of two identical subunits (MW = 45,000 each).

Investigators have dubbed the postulated mechanism of action of aconitase "ferrous wheel." According to this concept, *cis*-aconitate binds to the enzyme at three sites, including the iron. Stereospecific addition of water, to form either citrate or isocitrate, involves partial rotation of this "ferrous wheel."

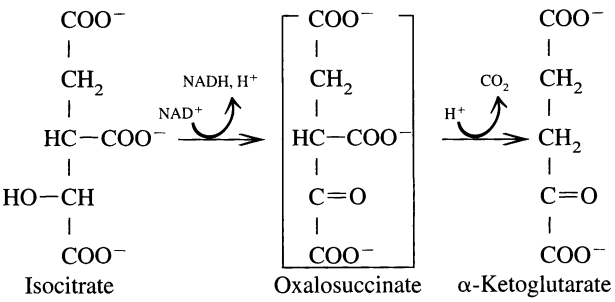
The aconitase reaction is of historical interest. When Ogston studied it in 1948, he found that the enzyme catalyzed a reaction with only one of the two identical parts of the symmetric citrate molecule. Consequently, only one of two possible chiral products formed. This finding led him to propose the *polyaffinity theory* for enzyme action (Section 4.2).

Aconitase constitutes the ultimate target for the inhibition produced by *fluoroacetate,* a compound that occurs in some South African plants. Fluoroacetate is one of the most toxic small molecules known. When ingested by animals, it is converted to fluoroacetyl CoA by *acetate thiokinase* (Figure 11.12). Citrate synthase can use fluoroacetyl CoA as an alternate substrate and catalyzes its conversion to *fluorocitrate,* a powerful inhibitor of aconitase. Fluorocitrate is used as a rodenticide. Its inhibition of aconitase is so strong that ingesting it can be lethal.

Fluoroacetate is an example of a **suicide substrate**—a substance that has no toxicity by itself but that is acted upon by an enzyme because of its similarity to a normal substrate. Enzymatic action converts the substance to a toxic form that inhibits a critical enzyme and causes cell death. The cell has died, or "committed suicide," as a result of a transformation of the originally nontoxic substance.

11.3.3. Isocitrate Dehydrogenase

Isocitrate dehydrogenase also catalyzes an *oxidative decarboxylation.* The enzyme is a pyridine-linked dehydrogenase that uses NAD^+ as coenzyme and forms **α-ketoglutarate,** a five-carbon compound:



The reaction proceeds through formation of an enzyme-bound intermediate, **oxalosuccinate,** so called be-

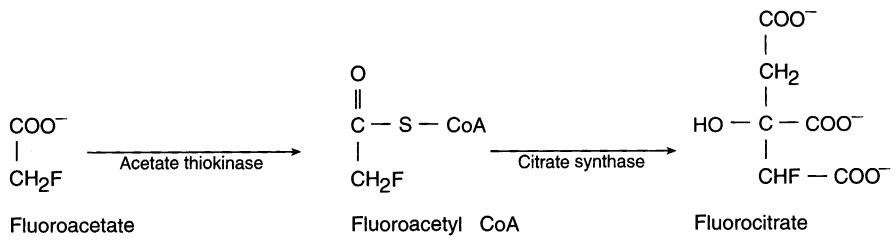


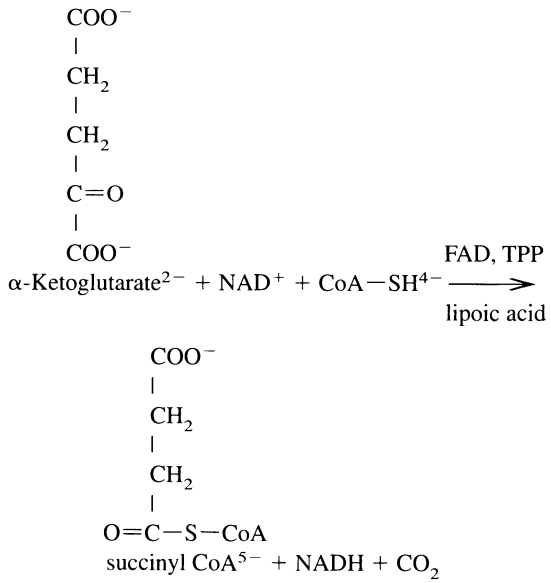
Figure 11.12. Conversion of fluoroacetate to fluorocitrate.

cause it has the structure of a condensation product between succinate and oxalate ($^{-}OOC\text{-}COO^{-}$). Isocitrate dehydrogenase is an allosteric enzyme, and its regulation constitutes the second major control point of the citric acid cycle. ATP and NADH inhibit the enzyme; ADP and NAD^+ activate it. As in the case of citrate synthase, these regulatory effects reflect the level of ATP in the cell. The mammalian enzyme is a tetramer ($\alpha_2\beta\gamma$) with a total molecular weight of about 190,000.

Most tissues contain two isocitrate dehydrogenases. The NAD^+-linked enzyme occurs only in mitochondria and functions in the citric acid cycle. A second, $NADP^+$-linked dehydrogenase occurs both in mitochondria and in the cytoplasm. Its main function appears to be generating reducing power in the form of NADPH for some cytoplasmic reactions.

11.3.4. α-Ketoglutarate Dehydrogenase Complex

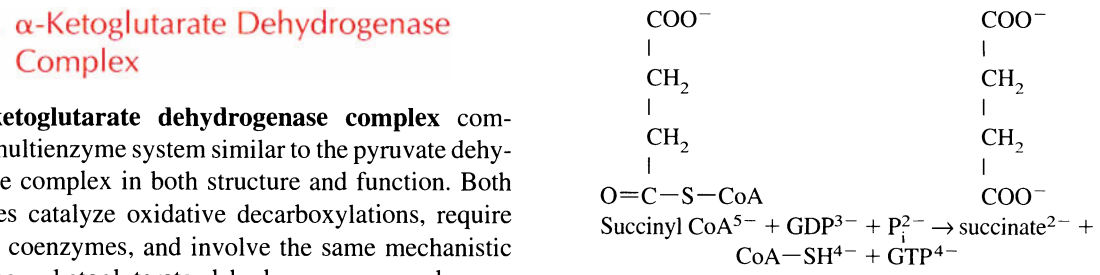
The **α-ketoglutarate dehydrogenase complex** comprises a multienzyme system similar to the pyruvate dehydrogenase complex in both structure and function. Both complexes catalyze oxidative decarboxylations, require the same coenzymes, and involve the same mechanistic steps. The α-ketoglutarate dehydrogenase complex accomplishes the shortening of α-ketoglutarate by one carbon, eliminated as CO_2. Because the product, **succinyl CoA,** has the structure of an acyl CoA, it constitutes an energy-rich compound:



The α-ketoglutarate dehydrogenase reaction represents the third major control point of the citric acid cycle. ATP and NADH inhibit the enzyme; ADP and NAD^+ activate it. As you saw for the other two regulatory sites of the citric acid cycle, activations and inhibitions reflect the intracellular level of ATP.

11.3.5. Succinate Thiokinase (Succinyl CoA Synthase)

Succinate thiokinase, or *succinyl CoA synthase,* catalyzes the cleavage of succinyl CoA to **succinate** and coenzyme A. Thus, the combined change for coenzyme A in this and the previous reaction equals zero. The exergonic cleavage of the energy-rich bond in succinyl CoA is coupled to the endergonic phosphorylation of guanosine 5′-diphosphate (GDP) to guanosine 5′-triphosphate (GTP):



Scientists believe that the reaction proceeds in three steps. In the first, succinyl phosphate forms from P_i and succinyl CoA. In the second step, the phosphate group is transferred to a histidine residue on the enzyme. Lastly, GTP forms by transfer of the phosphate group from the enzyme to GDP. Because the overall reaction involves breakage of one energy-rich bond and synthesis of another, the free energy change is close to zero. Production of GTP is equivalent to production of ATP, since the phosphate is readily transferred from GTP to ADP in a nucleotide exchange reaction, catalyzed by *nucleoside diphosphate kinase:*

$$GTP + ADP \rightleftarrows GDP + ATP$$

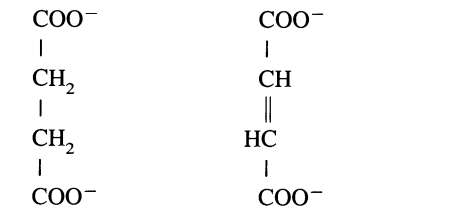
In plants and bacteria, ATP forms directly from succinyl CoA and ADP. Synthesis of GTP and ATP, coupled to cleavage of the energy-rich bond in succinyl CoA, represents another instance of *substrate-level phosphorylation,* similar to that occurring in glycolysis.

Note the difference between the pyruvate dehydrogenase complex and the α-ketoglutarate dehydrogenase complex as regards use of the energy contained in their respective acyl CoA compounds. Since the energy of acetyl CoA merely drives its condensation with oxaloacetate to form citrate, the energy appears to be "squandered." In re-

ality, it makes this reaction highly exergonic and thereby produces the committed step of the cycle. By contrast, the energy of succinyl CoA drives the synthesis of an ATP equivalent in the form of GTP. Hence, the energy of succinyl CoA is "saved" directly by forming an energy-rich compound.

11.3.6. Succinate Dehydrogenase

Succinate dehydrogenase catalyzes a *dehydrogenation* (i.e., oxidation) that converts succinate to **fumarate.** The enzyme, a flavoprotein, uses FAD as coenzyme. In this stereospecific reaction, fumarate—the *trans* isomer—forms with a 100% yield. Maleate, the *cis* isomer, does not form at all.

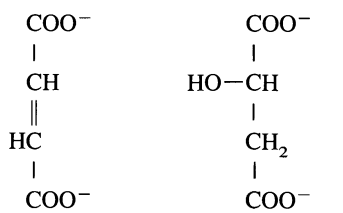


$$\text{Succinate}^{2-} + \text{FAD} \rightarrow \text{fumarate}^{2-} + \text{FADH}_2$$

Succinate dehydrogenase is an *iron–sulfur protein* (4Fe-4S type), located in the inner mitochondrial membrane, where it constitutes part of *respiratory complex II* (see Section 12.3). The mammalian enzyme is a dimer of two unequal subunits having molecular weights of 30,000 and 70,000. Because of its structural similarity to succinate, researchers have used *malonate* ($^{-}OOC-CH_2-COO^{-}$) as a competitive inhibitor of succinate dehydrogenase.

11.3.7. Fumarase

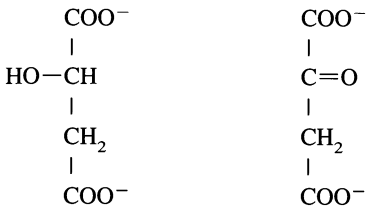
Fumarase, like succinate dehydrogenase, catalyzes a stereospecific reaction. In this reaction, only one of the two enantiomers of *malic acid* forms, namely, L-malate. The mammalian enzyme is a tetramer of four identical subunits, having a molecular weight of 49,000 each. Fumarase catalyzes a hydration reaction whereby water adds across the double bond of fumarate:



$$\text{Fumarate}^{2-} + H_2O \rightarrow \text{L-malate}^{2-}$$

11.3.8. Malate Dehydrogenase

Malate dehydrogenase catalyzes the last reaction of the citric acid cycle. The enzyme, a pyridine-linked dehydrogenase, uses NAD^+ as coenzyme. The mammalian enzyme is a dimer of identical subunits having a molecular weight of 35,000 each. Malate dehydrogenase catalyzes a *dehydrogenation* (i.e., oxidation) that converts malate to oxaloacetate:



$$\text{L-Malate}^{2-} + NAD^+ \rightarrow \text{oxaloacetate}^{2-} + NADH + H^+$$

Oxaloacetate can combine with a second molecule of acetyl CoA and start another round of the cycle.

11.3.9. Overall Reaction

By *combining the steps described in Sections 11.3.1 through 11.3.8,* we obtain the overall reaction of the citric acid cycle:

$$\text{Acetyl CoA}^{4-} + 3NAD^+ + FAD^{2-} + 2H_2O + GDP^{3-} + P_i^{2-} \downarrow \text{CoA}-SH^{4-} + 3NADH + 2H^+ + FADH_2^{2-} + GTP^{4-} + 2CO_2$$

When we *include the conversion of pyruvate to acetyl CoA,* the overall reaction becomes

$$\text{Pyruvate}^{-} + 4NAD^+ + FAD^{2-} + 2H_2O + GDP^{3-} + P_i^{2-} \downarrow 4NADH + 2H^+ + FADH_2^{2-} + GTP^{4-} + 3CO_2$$

11.4. MAJOR FEATURES OF THE CYCLE

Figure 11.10 illustrates key aspects of the citric acid cycle. The cycle is "fed" by entrance of acetyl CoA that condenses with oxaloacetate to form citrate. After releasing the coenzyme A component in the first step, the equivalent of the acetyl group is oxidized to two molecules of CO_2 in subsequent steps. Insofar as carbon atoms are concerned, complete balance exists between the input into the cycle and the output from it. However, the carbons released as CO_2 are not actually those of the incoming acetyl group. Instead, they originate from the oxaloacetate of the previous round. You can demonstrate

this experimentally by using suitably labeled acetyl CoA (see Figure 11.10).

The citric acid cycle contains four oxidation steps, each involving removal of two hydrogen atoms from a substrate in the form of a hydride ion and a proton ($H^- + H^+$). Thus, a total of 8H are removed as $4H^-$ and $4H^+$. Three of these steps require pyridine-linked dehydrogenases that use NAD^+ as coenzyme ($3NAD^+ \rightarrow 3NADH + 3H^+$). One step requires *succinate dehydrogenase,* a flavoprotein that uses FAD as coenzyme (FAD $\rightarrow$ $FADH_2$). Succinate dehydrogenase is embedded in the inner mitochondrial membrane. All other cycle enzymes occur in the mitochondrial matrix.

Operation of the cycle results in reduction of the coenzymes NAD^+ and FAD to NADH and $FADH_2$. The reduced coenzymes must be oxidized to regenerate NAD^+ and FAD. Cells cannot tolerate an accumulation of NADH and $FADH_2$, since then the citric acid cycle, glycolysis, and other metabolic systems requiring NAD^+ or FAD would cease. Oxidation of NADH and $FADH_2$ by means of the electron transport system, located in the inner mitochondrial membrane, results in ATP synthesis via oxidative phosphorylation. Because NADH is loosely bound to its dehydrogenase, it can diffuse from the matrix to the inner membrane to be oxidized. $FADH_2$, however, is tightly bound to its dehydrogenase and cannot be oxidized in this fashion. This explains why succinate dehydrogenase is the only cycle enzyme located directly in the inner membrane.

The last reaction of the electron transport system, Eq. (12.5), requires molecular oxygen. Because of this reaction, and because of the citric acid cycle's obligatory link to the electron transport system, the cycle constitutes an aerobic pathway. Even though none of the steps of the cycle proper requires oxygen directly, the cycle can operate only under aerobic conditions.

The citric acid cycle has intimate links to respiration, both to the exhalation of carbon dioxide and to the inhalation of oxygen. The CO_2 produced by the two decarboxylation reactions of the cycle is largely exhaled and constitutes the bulk of respiratory CO_2. NADH and $FADH_2$, produced by the cycle, are oxidized by means of molecular oxygen in the electron transport system. The required oxygen constitutes the bulk of respiratory oxygen. Thus, the citric acid cycle produces most of the respiratory carbon dioxide and uses, indirectly, most of the respiratory oxygen.

When NADH and $FADH_2$ become oxidized via the electron transport system, the final reaction involves the reduction of oxygen to water. Hence, when we write the overall reaction for the citric acid cycle linked to the electron transport system, we must modify the equations in Section 11.3.9 to include water as another product of the overall reaction. The combined two metabolic systems yield two products derived from the acetyl group of acetyl CoA: *CO_2 and H_2O.* Carbon dioxide forms in the citric acid cycle, and water forms in the electron transport system. In other words, the combined action of the citric acid cycle and the electron transport system results in *complete oxidation of the acetyl group to carbon dioxide and water.*

All of the intermediates of the cycle constitute weak acids. The first few are *tricarboxylic acids,* and the remaining ones are *dicarboxylic acids.* Conversion of succinyl CoA to succinate results in formation of GTP. Because GTP can phosphorylate ADP (yielding ATP), formation of GTP represents an effective production of a molecule of ATP.

You may wonder why the oxidation of a relatively simple acetyl group requires such a complex cycle in biological systems. The reason is that straight oxidation of acetic acid requires drastic conditions, incompatible with survival of a living cell. To circumvent the problem, a longer but gentler way evolved in the course of evolution that allows cells to carry out the necessary oxidation within the intracellular environment.

11.5. ENERGETICS AND CONTROL

11.5.1. Coupled Reactions

Biochemical standard free energy changes for individual steps of the citric acid cycle are given in Table 11.1. The overall reaction has a highly negative free energy change; it is strongly exergonic and proceeds spontaneously. If we include the pyruvate dehydrogenase reaction ($\Delta G^{\circ\prime} = -33.5$ kJ mol^{-1}) in the overall reaction, the free energy change becomes even more favorable. Beginning with pyruvate, the entire sequence has a free energy change of $\Delta G^{\circ\prime} = -92.5$ kJ mol^{-1}. Starting with pyruvate, the sequence includes four strongly exergonic reactions, four reactions for which $\Delta G^{\circ\prime}$ is close to zero, and one reaction (malate dehydrogenase) that has a pronounced positive free energy change. The latter reaction proceeds because of its coupling to the exergonic citrate synthase reaction. Indeed, all of the cycle's reactions constitute sets of *coupled reactions.* The product of one reaction serves as a reactant for the next; each product functions as a *common intermediate* of two reactions. In the citric acid cycle, any two coupled reactions are catalyzed by *different* enzymes, except those involving aconitase and isocitrate dehydrogenase.

Occurrence of strongly exergonic reactions means that these steps are irreversible, at least under biochemi-

cal standard conditions, ensuring progression of the sequence in the order indicated.

11.5.2. Efficiency

Most of the energy from acetyl CoA catabolism is produced by reducing NAD^+ and FAD in the citric acid cycle and subsequently oxidizing the reduced coenzymes via the electron transport system. As you will see (Section 12.3), oxidation of NADH yields 3 ATP, and oxidation of $FADH_2$ yields 2 ATP (Section 12.4). Additionally, one ATP forms via GTP produced in the succinate thiokinase reaction. Accordingly, the total number of molecules (moles) of ATP formed per molecule (mole) of acetyl coenzyme A processed via the citric acid cycle in conjunction with the electron transport system is:

$$\text{3 (NADH, H}^+\text{) lead to: } \frac{\text{3 (NADH, H}^+\text{)}}{\text{Acetyl CoA}} \times \frac{\text{3 ATP}}{\text{NADH, H}^+} = \text{9 ATP formed}$$

$$\text{1 FADH}_2\text{ leads to: } \frac{\text{1 FADH}_2}{\text{Acetyl CoA}} \times \frac{\text{2 ATP}}{\text{FADH}_2} = \text{2 ATP formed}$$

$$\text{1 GTP leads to: } \frac{\text{1 GTP}}{\text{Acetyl CoA}} \times \frac{\text{1 ATP}}{\text{GTP}} = \text{1 ATP formed}$$

$$\text{Total} \quad \text{12 ATP formed}$$

Oxidation of acetyl CoA has a $\Delta G^{\circ\prime}$ value of -870 kJ mol^{-1} (see Problem 11.5). Hydrolysis of ATP has a $\Delta G^{\circ\prime}$ value of -30.5 kJ mol^{-1}. We can, therefore, calculate the efficiency of energy conservation for the *combined* operation of the citric acid cycle and the electron transport system as follows:

$$\frac{(30.5 \text{ kJ mol}^{-1} \times 12)\ 100}{870 \text{ kJ mol}^{-1}} = 42.1\%$$

This represents an impressive efficiency in view of the complexity of the combined systems. Keep in mind, though, that this constitutes only a rough estimate since we have based our calculations on $\Delta G^{\circ\prime}$ rather than on $\Delta G'$ values.

11.5.3. Regulatory Ratios

As you saw, control of the citric acid cycle proper occurs at three steps involving the enzymes citrate synthase, isocitrate dehydrogenase, and α-ketoglutarate dehydrogenase complex (Figure 11.13). An important additional control point is located outside the cycle, at the conversion of pyruvate to acetyl CoA (pyruvate dehydrogenase complex). At all of these points, major control becomes exerted through the ATP/ADP and $NADH/NAD^+$ ratios.

An actively metabolizing cell uses up energy rapidly. Such a cell has a low ATP concentration and a small ATP/ADP ratio. The cell also has a low NADH concentration; NADH must be rapidly oxidized via the electron transport system in order to generate the required energy in the form of ATP. Thus, the cell has a small $NADH/NAD^+$ ratio.

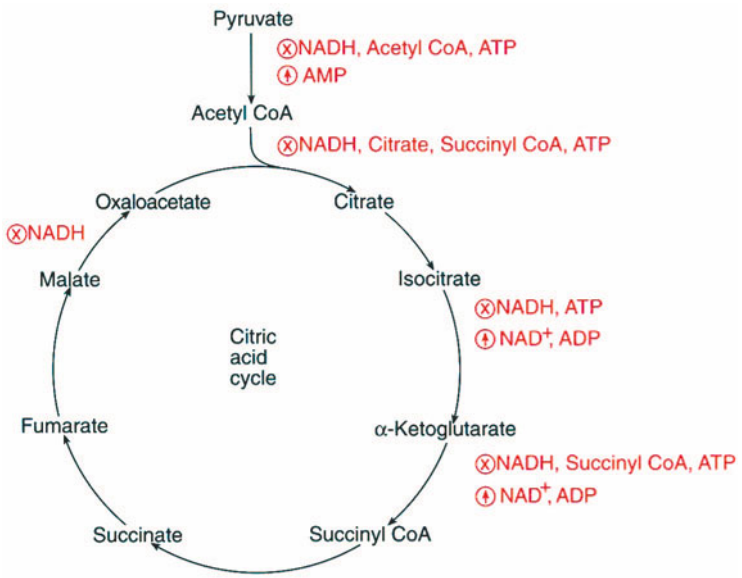


Figure 11.13. Controls of the citric acid cycle. ⊗ Inhibitor, ⊕ activator.

In a resting cell, the situation is reversed. The cell needs and uses relatively small amounts of energy. Such a cell has fairly high concentrations of both ATP and NADH so that the ATP/ADP and NADH/NAD^+ ratios have large values.

Both ATP and NADH inhibit the citric acid cycle. High concentrations block the operation of the cycle and, at the same time, slow down the electron transport system by limiting the supply of protons and electrons transferred to it from the cycle. ADP and NAD^+, on the other hand, serve as activators of the citric acid cycle and, at the same time, accelerate the electron transport system.

11.5.4. Anaplerotic Reactions

When the citric acid cycle functions in *catabolism,* acetyl CoA enters the cycle, and its acetyl moiety condenses with oxaloacetate to form citrate. Subsequent steps of the cycle result in regeneration of oxaloacetate. This means that the concentration of oxaloacetate, and that of all other cycle intermediates, *remains constant.* Processing of acetyl CoA by intermediates of the citric acid cycle is analogous to processing of metal by machines of a moving belt system for producing car bodies. The metal introduced is altered and shaped to form a car body, but the basic machinery of the moving belt system remains unchanged.

A different situation exists when the citric acid cycle functions in *anabolism.* Under these conditions, cycle intermediates are siphoned off from the pool and used as reactants for synthesis of biomolecules. Progress of such anabolic reactions lowers the concentrations of cycle intermediates. It becomes necessary to replenish the intermediates and restore their concentrations to normal levels. Special reactions called **anaplerotic reactions** (from the Greek, meaning "to fill up") accomplish this task. Typically, anaplerotic reactions involve a *carboxylation* step in which CO_2 is taken up by a compound. The conversion of pyruvate (CH_3-CO-COO^-) to oxaloacetate (^-OOC-CH_2-CO-COO^-), catalyzed by *pyruvate carboxylase,* illustrates anaplerotic reactions:

$$\text{Pyruvate}^- + CO_2 + ATP^{4-} + H_2O \rightarrow \text{oxaloacetate}^{2-} + ADP^{3-} + P_i^{2-} + 2H^+$$

Pyruvate carboxylase requires **biotin,** a member of the vitamin B complex. The coenzyme form of biotin, much like that of lipoic acid, consists of the vitamin attached via an amide bond to the ϵ-amino group of lysine in an enzyme (Figure 11.14). The complex is called **biocytin** (**biotinyllysine**) and serves as a carrier of CO_2.

Pyruvate carboxylase is an allosteric enzyme for which acetyl CoA serves as positive allosteric effector. Consider what this means for operation of the citric acid cycle. When the cycle functions in anabolism, concentra-

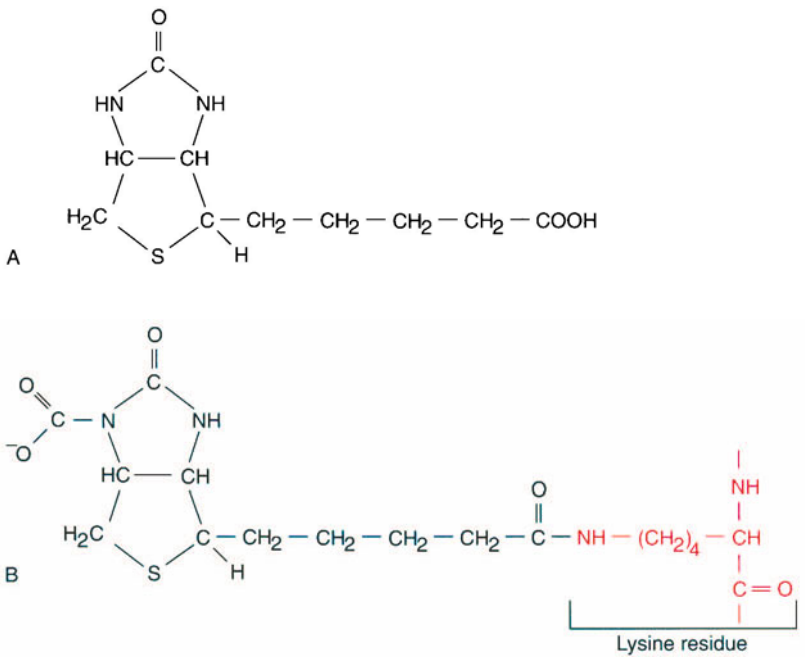


Figure 11.14. (A) Biotin. (B) Carboxybiocytin (carboxybiotinyllysine). In biocytin, biotin is linked covalently to the ϵ-NH_2 group of a lysine residue in an enzyme. Biocytin serves as a carrier of CO_2.

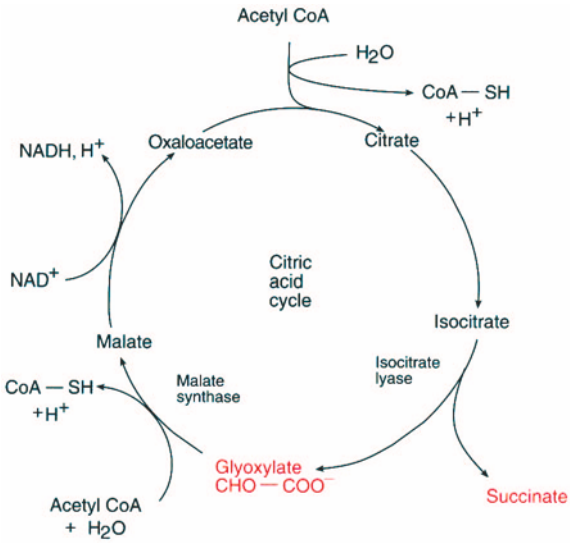


Figure 11.15. The glyoxylate cycle.

tions of cycle intermediates decrease. Acetyl CoA, derived from catabolism, cannot be processed as rapidly as it forms and begins to accumulate. As the concentration of acetyl CoA builds up, it begins to exert its allosteric effect on pyruvate carboxylase. Activation of this enzyme results in conversion of pyruvate to oxaloacetate, thereby raising the level of all cycle intermediates and allowing the cycle to metabolize the accumulated acetyl CoA.

11.6. GLYOXYLATE CYCLE

In animals, acetyl CoA cannot be used for direct synthesis of carbohydrates. Such synthesis requires conversion of acetyl CoA to pyruvate, followed by conversion of pyruvate to glucose via *gluconeogenesis*. In animals, this pathway is blocked because the reaction catalyzed by the pyruvate dehydrogenase complex has a highly negative free energy change and cannot be reversed.

Some plants and bacteria, however, can use acetyl CoA for direct synthesis of carbohydrates. In these organisms, carbohydrate synthesis proceeds via a modified citric acid cycle called the **glyoxylate cycle** (Figure 11.15). In plants, the glyoxylate cycle occurs in specialized cytoplasmic organelles named **glyoxysomes.** The cycle, proposed by Krebs and Kornberg in 1957, employs some citric acid cycle reactions as well as two unique reactions. One, catalyzed by **isocitrate lyase,** results in cleavage of isocitrate to succinate and *glyoxylate*. The other, catalyzed by **malate synthase,** leads to condensation of acetyl CoA

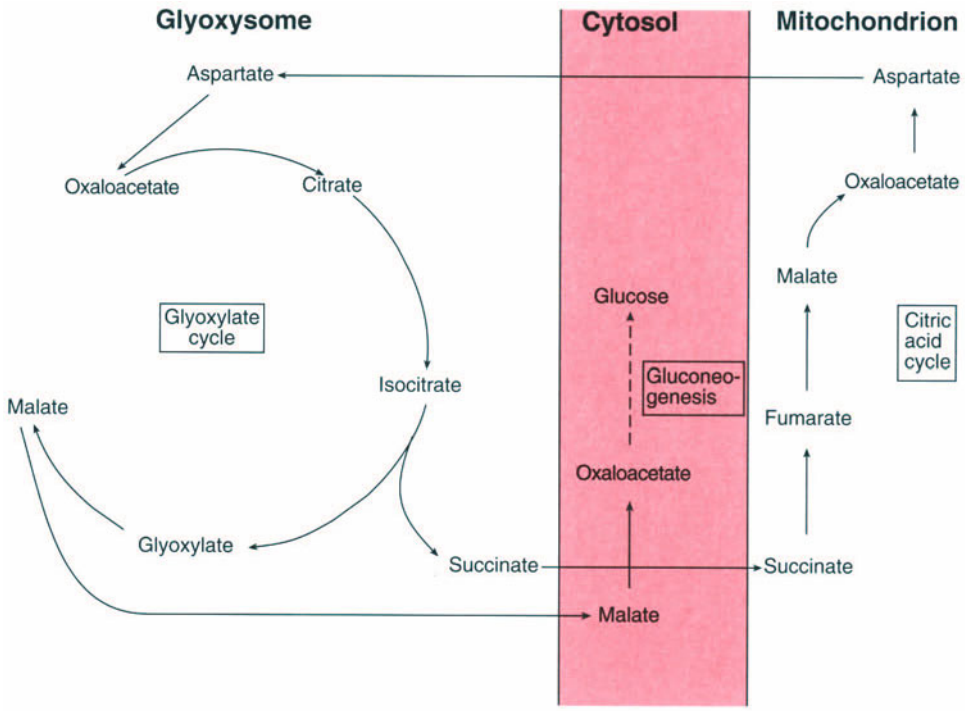


Figure 11.16. Transfer of metabolites between the glyoxysome, the mitochondrion, and the cytosol. Combined operation of the glyoxylate cycle, the citric acid cycle, and gluconeogenesis requires transfer of metabolites between these three compartments. Malate dehydrogenase catalyzes the conversion of malate to oxaloacetate in both the cytosol and the mitochondrion. Aspartate aminotransferase catalyzes the interconversion of aspartate and oxaloacetate (transamination) in both the glyoxysome and the mitochondrion.

with glyoxylate to form malate. These two steps effectively bypass the two decarboxylation reactions of the citric acid cycle.

Another significant difference between the glyoxylate and citric acid cycles is that the glyoxylate cycle requires input of *two* molecules of acetyl CoA. In fact, the cycle achieves a net conversion of two two-carbon fragments (the acetyl moiety of acetyl CoA) to one four-carbon compound (succinate). You can see this from the balanced equation:

$$2\ \text{Acetyl CoA}^{4-} + 2H_2O + NAD^+ \rightarrow \text{succinate}^{2-} + 2\ \text{CoA—SH}^{4-} + NADH + 3H^+$$

No carbon atoms of acetyl CoA are released as CO_2 during the operation of the glyoxylate cycle. The succinate formed in the glyoxysomes is transported into the mitochondria, where citric acid cycle reactions convert it to oxaloacetate. Because oxaloacetate serves as precursor for glucose synthesis by gluconeogenesis, succinate formed in the glyoxylate cycle leads to net synthesis of carbohydrates. However, gluconeogenesis takes place in the cytosol, and oxaloacetate, formed in the mitochondria, cannot pass directly across the mitochondrial membrane. To link these different pathways requires transfer of metabolites between three compartments: the glyoxysome, the mitochondrion, and the cytosol. As you can see from Figure 11.16, transfers of succinate, aspartate, and malate provide the necessary links.

Plants and bacteria, therefore, use acetate both as an energy source (citric acid cycle) and as a carbon source for synthesis of carbohydrates (glyoxylate cycle). However, the glyoxysomal pool of acetyl CoA, which drives carbohydrate synthesis via gluconeogenesis, is entirely separate from the mitochondrial pool, which drives generation of energy via the citric acid cycle.

SUMMARY

The citric acid cycle functions in both catabolism and anabolism. It constitutes the major pathway for oxidation of nutrients and for the production of metabolic energy. It also serves as an important source of intermediates for biosynthetic reactions. The cycle consists of eight consecutive, enzyme-catalyzed reactions. Any two successive reactions are energetically coupled. In addition to NAD^+ and FAD, enzymes of the cycle use three other coenzymes—coenzyme A, lipoic acid, and thiamine pyrophosphate.

The citric acid cycle is "fed" by entry of acetyl CoA. Lipid catabolism produces acetyl CoA directly. Carbohydrate catabolism yields pyruvate, which is converted to acetyl CoA by the pyruvate dehydrogenase complex, a multienzyme system.

One turn of the cycle leads to loss of two molecules of CO_2 and eight hydrogen atoms. Six hydrogens are removed in the form of 3 ($NADH, H^+$) and two in the form of $FADH_2$. These reduced coenzymes enter the electron transport system, where they are oxidized and lead to synthesis of ATP via oxidative phosphorylation. In addition, the cycle yields a molecule of GTP (ATP) by substrate-level phosphorylation. The total yield of ATP from operation of the citric acid cycle and the electron transport system is 12 molecules of ATP per molecule of acetyl CoA entering the cycle.

Major controls of the cycle occur at three steps and involve primarily the ratios ATP/ADP and $NADH/NAD^+$. ATP and NADH inhibit the cycle; ADP and NAD^+ activate it. When anabolic reactions lower the levels of cycle intermediates, anaplerotic reactions come into play that restore these concentrations to their normal levels. Plants and bacteria use a modified form of the citric acid cycle, called the glyoxylate cycle, that allows them to use acetyl CoA directly for synthesis of carbohydrates.

SELECTED READINGS

Barron, J. T., Kopp, S. J., Tow, J., and Parrillo J. E., Fatty acid, tricarboxylic acid cycle metabolites, and energy metabolism in vascular smooth muscle, *Am. J. Physiol.* 267:H764–H769 (1994).

Beevers, H., The role of the glyoxylate cycle, in *The Biochemistry of Plants,* Vol. 4 (P. K. Stumpf and E. E. Conn, eds.), Academic Press, New York (1980).

Beylot, M., Soloviev, M. V., David, F., Landau, B. R., and Brunengraber, H., Tracing hepatic gluconeogenesis relative to citric acid cycle activity *in vitro* and *in vivo, J. Biol. Chem.* 270:1509–1514 (1995).

DiDonato, L., Des Rosiers, C., Montgomery, J. A., David, F., Garneau, M., and Brunengraber, H., Rates of gluconeogenesis and citric acid cycle in perfused livers, assessed from the mass spectrometric assay of the carbon-13-labeling pattern of glutamate, *J. Biol. Chem.* 268:4170–4180 (1993).

Drozdov, L. N., *et al.,* The optimal structure of the tricarboxylic acid cycle multienzyme system of *E. coli* for the cases of cell growth on various carbon sources, *Biokhimiya (Moscow)* 59:368–380 (1994).

Durschlag, H., *et al.,* Structural changes of citrate synthase upon ligand binding and upon denaturation, *Prog. Colloid Polym. Sci.* 93:222–230 (1993).

Hiromasa, Y., Aso, Y., Yamashita, S., and Aikawa, Y., Homogeneity of the pyruvate dehydrogenase multienzyme complex from *Bacillus stearothermophilus, J. Biochem (Tokyo)* 117:467–470 (1995).

Kleczkowski, L. A, Kinetics and regulation of the NAD(P)H-dependent glyoxylate-specific reductase from spinach leaves, *Z. Naturforsch. C* 50:21–28 (1995).

Kornberg, H. L., Tricarboxylic acid cycles, *BioEssays* 7:236–238 (1987).

Krebs, H. A., The history of the tricarboxylic acid cycle, *Perspect. Biol. Med.* 14:154–170 (1970).

Mason, G. F., *et al.,* Simultaneous determination of the rates of the TCA cycle, glucose utilization, α-ketoglutarate/glutamate exchanges, and glutamine synthesis in human brain by NMR, *J. Cereb. Blood Flow Metab.* 15:12–25 (1995).

Patel, M. S., and Roche, T. E., Molecular biology and biochemistry of pyruvate dehydrogenase complexes, *FASEB J.* 4:3224–3233 (1990).

REVIEW QUESTIONS

A. Define each of the following terms:

Amphibolic pathway
Anaplerotic reaction
Biocytin
Lipoamide
Oxidative decarboxylation
Pyruvate dehydrogenase complex
Succinyl CoA
Suicide substrate

B. Differentiate between the two terms in each of the following pairs:

Succinate thiokinase/succinate dehydrogenase
Isocitrate lyase/isocitrate dehydrogenase
Succinate/oxalosuccinate
Citrate/isocitrate
Citric acid cycle/glyoxylate cycle
Malate synthase/malate dehydrogenase
Coenzyme A/acetyl coenzyme A
Lipoic acid/lipoyllysine

C. (1) Write out the eight individual reactions that, together, constitute the citric acid cycle. Include the names of the enzymes, abbreviations for the coenzymes, and names and structures of the reactants and products.
(2) What are the major functions of the citric acid cycle?
(3) What is the function of the glyoxylate cycle and how is it achieved?

PROBLEMS

11.1. Compare the ratios of NADH/NAD^+ and ATP/ADP in heart muscle during periods of sleep and jogging.

11.2. (a) Prepare a plot of 1/v versus 1/[S] for succinate dehydrogenase (v = velocity; [S] = substrate concentration). Draw 2 graphs for the following conditions on the same set of axes: (1) no inhibition; (2) in the presence of malonate. (b) Would it be possible to decrease the inhibition of succinate dehydrogenase by malonate in a liver homogenate by adding oxaloacetate? If so, explain the reason for this effect.

11.3. Write the overall equations for the following partial

sequences in both the citric acid cycle and the glyoxylate cycle: (a) citrate → malate; (b) acetyl CoA → citrate; (c) isocitrate → succinate.

11.4.* Use the data in Table 11.1 to calculate the relative concentrations of citrate and isocitrate at pH 7.0.

11.5.* Calculate $\Delta G^{\circ\prime}$ for the oxidation of acetyl CoA by coupling the following two reactions:

$$CH_3COOH + CoA\text{-}SH \rightleftarrows CH_3CO\text{-}S\text{-}CoA + H_2O \qquad \Delta G^{\circ\prime} = +31.4 \text{ kJ mol}^{-1}$$

$$CH_3COOH + 2O_2 \rightleftarrows 2CO_2 + 2H_2O \qquad \Delta G^{\circ\prime} = -837.0 \text{ kJ mol}^{-1}$$

11.6. (a) Why does it make good metabolic sense to have ATP and NADH function as inhibitors of the citric acid cycle? (b) Most of this inhibition is concentrated in the first half of the cycle (Figure 11.13). Is that advantageous? Why?

11.7. Adding the dicarboxylic acids succinate, fumarate, and oxaloacetate has been shown to increase the rate of CO_2 production by liver homogenates. What is the reason for this effect?

11.8. Can you achieve *net synthesis* of oxaloacetate by adding acetyl CoA to a cell-free extract that contains only the enzymes, coenzymes, and intermediates of the citric acid cycle?

11.9. According to Figure 11.13, AMP serves as an activator of the pyruvate dehydrogenase complex. Why is this metabolically desirable?

11.10.* A researcher labels acetyl CoA completely with ^{14}C in each of the two carbons of its acetyl group and uses the acetyl CoA to study the citric acid cycle. What fractional intensity of released $^{14}CO_2$ can the researcher expect after one, two, and three complete turns of the citric acid cycle? Assume that any acetyl CoA entering the cycle in the second and third turns does *not* carry any radioactive label.

11.11.* Suppose that a mutation in *E. coli* resulted in an altered aconitase that *binds citrate symmetrically.* What fraction of the labeled carbon atoms, introduced as the acetyl group of acetyl CoA, would be released as labeled CO_2 in one complete turn of the citric acid cycle? (Hint: Refer to Figure 11.10.)

11.12. What type of inhibition do you expect to be involved when fluorocitrate acts on aconitase? Given that fluorocitrate is a potent inhibitor, would its K_i be large or small?

11.13. How many molecules of ATP can be synthesized by the electron transport system *per molecule of acetyl CoA* (a) in one complete turn of the citric acid cycle and (b) via the glyoxylate cycle? For (b), consider only the mitochondrial reactions shown in Figure 11.16.

11.14.* What is the minimum [citrate]/[isocitrate] ratio required for the aconitase reaction to proceed spontaneously at pH 7.0? (Refer to Table 11.1.)

11.15. Succinyl CoA and citrate both inhibit the enzymes involved in their own synthesis (Figure 11.13). Name this type of inhibition.

11.16. Using the data of Table 11.1 and a value of $\Delta G^{\circ\prime} = -30.5$ kJ mol^{-1} for the hydrolysis of GTP to GDP and P_i, calculate $\Delta G^{\circ\prime}$ for the hydrolysis of succinyl CoA to succinate and CoA-SH.

11.17. Based on your calculation from the previous problem and the data of Table 11.1, calculate the efficiency of substrate-level phosphorylation (GTP synthesis) for the combined α-ketoglutarate dehydrogenase complex and succinate thiokinase reactions.

11.18. How many ATPs could theoretically be formed from coupling the first three reactions of the citric acid cycle, based on $\Delta G^{\circ\prime}$ values?

11.19. Assume that biosynthesis of purines has lowered the α-ketoglutarate concentration. How will this affect the concentrations of isocitrate and fumarate?

11.20.* If labeled oxaloacetate ($^-OOC^*-CH_2-CO-COO^-$) were to be processed through the citric acid cycle, would the CO_2 released during the first turn be labeled or not? If labeled, at what step is it released?

11.21. A student was asked to list some examples of anaplerotic reactions. The student listed the glyoxylate cycle as one answer. Is that correct? Why or why not?

11.22. Malic enzyme catalyzes the following reaction:

$$Malate^{2-} + NADP^+ \rightarrow pyruvate^- + CO_2 + NADPH$$

Could the addition of malic enzyme to a liver homogenate be used to lower the concentration of malate in the citric acid cycle?

11.23.* The hydroxyethyl group linked to TPP in Figure 11.11 is known as *active acetaldehyde.* Active acetaldehyde forms by loss of the hydrogen bonded to carbon in the thiazole ring of TPP (see Figure 11.9), forming a carbanion that attacks the carbonyl carbon of pyruvate. An input of $2H^+$ and a loss of CO_2 yields the final product. Write out the structures for the sequence of events just described. What does this mechanism tell you about the role of vitamin B_1 in TPP?

11.24. None of the reactions of the citric acid cycle requires oxygen as a reactant. Why, then, does the cycle constitute an aerobic metabolic pathway?

Electron Transport and Oxidative Phosphorylation

12

Free energy changes of chemical reactions are related to a number of parameters: entropy and enthalpy changes (Eq. 9.1), equilibrium constants (Eq. 9.2), and oxidation–reduction potentials. Of these, equilibrium constants and oxidation–reduction potentials are of primary importance for computing free energy changes of metabolic reactions. Relationships between free energy changes and equilibrium constants were explored in Chapter 9. Now we will focus on the relationships between free energy changes and oxidation–reduction potentials.

As you know, *oxidation–reduction* (or redox) reactions form the core of stage III of catabolism or cellular respiration (Figure 8.2). Oxidation of metabolites in the citric acid cycle results in removal of hydride ions (H^- or H:) and protons that are channeled into the *electron transport system* and ultimately reduce molecular oxygen to water. Energy derived from this flow of electrons leads to ATP synthesis via *oxidative phosphorylation.*

Electron transport and oxidative phosphorylation constitute the culminating events of cellular respiration. The degradative pathways of carbohydrates, lipids, and proteins converge into this final stage that provides the major metabolic synthesis of ATP in aerobic organisms. Cell respiration has been the subject of intensive research since the days of Lavoisier (1777).

In eukaryotes, electron transport and oxidative phosphorylation occur in the mitochondria, the "power houses" of the cell. Mitochondria were first described in the period 1850–1890. Techniques for isolating these organelles were perfected in the late 1940s, and between 1948 and 1950 Eugene Kennedy and Albert Lehninger discovered that the citric acid cycle, fatty acid oxidation, and oxidative phosphorylation take place in the mitochondria.

12.1. OXIDATION–REDUCTION POTENTIALS

12.1.1. Standard Redox Potentials

The nature of an overall oxidation–reduction reaction depends on the electron affinities of the oxidants in the two component half-reactions. *The half-reaction whose oxidant has the greater affinity for electrons will proceed as a reduction; the other half-reaction will proceed as an oxidation.* We describe the electron affinity of an oxidant by an **oxidation–reduction (redox) potential** for the half-reaction involved. In biochemistry, we use *reduction poten-*

tials, redox potentials based on the tendency for *gaining* electrons. Reduction potentials are measured by reference to the hydrogen half-reaction, or *hydrogen electrode:*

$$H^+ + e^- \rightleftarrows \tfrac{1}{2}H_2$$

A hydrogen electrode consists of platinum metal in contact with hydrogen gas and a solution containing hydrogen ions. We assign this electrode a potential of 0.0000 V (volts) when $[H^+] = 1.0M$, the pressure of hydrogen gas is 1.0 atm, and the temperature is 25°C. This constitutes the **standard hydrogen electrode.**

We define standard conditions for other half-reactions as we did for free energy changes (see Section 9.1): Reactant and product concentrations are $1.0M$ each, the temperature is 25°C, and the pressure is 1.0 atm. Under these conditions, any half-reaction that has a greater tendency to act as a reducing agent (that is, to generate electrons) than the standard hydrogen electrode is credited with a *negative* standard reduction potential. Half-reactions that have smaller tendencies to act as reducing agents are credited with *positive* standard reduction potentials. We designate a *standard reduction potential* as $E°$.

Because standard conditions include $[H^+] = 1.0M$ (pH 0), they are as inappropriate for evaluating biochemical potential changes as they are for evaluating biochemical free energy changes (see Section 9.1 again). We, therefore, use a different reference potential, called the **biochemical standard reduction potential,** and designate it $E°'$.

> $E°'$ is the reduction potential of a half-reaction that develops when all reactants and products are at an initial concentration of $1.0M$ each, except for protons, the initial concentration of which, unless otherwise specified, is taken as $[H^+] = 10^{-7}M$ (pH 7.0), the temperature is 25°C, and the pressure is 1.0 atm.

As we did for for free energy changes, we can define redox potentials in two ways, conceptually and mathematically. The above statement constitutes the *conceptual definition* of $E°'$ and describes its physical meaning. The *mathematical definition* of $E°'$ is given by the following equation, which permits a calculation of the value of $E°'$:

$$E°' = \frac{2.303RT}{nF} \log K'_{bio} = \frac{2.303RT}{nF} \log \frac{[\text{reductant}]}{[\text{oxidant}]} \qquad (12.1)$$

where R is the gas constant ($8.314\ J\ deg^{-1}\ mol^{-1} = 1.987$ cal $deg^{-1}\ mol^{-1}$), T is the absolute temperature (K), n is the number of electrons in the half-reaction, F is the Faraday constant (96,491 J V^{-1} mol^{-1} = 23,062 cal V^{-1} mol^{-1}), and K'_{bio} is the equilibrium constant at pH 7.0.

Generally, K'_{bio} consists of only two terms, [reductant] and [oxidant]; electrons are omitted from the expression, and protons do not appear in it because we calculate K'_{bio} for a fixed pH of 7.0. The concentrations of all components are those existing *at equilibrium.* At 25°C (T = 298.2 K), the coefficient of the logarithmic term equals $0.06/n$.

One can determine biochemical standard reduction potentials by measuring the potential, or *electromotive force,* developed in an electrochemical cell composed of two half-reactions, a standard hydrogen electrode and the half-reaction of interest (see Appendix D, Figure D.2). Table 12.1 lists $E°'$ values for some half-reactions of biological interest.

12.1.2. Actual Redox Potentials

Reduction potentials developed under nonstandard conditions constitute *actual reduction potentials.* The potential corresponding to $E°$ is designated E, and that corresponding to $E°'$ is designated E'. E' is the **biochemical actual**

Table 12.1. Biochemical Standard Reduction Potentials ($E°'$) of Some Half-Reactions of Biological Relevance

Half-reaction[a]	$E°'$ (volts)
Ferredoxin (Fe^{3+}) + e^- ⇌ ferredoxin (Fe^{2+})	−0.43
$H^+ + e^- \rightleftarrows \frac{1}{2}H_2$	−0.42
α-Ketoglutarate + CO_2 + $2H^+$ + $2e^-$ ⇌ isocitrate	−0.38
$NAD^+ + 2H^+ + 2e^- \rightleftarrows NADH + H^+$	−0.32
$NADP^+ + 2H^+ + 2e^- \rightleftarrows NADPH + H^+$	−0.32
1,3-Bisphosphoglycerate + $2H^+$ + $2e^-$ ⇌ G-3-P + P_i	−0.29
Acetaldehyde + $2H^+$ + $2e^-$ ⇌ ethanol	−0.20
Pyruvate + $2H^+$ + $2e^-$ ⇌ lactate	−0.19
FAD + $2H^+$ + $2e^-$ ⇌ $FADH_2$ (free coenzyme)	−0.18
FMN + $2H^+$ + $2e^-$ ⇌ $FMNH_2$ (free coenzyme)	−0.18
Oxaloacetate + $2H^+$ + $2e^-$ ⇌ malate	−0.17
Fumarate + $2H^+$ + $2e^-$ ⇌ succinate	+0.03
Myoglobin (Fe^{3+}) + e^- ⇌ Myoglobin (Fe^{2+})	+0.05
DHA + $2H^+$ + $2e^-$ ⇌ ascorbic acid	+0.06
Cyt *b* (Fe^{3+}) + e^- ⇌ Cyt *b* (Fe^{2+})	+0.07
CoQ + $2H^+$ + $2e^-$ ⇌ $CoQH_2$	+0.05
Hemoglobin (Fe^{3+}) + e^- ⇌ hemoglobin (Fe^{2+})	+0.17
Cyt c_1 (Fe^{3+}) + e^- ⇌ Cyt c_1 (Fe^{2+})	+0.22
Cyt *c* (Fe^{3+}) + e^- ⇌ Cyt *c* (Fe^{2+})	+0.25
Cyt *a* (Fe^{3+}) + e^- ⇌ Cyt *a* (Fe^{2+})	+0.29
$O_2 + 2H^+ + 2e^- \rightleftarrows H_2O_2$	+0.30
$NO_3^- + 2H^+ + 2e^- \rightleftarrows NO_2^- + H_2O$	+0.42
Cyt a_3 (Fe^{3+}) + e^- ⇌ Cyt a_3 (Fe^{2+})	+0.39
$Fe^{3+} + e^- \rightleftarrows Fe^{2+}$	+0.77
$\frac{1}{2}O_2 + 2H^+ + 2e^- \rightleftarrows H_2O$	+0.82

[a]Abbreviations: G-3-P, glyceraldehyde 3-phosphate; DHA, dehydroascorbic acid; Cyt, cytochrome.

reduction potential. E is identical to E', much as ΔG is identical to $\Delta G'$ (see Section 9.1). There exists only *one* actual reduction potential for a given set of nonstandard conditions, whether we base it on E° and designate it E or base it on $E^{\circ\prime}$ and designate it E'. To be consistent with free energy designations, we will use $E^{\circ\prime}$ and E' for reduction potentials.

The relationship of E' to $E^{\circ\prime}$ is described by the **Nernst equation:**

$$E' = E^{\circ\prime} - \frac{2.303RT}{nF} \log \frac{[\text{reductant}]}{[\text{oxidant}]} \qquad (12.2)$$

where [reductant] and [oxidant] represent the *actual initial* (not equilibrium) *concentrations* of the reduced and oxidized forms. Note that just as $\Delta G'$ is related to $\Delta G^{\circ\prime}$ by a concentration term, so E' is related to $E^{\circ\prime}$ by a similar term.

For redox reactions, free energy changes are related to potentials by two equations:

$$\Delta G^{\circ\prime} = -n\mathrm{F}\Delta E^{\circ\prime} \qquad (12.3)$$

$$\Delta G' = -n\mathrm{F}\Delta E' \qquad (12.4)$$

where $\Delta E^{\circ\prime}$ and $\Delta E'$ represent the overall change in reduction potential when two half-reactions are combined, and n is the number of electrons generated by one half-reaction and consumed by the other half-reaction.

These equations show that a change in reduction potential is *equivalent* to a change in free energy. It is this interdependence that accounts for the importance of redox reactions in biochemical energetics.

The overall redox reaction resulting from the combination of two half-reactions represents a special case of *coupled reactions* (Section 9.3) in which electrons constitute the *common intermediate.* Electrons are a product of one half-reaction and serve as a reactant for the second half-reaction. How to couple two half-reactions under standard or actual conditions is described in detail in Appendix D. As pointed out there, *what ultimately determines the direction of an overall redox reaction are the intracellular concentrations of reactants and products.*

12.2. BIOLOGICAL ELECTRON CARRIERS

In Chapter 11 you saw that two types of dehydrogenases—pyridine-linked dehydrogenases and flavoproteins—catalyze the oxidation of metabolites in the citric acid cycle. The redox reactions are mediated by NAD^+ and FAD, the coenzymes of these dehydrogenases. When the metabolite undergoes oxidation, the coenzyme undergoes reduction; NAD^+ is reduced to NADH, and FAD is reduced to $FADH_2$. The process requires removal of two hydrogen atoms from the metabolite in the form of a hydride ion (H^- or H:) and a proton (H^+), one or both of which are transferred to the coenzyme. Reduction of NAD^+ requires transfer of H^-, and reduction of FAD requires transfer of both H^- and H^+. From NADH and $FADH_2$, the abstracted hydride ions and protons are transferred to other redox coenzymes and special components that function as consecutive **electron carriers.** Electrons flow through this chain of carriers to oxygen. The entire array of specific enzymes, coenzymes, and other components constitutes the **electron transport system (ETS).**

In eukaryotes, the electron transport system is located in the inner mitochondrial membrane. Prokaryotes have similar electron transport systems but composed of different electron carriers. All prokaryotic components are attached to the plasma membrane so that the reactions occur at the cell periphery.

The mitochondrial ETS consists of five groups of biological electron carriers: *pyridine-linked dehydrogenases, flavoproteins, coenzyme Q (ubiquinone), cytochromes,* and *iron–sulfur proteins.* We discussed pyridine-linked dehydrogenases and flavoproteins in Section 11.1. The remaining electron carriers are described below.

12.2.1. Coenzyme Q (Ubiquinone)

Coenzyme Q or **ubiquinone,** abbreviated **CoQ** or **Q,** is a generic term for a group of electron carriers discovered in the late 1950s and characterized by *quinone*-like structures and *ubiquitous* occurrence (Figure 12.1).

Despite their names, ubiquinones are *not* coenzymes and are not linked to proteins. Rather, they comprise lipid-soluble compounds present in the lipid phase of the mitochondrial membrane, where they function as **mobile electron carriers.** Ubiquinones owe their nonpolar, *lipophilic* character to the *isoprenoid side chain.* Like the flavoproteins, ubiquinones can participate in either a one-electron or a two-electron transfer. Ubiquinone forms a free radical that has a stable semiquinone-type structure (Figure 12.2).

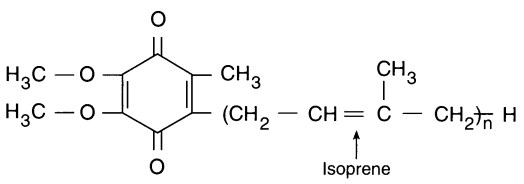


Figure 12.1. Coenzyme Q (ubiquinone). The length of the isoprenoid chain varies. Typically, n has a value of 10 for eukaryotes and 6 for bacteria.

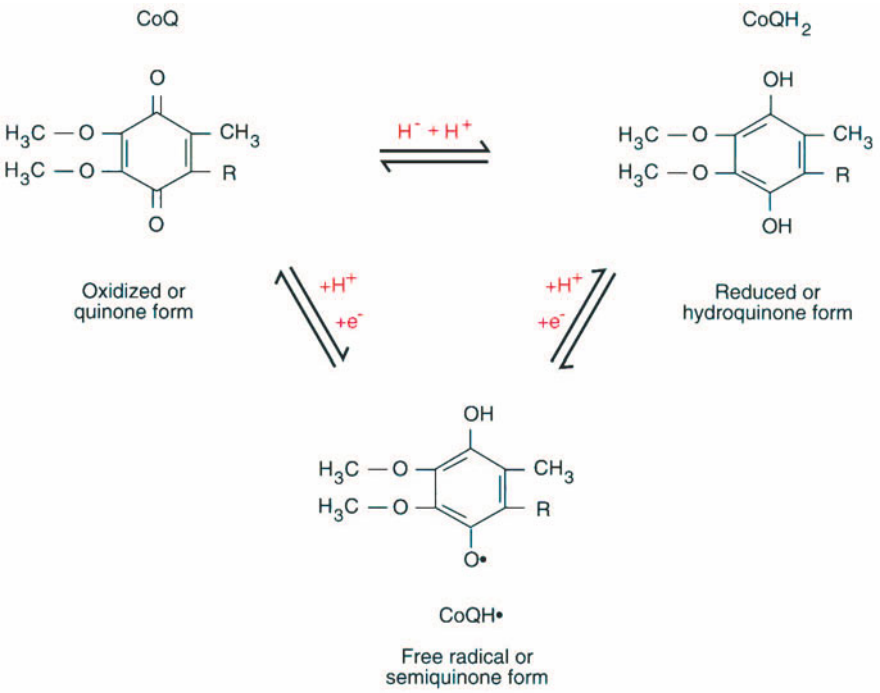


Figure 12.2. Oxidation and reduction of coenzyme Q.

12.2.2. Cytochromes

Cytochromes ("cellular pigments") were first identified in the late 19th century. We now know that cytochromes occur in all types of cells. They are located primarily in membranes and, in eukaryotes, specifically in the mitochondrial membrane.

Cytochromes are red-brown, conjugated proteins that contain an *iron–porphyrin complex* or *heme*. Some cytochromes contain a heme identical to that present in hemoglobin and myoglobin. Investigators have identified some 30 different cytochromes. They vary in the substituents of the tetrapyrrole ring system (Figure 12.3), the amino acid sequence of the conjugated protein, and the linkage of heme to protein. Of all the cytochromes, only cytochrome *c* is readily extracted from the mitochondrial membrane; the other cytochromes are *integral membrane proteins*.

Cytochromes have characteristic spectral properties and absorb strongly in the visible range (Figure 12.4). Cytochromes accept or donate one electron by virtue of reversible oxidation–reduction of the heme iron:

$$Fe^{3+} \underset{-e^-}{\overset{+e^-}{\rightleftharpoons}} Fe^{2+}$$

Although all cytochromes share this fundamental redox reaction, they do not have the same $E^{\circ\prime}$ value (see Table 12.1) because reduction of the iron atom occurs within different environments. The polypeptide chains that surround the hemes of different cytochromes vary with respect to the type, number, and position of charged functional groups. Additionally, cytochromes may differ in the structure of the heme and in its attachment to the polypeptide chain. Hence, the local environment in which the reduction of iron takes place varies from compound to compound. Reduction proceeds more readily in some cytochromes than in others, a property reflected in their nonidentical $E^{\circ\prime}$ values.

12.2.3. Iron–Sulfur Proteins

Iron–sulfur proteins (Fe-S), or **nonheme iron proteins (NHI proteins),** are conjugated proteins that contain complexes of iron and sulfur but no heme. Biochemists recognize several types of iron–sulfur proteins (Figure 12.5), but in all cases each iron atom is coordinated to four sulfur atoms. Some sulfur atoms constitute components of cysteine's sulfhydryl groups; others are "acid-labile" atoms that can be released from the iron–sulfur protein at pH $\approx$ 1 in the form of H_2S. Iron–sulfur proteins undergo oxidation–reduction reactions by virtue of reversible oxidation of their iron. The oxidized and reduced states of all iron–sulfur proteins differ by one charge, regardless of the number of iron atoms they contain. Hence, much like cytochromes, iron–sulfur proteins function in one-electron transfer reactions.

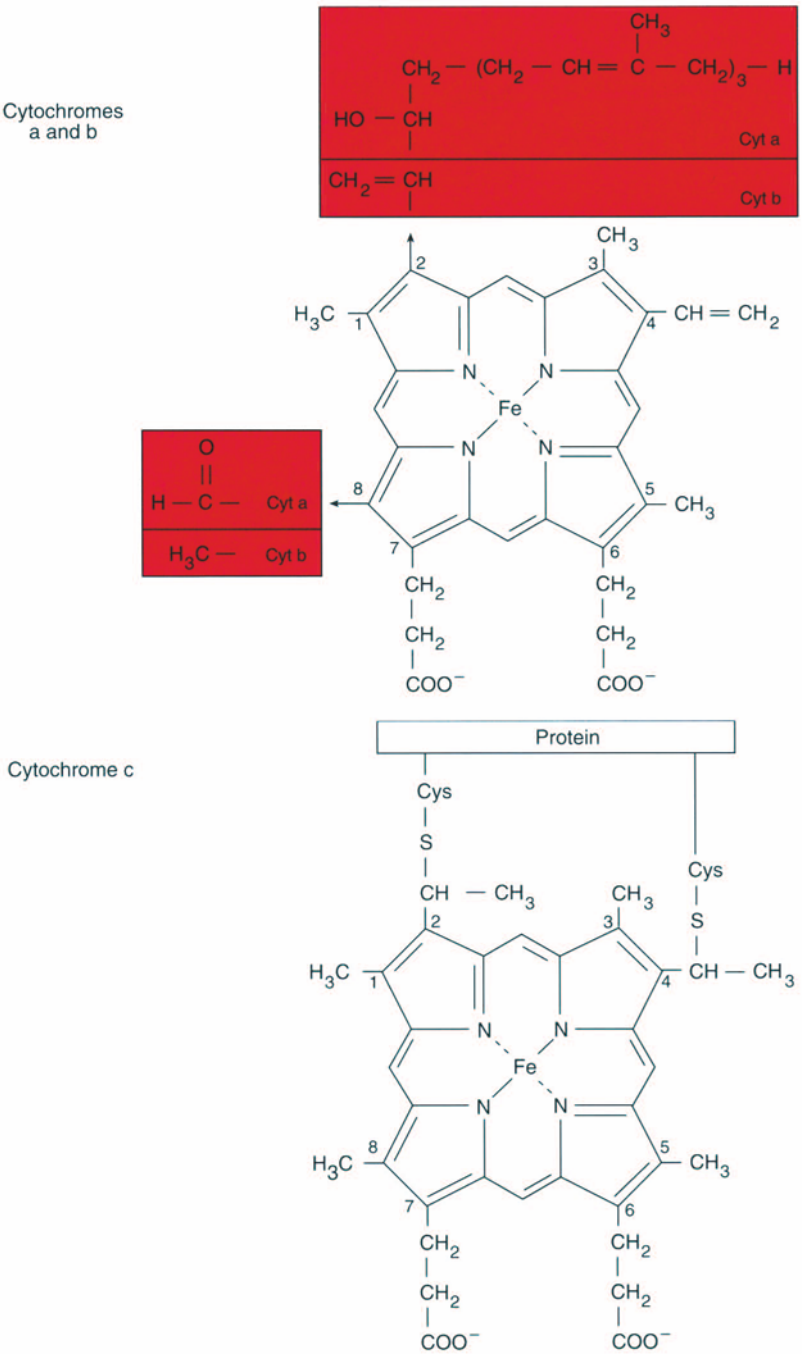


Figure 12.3. Structures of the heme groups of cytochromes *a*, *b*, and *c*. Cytochrome *b* heme is identical to that occurring in hemoglobin and myoglobin (iron-protoporhyrin IX).

12.2.4. Comparing the Electron Carriers

Oxidation–reduction characteristics of cytochromes and iron–sulfur proteins differ from those of pyridine nucleotides, flavins, and ubiquinones in two respects. First, redox reactions of cytochromes and iron–sulfur proteins involve only electrons, not electrons plus varying forms of hydrogen atoms. Second, cytochromes and iron–sulfur proteins are restricted to one-electron transfer reactions. Pyridine nucleotides, on the other hand, require a two-electron transfer (H^-), and flavins and coenzyme Q can function in either one-electron or two-electron transfers.

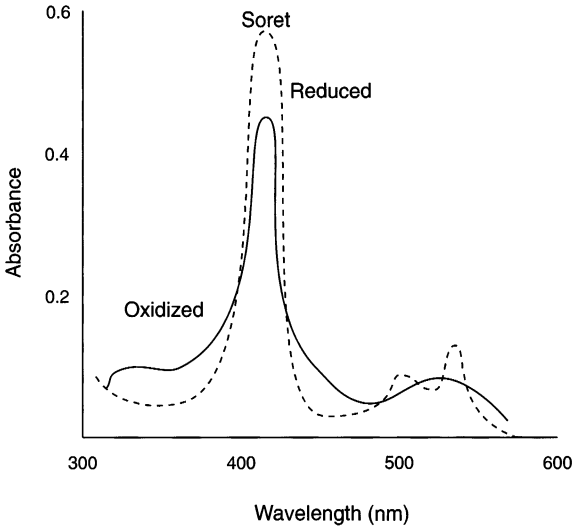


Figure 12.4. Absorption spectra of cytochrome *c* at pH 7.0. The characteristic absorption band at about 400 nm is called the *Soret band.* Other cytochromes have similar absorption spectra.

Oxidation of metabolites in the citric acid cycle generates the electrons that flow through the electron transport system. These redox reactions involve the transfer of two electrons to NAD^+ or FAD. As the two electrons pass from NADH or $FADH_2$ to a cytochrome or an iron–sulfur protein, either one electron carrier molecule must be reduced twice consecutively, accepting one electron at a time, or two electron carrier molecules must function in concert and be reduced simultaneously, with each accepting a single electron. Researchers have not yet definitively established which of these two mechanisms applies in each case.

The capacity of flavoproteins and ubiquinones to function as electron carriers in *either* one-electron or two-electron transfer reactions is of crucial importance for operation of the electron transport system. Because of this flexibility, flavoproteins and ubiquinones function as electron conduits between the two-electron donor, NADH, and the one-electron acceptors, cytochromes and iron–sulfur proteins.

12.3. ELECTRON TRANSPORT SYSTEM (ETS)

We are now ready to link together the reactions of the various electron carriers and construct the mitochondrial electron transport system. In this system, electrons flow from metabolite to molecular oxygen, which is reduced to water. Electron flow leads to pumping of protons across the inner mitochondrial membrane. Proton pumping occurs via three complexes located in the membrane and results in establishment of a proton gradient that provides the energy for ATP synthesis.

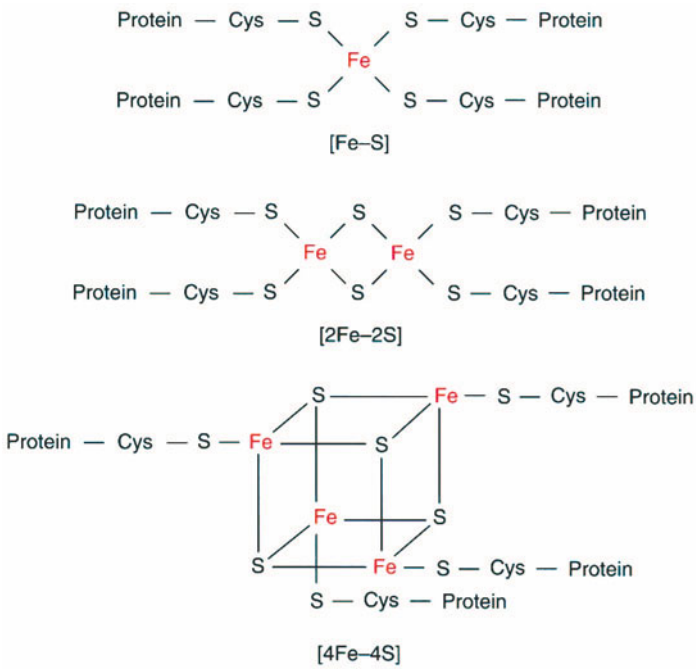


Figure 12.5. Common forms of iron-sulfur clusters in nonheme iron proteins.

12.3.1. Chain of Electron Carriers

12.3.1A. From Metabolite to Cytochrome *c*. The initial step in the sequence involves the oxidation of a substrate of a pyridine-linked dehydrogenase. Oxidation of isocitrate, the substrate of isocitrate dehydrogenase and a citric acid cycle intermediate, represents a typical example.

Let us designate the general substrate as some metabolite, MH_2, from which two hydrogens can be removed. Oxidation of MH_2 is *coupled* to the reduction of NAD^+ to NADH:

$$MH_2 + NAD^+ \rightleftharpoons M + NADH + H^+$$

As noted earlier, the cell cannot tolerate extensive depletion of NAD^+. If NAD^+ is not regenerated from NADH, operation of both glycolysis and the citric acid cycle will come to a halt. NAD^+ regeneration is accomplished by coupling the oxidation of NADH to the reduction of a flavoprotein; specifically, flavin mononucleotide (FMN) is reduced to $FMNH_2$ as NADH is oxidized to NAD^+:

$$NADH + H^+ + FMN \rightleftarrows NAD^+ + FMNH_2$$

Much as the cell cannot tolerate depletion of NAD^+, it can also not tolerate depletion of FMN. To regenerate FMN, electrons flow from $FMNH_2$ to an iron–sulfur protein; the oxidation of $FMNH_2$ is coupled to reduction of an iron–sulfur protein. Because full oxidation of $FMNH_2$ requires removal of two electrons, one iron–sulfur protein must be reduced twice consecutively or two iron–sulfur proteins must be reduced simultaneously:

$$FMNH_2 + 2Fe^{3+}\text{–}S \rightleftarrows FMN + 2Fe^{2+}\text{–}S + 2H^+$$

The reduced iron–sulfur protein is reoxidized by transferring its electrons to coenzyme Q. Full reduction of coenzyme Q requires the transfer of two electrons:

$$2Fe^{2+}\text{–}S + CoQ + 2H^+ \rightleftarrows 2Fe^{3+}\text{–}S + CoQH_2$$

Following the reduction of coenzyme Q, the remaining electron carriers—another iron–sulfur protein and several cytochromes—can participate only in one-electron transfers. The first electron carrier is a cytochrome, designated cytochrome *b,* that oxidizes $CoQH_2$ back to CoQ. In the process, the ferric iron (Fe^{3+}) of cytochrome *b* is reduced to the ferrous state (Fe^{2+}):

$$CoQH_2 + 2\ \text{Cyt}\ b\ (Fe^{3+}) \rightleftarrows CoQ + 2\ \text{Cyt}\ b\ (Fe^{2+}) + 2H^+$$

Reduced cytochrome *b* is oxidized by means of an iron–sulfur protein:

$$2\ \text{Cyt}\ b\ (Fe^{2+}) + 2Fe^{3+}\text{–}S \rightleftarrows 2\ \text{Cyt}\ b\ (Fe^{3+}) + 2Fe^{2+}\text{–}S$$

From the iron–sulfur protein, the electrons flow through a series of cytochromes. The first cytochrome reoxidizes the iron–sulfur protein and is itself reduced. The reduced cytochrome is then reoxidized by means of another cytochrome that follows it in the series and is reduced in the process. In principle, this reduction and reoxidation of cytochromes could go on until all the potential free energy of the transfer process is expended. Instead, the chain of cytochromes and the entire ETS is brought to a useful termination point by a significant change in the mechanism.

12.3.1B. Cytochrome Oxidase. The change comes in the form of **cytochrome oxidase,** the last cytochrome in the series. Like all oxidases, cytochrome oxidase catalyzes the *direct combination of a substrate with molecular oxygen.* The cytochrome oxidase step represents one of the few, but critical, biochemical redox reactions in which oxidation occurs by addition of oxygen rather than by removal of hydrogen. Cytochrome oxidase is a multienzyme complex. Mammalian cytochrome oxidase consists of 10 subunits and contains two different cytochromes, *a* and a_3. The enzyme complex also has two essential copper ions (designated Cu_A and Cu_B) that can alternate between +1 and +2 states as they participate in the redox reactions. One copper ion is associated with each heme. Heme *a* is close to Cu_A in subunit II, and heme a_3 is close to Cu_B in subunit I.

The oxidized form of cytochrome oxidase is regenerated from the reduced enzyme by reaction with molecular oxygen. In the process, molecular oxygen is reduced to water. We can formulate the final step of the ETS as follows:

$$2\ \text{Cyt oxidase}\ (Fe^{2+}) + \tfrac{1}{2}O_2 + 2H^+ \rightarrow 2\ \text{Cyt oxidase}\ (Fe^{3+}) + H_2O \qquad (12.5)$$

By being able to use molecular oxygen, an *external source of oxidizing power,* the chain of electron carriers comes to an end. As in all previous steps, the oxidizing electron carrier is reduced. However, in this step, as distinct from all previous steps, depletion of the oxidized form of the final electron carrier (O_2) and accumulation of its reduced form (H_2O) present no problem. There is no need to regenerate oxygen by oxidation of water since the supply of molecular oxygen is virtually unlimited. Likewise, accumulation of water can be tolerated since the water formed is readily removed by excretion.

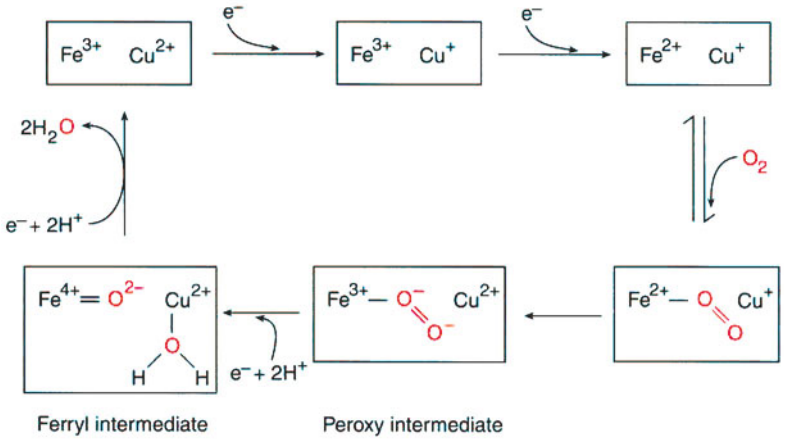


Figure 12.6. Proposed cyclic mechanism for the four-electron reduction of O_2 as catalyzed by cytochrome oxidase at the (heme a_3–Cu_B) center of the enzyme. [Adapted, with permission, from M. Wikstrom and J. E. Morgan, *J. Biol. Chem.* 267:10266–10273 (1992).]

The cytochrome oxidase reaction, Eq. (12.5), constitutes the key reaction of respiration. In addition to providing a useful end to the series of redox steps, the reaction also requires $2H^+$. These protons are available from the oxidation of $CoQH_2$ to CoQ. Inclusion of these protons provides for complete material balance of the sequence. However, note that, although the reaction is stoichiometrically correct, the actual cytochrome oxidase reaction must involve one (not one-half) molecule of oxygen and is, therefore a *four-electron* process:

$$\begin{aligned} O_2 + 4H^+ + 4e^- &\rightarrow 2H_2O \\ 4\ \text{Cyt oxidase}\ (Fe^{2+}) + O_2 + 4H^+ &\rightarrow \\ 4\ \text{Cyt oxidase}\ (Fe^{3+}) + 2H_2O & \end{aligned} \qquad (12.5a)$$

Reduction of one molecule of oxygen by four electrons must proceed without formation of intermediates in which oxygen is incompletely reduced. Such oxygen forms (see Section 12.5) are extremely toxic to living systems. Figure 12.6 shows the proposed mechanism by which cytochrome oxidase accomplishes its task. Reduced cytochrome *c* donates its electron to the heme *a*–Cu_A center of cytochrome oxidase. An electron is then transferred to the heme a_3–Cu_B center of the enzyme, where O_2 is reduced in a cyclic series of steps to two molecules of water.

The cycle begins with the heme a_3–Cu_B center in its fully oxidized state (Fe^{3+}–Cu^{2+}). The first electron relayed from reduced cytochrome *c* reduces Cu^{2+} to Cu^+. The second electron relayed reduces Fe^{3+} to Fe^{2+}, producing a fully reduced iron–copper center (Fe^{2+}–Cu^+). An oxygen molecule now binds to the iron of the center and abstracts an electron from both center ions to form a *peroxy* intermediate. The peroxide is cleaved when a third electron and two protons are taken up. At this point, one oxygen atom is bound as water to Cu^{2+} while the second oxygen atom is bound to iron in the *ferryl* state (Fe^{4+}). Input of a fourth electron and two more protons leads to release of oxygen in the form of two water molecules and to regeneration of the fully oxidized iron–copper center (Fe^{3+}–Cu^{2+}). The last two stages of the cycle involve proton pumping across the mitochondrial membrane. Four protons are pumped into the intermembrane space as two electrons flow through the oxidase.

The cytochrome oxidase reaction makes breathing an essential aspect of life for plants and animals. Inhaled oxygen is required to regenerate the oxidized form of cytochrome oxidase and thereby permit operation of the electron transport system. Because of this link to respiration, we also refer to the mitochondrial electron transport

Table 12.2. Reactions of the Electron Transport System[a]

$MH_2 + NAD^+ \rightleftharpoons M + NADH + H^+$
$NADH + H^+ + FMN \rightleftharpoons NAD^+ + FMNH_2$
$FMNH_2 + 2Fe^{3+}\text{–S} \rightleftharpoons FMN + 2Fe^{2+}\text{–S} + 2H^+$
$2Fe^{2+}\text{–S} + CoQ + 2H^+ \rightleftharpoons CoQH_2 + 2Fe^{3+}\text{–S}$
$CoQH_2 + 2\ \text{Cyt}\ b\ (Fe^{3+}) \rightleftharpoons CoQ + 2\ \text{Cyt}\ b\ (Fe^{2+}) + 2H^+$
$2\ \text{Cyt}\ b\ (Fe^{2+}) + 2Fe^{3+}\text{–S} \rightleftharpoons 2\ \text{Cyt}\ b\ (Fe^{3+}) + 2Fe^{2+}\text{–S}$
$2Fe^{2+}\text{–S} + 2\ \text{Cyt}\ c_1\ (Fe^{3+}) \rightleftharpoons 2Fe^{3+}\text{–S} + 2\ \text{Cyt}\ c_1\ (Fe^{2+})$
$2\ \text{Cyt}\ c_1\ (Fe^{2+}) + 2\ \text{Cyt}\ c\ (Fe^{3+}) \rightleftharpoons 2\ \text{Cyt}\ c_1 (Fe^{3+}) + 2\ \text{Cyt}\ c\ (Fe^{2+})$
$2\ \text{Cyt}\ c\ (Fe^{2+}) + 2\ \text{Cyt}\ a\ (Fe^{3+}) \rightleftharpoons 2\ \text{Cyt}\ c\ (Fe^{3+}) + 2\ \text{Cyt}\ a\ (Fe^{2+})$
$2\ \text{Cyt}\ a\ (Fe^{2+}) + 2Cu^{2+} \rightleftharpoons 2\ \text{Cyt}\ a\ (Fe^{3+}) + 2Cu^+$
$2Cu^+ + 2\ \text{Cyt}\ a_3\ (Fe^{3+}) \rightleftharpoons 2Cu^{2+} + 2\ \text{Cyt}\ a_3\ (Fe^{2+})$
$2\ \text{Cyt}\ a_3\ (Fe^{2+}) + \frac{1}{2}O_2 + 2H^+ \rightarrow 2\ \text{Cyt}\ a_3\ (Fe^{3+}) + H_2O$

Overall reaction: $MH_2 + \frac{1}{2}O_2 \rightarrow M + H_2O$
Actual reaction: $2MH_2 + O_2 \rightarrow 2M + 2H_2O$

[a]Fe–S, Iron–sulfur protein; Cyt, cytochrome.

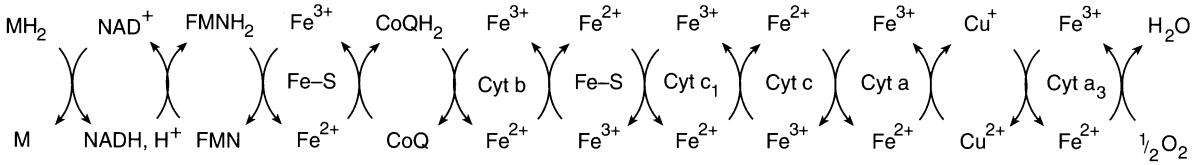


Figure 12.7. Sequence of the electron carriers in the mitochondrial electron transport system (ETS). Fe–S represents iron–sulfur proteins. See Figure 12.6 for the detailed mechanism of cytochrome oxidase.

system as the **respiratory chain.** Cyanide combines avidly with ferric iron (Fe^{3+}) of cytochrome oxidase, thus blocking the electron transport system and the associated synthesis of ATP via oxidative phosphorylation. This is why cyanide is such a strong poison. Carbon monoxide also inhibits cytochrome oxidase; it binds to the ferrous iron (Fe^{2+}) of the enzyme.

12.3.1C. Pathways of Electron Transport. Table 12.2 provides a summary of the entire sequence of reactions. The detailed mechanism corresponding to the last three reactions in the table was described above. Figure 12.7 schematically represents the coupled reactions by means of curved arrows.

Many metabolites are oxidized via the pathway shown in Figure 12.7. Succinate, however, is processed by means of a second pathway in which a flavoprotein (FAD) carries out the initial metabolite oxidation. Reduced $FADH_2$ is then oxidized, to regenerate FAD, by reaction with an iron–sulfur protein. The next electron carrier is a cytochrome *b,* different from that in Figure 12.7, which subsequently reduces coenzyme Q and thereby enters the major pathway. Figure 12.8 outlines the two pathways of electron transport. In the overall scheme, NAD^+ and CoQ serve as central electron-collecting compounds; they channel the electrons removed from metabolites through the remaining portion of the ETS. The four complexes shown in Figure 12.8 are discussed below.

12.3.2. Sequence of ETS Components

Researchers have elucidated the sequence of electron carriers in the ETS by means of three main experimental approaches. These involve (1) characterizing respiratory complexes, (2) using artificial electron acceptors, and (3) using inhibitors.

12.3.2A. Respiratory Complexes. Careful fragmentation of the inner mitochondrial membrane, site of the ETS, has yielded a number of macromolecular aggregates that have the capacity to carry out specific portions of the entire sequence. In particular, investigators have characterized four such functional assemblies, called **respiratory complexes** (Table 12.3). Each complex represents a multienzyme system that catalyzes a number of consecutive reactions. Electron flow through three of these transmembrane complexes (I, III, and IV) leads to proton pumping across the mitochondrial membrane (Figure 12.9). The resultant proton gradient drives ATP synthesis, but actual ATP synthesis occurs at the active site of the enzyme ATP synthase.

Analysis of each complex for its content of electron carriers and for the reactions catalyzed by them has helped to establish the sequence of ETS reactions. These studies assume the likelihood that electron carriers located close to each other participate in reactions that are closely linked mechanistically.

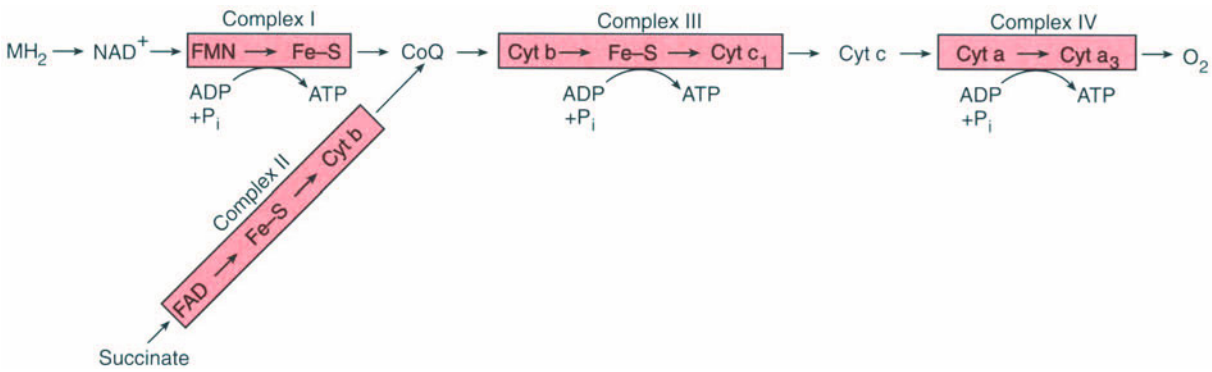


Figure 12.8. The two pathways of electron transport in mitochondria. Cytochrome *b* of Complex II is not identical to that of Complex III.

Table 12.3. Respiratory Complexes of the Electron Transport System

	Complex I	Complex II	Complex III	Complex IV
Name	NADH-CoQ reductase	Succinate-CoQ reductase[a]	Cytochrome reductase	Cytochrome oxidase
Reaction sequence	$NAD^+ \rightarrow CoQ$	Succinate → CoQ	CoQ → cytochrome c	Cytochrome c → O_2
Molecular weight	850,000	127,000	280,000	200,000
No. of subunits	26	5	10	13
Fe–S protein	Yes	Yes	Yes	—
$\Delta E^{\circ\prime}$ (volts)	+0.37	+0.02	+0.20	+0.57
ATP synthesis	Yes	—	Yes	Yes

[a]Also known as *succinate dehydrogenase complex.*

12.3.2B. Artificial Electron Acceptors. A second approach for determining the sequence of electron carriers is based on using **artificial electron acceptors.** An artificial electron acceptor is a compound that, like naturally occurring electron carriers, undergoes oxidation–reduction reactions as a function of its specific reduction potential. When such a compound is added to the ETS, it can be reduced by an electron carrier with a lower reduction potential. In the presence of an artificial electron acceptor, then, electrons passing through the ETS are *diverted* and used to reduce the added compound.

We can pinpoint the site at which this takes place. Siphoning off electrons from the ETS by the added compound results in fewer electrons being available for the remaining portion of the ETS. With fewer electrons, less reduction of electron carriers takes place in the remaining ETS segment. Hence, diversion of electrons leads to a *decrease in the concentration of reduced forms of electron carriers subsequent to the point of electron diversion* (Figure 12.10).

Put somewhat differently, the ratio of the oxidized to the reduced form increases for carriers subsequent to the point of electron diversion. Changes in these ratios can be detected spectrophotometrically, since reduced and oxidized forms of many electron carriers have different absorption properties (see Figure 12.4). By using a variety of artificial electron acceptors, we can elucidate the sequence of electron carriers much as we can elucidate a metabolic pathway by using mutants (Section 8.3). Figure 12.11 indicates the points of action of a number of artificial electron acceptors.

12.3.2C. Inhibitors. The third approach for determining the sequence of carriers in the ETS involves using inhibitors that *block* specific steps in the sequence (Figure 12.11). The effect of adding an inhibitor is identi-

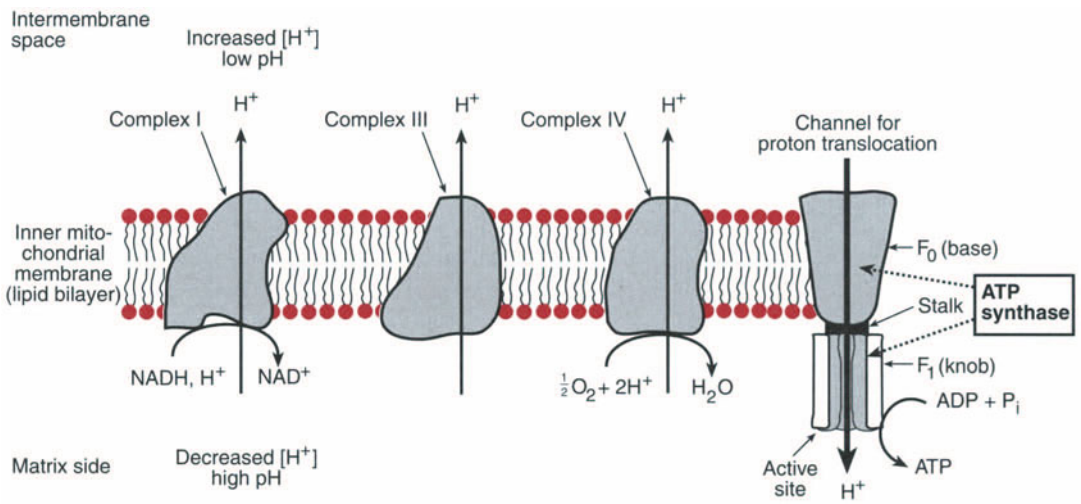


Figure 12.9. Localization of the ETS in the inner mitochondrial membrane. Three respiratory complexes pump protons from the membrane's matrix side to the intermembrane space. The resultant proton gradient drives ATP synthesis, catalyzed by ATP synthase. The enzyme, a transmembrane protein, has an active site on the matrix side and contains a channel through which protons move. Proton translocation is coupled to ATP synthesis. [Adapted from D. R. Ort and N. E. Good, *Trends Biochem. Sci.* 13:467–469 (1988) with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.]

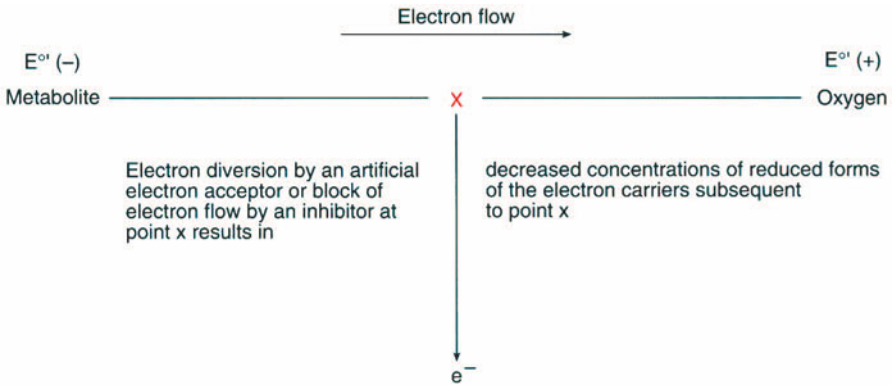


Figure 12.10. Principle of action of artificial electron acceptors and inhibitors when used with the ETS.

cal to that of adding an artificial electron acceptor. An inhibitor also leads to a *decrease in the concentration of reduced forms of electron carriers subsequent to the point of inhibition.* Therefore, points of action of inhibitors can likewise be pinpointed by spectrophotometric measurements of the electron carriers.

12.3.2D. Reduction Potentials of Electron Carriers. We generally discuss the sequence of electron carriers in terms of the carriers' biochemical standard reduction potentials ($E^{\circ\prime}$). Table 12.4 lists these, beginning with isocitrate as the metabolite undergoing oxidation to α-ketoglutarate:

$$\text{Isocitrate}^{3-} + \text{NAD}^+ \rightarrow \alpha\text{-ketoglutarate}^{2-} + \text{NADH} + \text{CO}_2$$

In the ETS, electrons flow from a metabolite, the initial electron donor, to oxygen, the terminal electron acceptor. In this flow, each carrier must be a stronger reducing agent than the one following it in the chain; electron flow must be from a smaller to a larger *biochemical actual reduction potential* (E'). As you can see from Table 12.4, the sequence of electron carriers is *essentially* in agreement with that predicted on the basis of their *biochemical standard reduction potentials* ($E^{\circ\prime}$).

The reason for the agreement between carrier sequence based on E' and that based on $E^{\circ\prime}$ derives from the intracellular concentrations of electron carriers. For most of the electron carriers, the concentrations of their oxidized and reduced forms are similar so that $E^{\circ\prime} \approx E'$ (see Eq. 12.2). Remember, of course, that *what really determine the sequential arrangement of the electron carriers in the ETS are their biochemical actual reduction potentials* (E'), the potentials existing under intracellular conditions of reactant and product concentrations. These must indeed progress consistently from the smaller to the larger potential.

Note what the above tells you about the efficiency of the ETS. An electron carrier has *maximum capacity* for accommodating *either* loss of electrons (oxidation) or addition of electrons (reduction) when the concentration of its reduced form equals that of its oxidized form. Therefore, the fact that $E^{\circ\prime} \approx E'$ means that the ETS operates rather efficiently.

We encountered an analogous situation when we considered the efficiency of a buffer (Section 1.3). Recall that a buffer has *maximum capacity* for addition of *either* acid or base when the concentration of its dissociated form equals that of the undissociated form: $[A^-] = [HA]$.

12.3.3. Energetics of the ETS

Electrons flow through the chain of ETS carriers much as electrons flow through a wire; the flow is a function of the potential drop. In the example of Table 12.4, the total po-

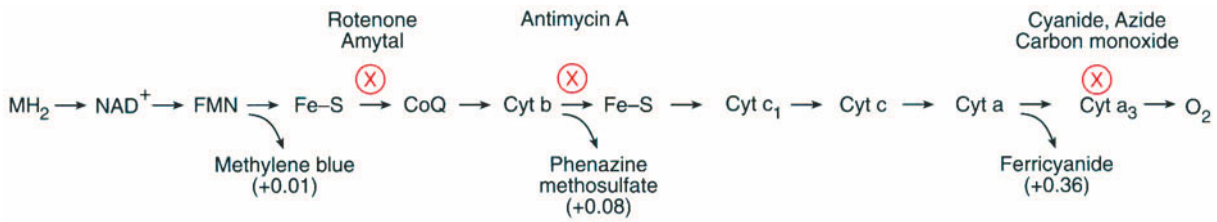


Figure 12.11. Points of action in the ETS of some specific artificial electron acceptors ($E^{\circ\prime}$ values in parentheses) and inhibitors (×).

Table 12.4. Reduction Potentials of Electron Carriers

Redox pair[a]	$E^{\circ\prime}$ (volts)	
α-Ketoglutarate/isocitrate	−0.38	
NAD^+/NADH	−0.32	
FMN/$FMNH_2$ (enzyme bound)	−0.30	
Fe^{3+}–S/Fe^{2+}–S (average)	−0.24	
CoQ/$CoQH_2$	+0.05	
Cyt *b* (Fe^{3+})/Cyt *b* (Fe^{2+})	+0.07	Electron flow
Fe^{3+}–S/Fe^{2+}–S	+0.28	
Cyt c_1 (Fe^{3+})/Cyt c_1 (Fe^{2+})	+0.22	
Cyt *c* (Fe^{3+})/Cyt *c* (Fe^{2+})	+0.25	
Cyt *a* (Fe^{3+})/Cyt *a* (Fe^{2+})	+0.29	
Cu^{2+}/Cu^+ (average)	+0.29	
Cyt a_3 (Fe^{3+})/Cyt a_3 (Fe^{2+})	+0.39	
$\frac{1}{2}O_2/H_2O$	+0.82	

[a]Fe–S, Iron–sulfur protein; Cyt, cytochrome.

tential drop from isocitrate to oxygen equals $+0.82 - (-0.38) = +1.20$ V. This difference in potential is equivalent to a change in free energy (Eq. 12.3). Since $\Delta E^{\circ\prime}$ is positive, $\Delta G^{\circ\prime}$ is negative. The entire ETS, therefore, represents an exergonic set of reactions accompanied by the release of free energy. Remember, though, that intracellularly $\Delta E'$ must be positive in order for $\Delta G'$ to be negative.

12.3.3A. ATP Synthesis. Free energy released during the operation of the electron transport system drives the synthesis of ATP. We call the combined reactions *oxidative phosphorylation.* Linking the exergonic ETS with the endergonic synthesis of ATP constitutes a special case of *energetically coupled reactions.* ATP synthesis from ADP and P_i requires an input of +30.5 kJ mol^{-1} (Table 9.2). In coupling ATP synthesis to the electron transport system, the ETS must generate 30.5 kJ mol^{-1} for every mole of ATP synthesized. You can compute the necessary potential drop, equivalent to 30.5 kJ mol^{-1}, from Eq. (12.3):

$$\Delta G^{\circ\prime} = -n\mathrm{F}\Delta E^{\circ\prime}$$

so that

$$-30{,}500 = -(2)(96{,}491)\Delta E^{\circ\prime}$$

and

$$\Delta E^{\circ\prime} = +0.16\ \mathrm{V}$$

Thus, a minimum potential drop of 0.16 V is required for synthesis of one mole (molecule) of ATP per mole (molecule) of metabolite oxidized. This requirement is met at the three respiratory complexes where proton pumping and ATP synthesis occur during oxidative phosphorylation (Figure 12.12). At each complex, the potential drop exceeds 0.16 V.

Because it takes 0.16 V to synthesize one mole of ATP, and because the total potential drop in the above sequence is 1.20 V, there could have occurred, theoretically, a synthesis of $1.20/0.16 \approx 8$ moles of ATP per mole of metabolite oxidized. This calculation assumes a 100% efficiency of energy conservation, an unlikely occurrence, especially in such a complex biological system. The experimentally determined maximum yield is 3 moles of ATP per mole of metabolite (that is, per mole NADH oxidized). Thus, of the 1.2 V available, only $3 \times 0.16 = 0.48$ V serve for ATP synthesis. Hence, we can calculate the efficiency of ATP synthesis as

$$\frac{3 \times 0.16\ \text{Volts}}{1.20\ \text{Volts}} \times 100 = 40\%$$

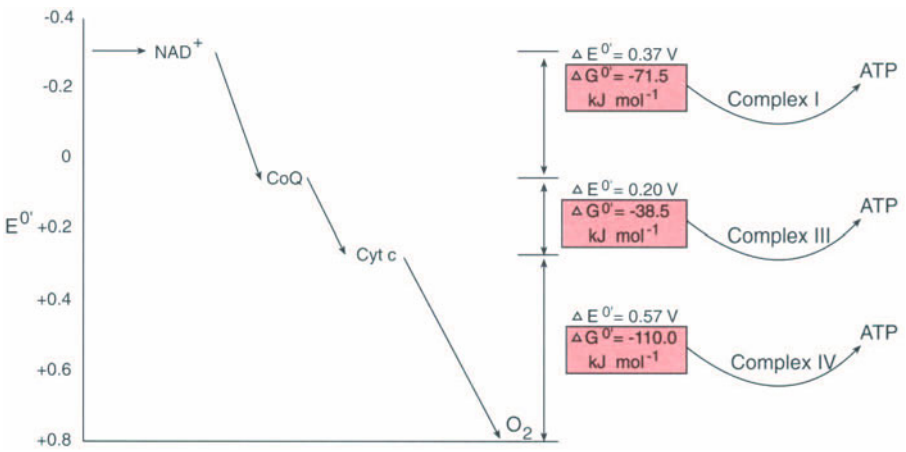


Figure 12.12. Schematic diagram of the energetics of the ETS. Potential drops at Complexes I, III, and IV generate sufficient energy for ATP synthesis. Actual ATP synthesis occurs at the active site of ATP synthase and is driven by the proton gradient produced by the respiratory complexes.

12.3.3B. Energy Conservation. The ETS, with its coupled ATP synthesis, illustrates an important principle of energy conservation. In order for ATP synthesis to be efficient, an ideal exergonic pathway should consist of a large number of steps, with each step releasing just enough energy to synthesize one mole of ATP. In the idealized case of isocitrate oxidation, the ETS should consist of eight steps, with each step having a potential drop of 0.16 V.

Two reasons account for constructing an idealized system in this fashion. First, larger potential drops, but not large enough to yield another mole of ATP, result in wasted energy. Assume that the potential drop of a step were 0.20 V, enough energy to synthesize one mole, or one molecule, of ATP (0.16 V) but not enough to synthesize two molecules. Because a fraction of a molecule cannot be synthesized, the excess energy (0.04 V) would be dissipated.

Second, multiples of the required potential drop, in principle available for the synthesis of several moles of ATP, likewise result in wasted energy. Consider a step that has a potential drop of 0.48 V, enough energy to synthesize three moles, or three molecules, of ATP (0.48/0.16 = 3). In order for this to occur, it would be necessary to have the more or less simultaneous interaction of a minimum of 11 entities (three molecules of ADP, three molecules of P_i, three protons, and two molecules of electron carriers). Such multiple molecular collisions are extremely unlikely. In practice, a step with a large potential drop results in synthesis of only one molecule of ATP while the excess energy is dissipated.

The ETS scheme approaches the ideal case in that it consists of a relatively large number of steps, with each step having only a small change in potential and free energy. For identical reasons, other metabolic pathways associated with the synthesis or utilization of ATP typically consist of many steps, with each step involving a relatively small change in free energy.

Keep in mind, however, that we made all of the above calculations using *biochemical standard reduction potentials* ($E^{\circ\prime}$) of the carriers. Our conclusions are, therefore, only tentative. Correct evaluations of the energetics must involve the use of *biochemical actual reduction potentials* (E').

12.4. OXIDATIVE PHOSPHORYLATION

12.4.1. P/O Ratio

Oxidative phosphorylation, *synthesis of ATP coupled to operation of the electron transport system,* constitutes the major mechanism for ATP synthesis in all nonphotosynthetic organisms. Coupling of these two processes was discovered early on in the development of biochemistry when researchers found that uptake of oxygen could be correlated with ATP synthesis. Manometric measurements of tissue homogenates and tissue slices show that, as oxygen is taken up, the amount of ATP increases concomitantly. You can assess ATP synthesis by measuring ATP's incorporation of radioactively labeled inorganic phosphate. Under optimal conditions, three *molecules* of phosphate are taken up—that is, three molecules of ATP are synthesized—for every one *atom* of oxygen consumed. We can characterize these results by a *P/O ratio* (or a $P/2e^-$ ratio) of 3. We now know that ATP synthesis is coupled to proton pumping of three respiratory complexes (Figure 12.9).

P/O ratios may have values less than 3 depending upon the point at which metabolite oxidation links up with the ETS. The succinate pathway, for example, connects with the main ETS sequence at the level of CoQ. This means that only Complexes III and IV are available for proton pumping and ATP synthesis. These complexes span a smaller potential drop than the complete ETS, which includes complex I. Under these conditions, only two molecules of ATP are synthesized per atom of oxygen consumed (that is, per $FADH_2$ oxidized), and the P/O ratio equals 2.

A metabolite whose oxidation is linked to the ETS after respiratory complex III yields a P/O ratio of 1. Ascorbate can donate electrons to tetramethyl-*p*-phenylenediamine (TMPD), which can reduce cytochrome *c* directly. Thus, addition of ascorbic acid and TMPD to an electron transport system results in a P/O ratio of 1.

12.4.2. Chemiosmotic Coupling

Coupling of ATP synthesis to the operation of the electron transport system has been the subject of intensive research. The earliest proposal, called the *chemical coupling hypothesis,* suggested a classical mechanism of energetically coupled reactions. According to this hypothesis, operation of the ETS results in formation of covalent *energy-rich compounds* that function as *common intermediates* to drive the synthesis of ATP.

A subsequent proposal, termed the *conformational coupling hypothesis,* postulated that operation of the ETS results in an energized conformation of one or more proteins of the inner mitochondrial membrane. When the membrane proteins return to their low-energy conformation, the energy released energizes the enzyme catalyzing ATP synthesis.

Investigators attempted for many years to isolate the energy-rich intermediates postulated by these hypotheses but were unsuccessful. A completely different mechanism was proposed by Peter Mitchell in 1961 (Nobel Prize,

1978). Mitchell's proposal, termed the **chemiosmotic coupling hypothesis,** has received extensive experimental support and constitutes the currently accepted mechanism of oxidative phosphorylation.

According to this hypothesis, the driving force for ATP synthesis consists of an *electrochemical gradient* that comprises both a pH and an electrical potential component. *Pumping* of protons across the inner mitochondrial membrane by the respiratory complexes sets up a pH gradient across the membrane; the matrix side becomes more basic, and the intermembrane space becomes more acidic. Complexes I, III, and IV contribute to proton gradient formation (Figure 12.13). In Complex III, a molecule of coenzyme Q is involved twice in proton translocation. The resultant mechanism, called the **Q-cycle,** involves the semiquinone form of coenzyme Q. The Q-cycle results in the transport of $4H^+$ across the mitochondrial membrane and the transfer of a pair of electrons from $CoQH_2$ to cytochrome *c*.

Complex II does not contribute to the proton gradient but serves to channel electrons into the ETS at the level of coenzyme Q (Figure 12.14).

Movement of the positive charges across the membrane also generates an electrical potential. The outer (intermembrane space) side becomes positive relative to the inner (matrix) side, developing a potential of about 0.15 V. The concentration gradient (ΔpH) produces an electrical potential across the membrane. The combined effects of the difference in pH and the difference in potential constitute an electrochemical gradient. Note that the mitochondrial membrane is impermeable to ions such as H^+, OH^-, K^+, and Cl^- whose free diffusion would discharge the potential of such a gradient. We can describe the electrochemical gradient by the equation:

$$\Delta\mu_H = \Delta\psi - 2.303RT\Delta pH/F$$

where $\Delta\mu_H$ is the electrochemical proton potential, $\Delta\psi$ is the membrane potential (potential gradient), the ΔpH term is the proton potential (pH gradient), and F is the Faraday constant.

According to the chemiosmotic hypothesis, dissipation of the electrochemical gradient drives ATP synthesis. Recall that ATP synthesis from ADP requires an input of protons (Figure 9.4). The pH gradient furnishes these protons. As the protons move back across the inner mitochondrial membrane, both the pH gradient and the poten-

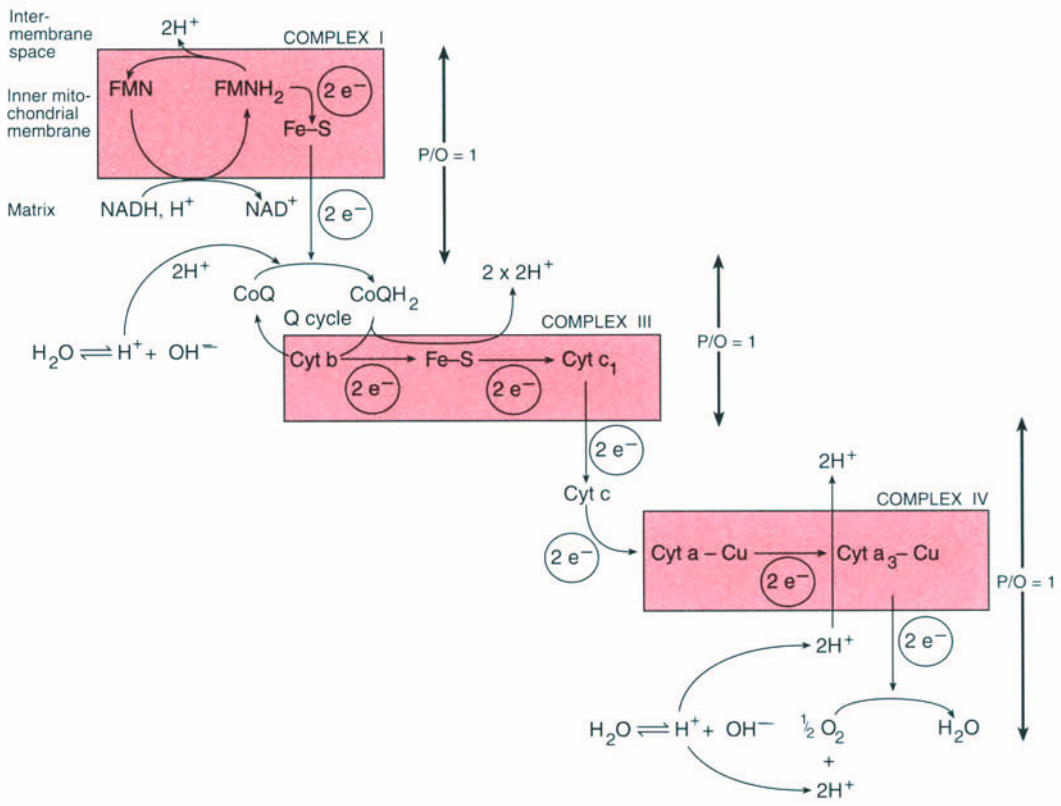


Figure 12.13. Proposed electron flow through Complexes I, III, and IV. Electrons pass from NADH to coenzyme Q via FMN and Fe-S, then via the Q-cycle to cytochrome *b* and Fe-S, and finally through a series of other cytochromes to oxygen. (2e) denotes two one-electron transfers.

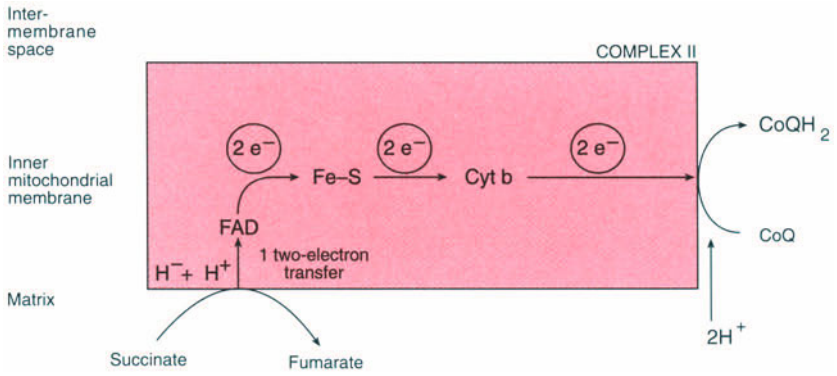


Figure 12.14. Proposed electron flow through Complex II. (2e) denotes two one-electron transfers.

tial difference across the membrane are being lost. The entire electrochemical gradient is being dissipated at the expense of ATP synthesis. Thus, the gradient serves as the *common intermediate* for the coupled processes of electron transport and ATP synthesis.

12.4.3. ATP Synthase

Actual ATP synthesis occurs at the active site of **ATP synthase,** a complex protein oligomer, located in the inner mitochondrial membrane. Because it consists of two major parts (F_0 and F_1) and serves to couple the ETS and ATP synthesis, the synthase is known as **F_0F_1-coupling factor.** Lastly, because the synthase, like all enzymes, catalyzes both the forward and the reverse reaction, we also refer to the enzyme complex as *ATPase, F_0F_1-ATPase,* or *proton-pumping ATPase.*

The enzyme complex has the appearance of a knob attached to a base by means of a stalk (see Figure 12.9). The base (F_0) is embedded in and spans the membrane; the knob (F_1) protrudes into the matrix. F_0 is a water-insoluble transmembrane protein that consists of four different polypeptide chains and contains a channel for proton translocation. F_1 is a water-soluble peripheral membrane protein. It consists of five different polypeptide chains in the ratio of $\alpha_3\beta_3\gamma\delta\epsilon$ and contains the active site. The stalk consists of two proteins.

Total molecular weight of the F_0F_1-ATPase is 450,000. *Oligomycin,* an antibiotic produced by *Streptomyces,* is an inhibitor of the enzyme. The antibiotic binds to a subunit of F_0, thereby interfering with proton transport through F_0. Investigators believe that proton movement through the channel in F_0 drives ATP synthesis by ATP synthase. Exactly how the enzyme accomplishes the synthesis is not yet clear. It appears that the mechanism does not consist of merely shifting the equilibrium toward ATP synthesis (Figure 9.4). Instead, researchers have proposed that the flow of protons triggers a conformational change in the enzyme, possibly by means of cooperative interactions among multiple identical active sites.

12.5. CONTROL MECHANISMS

12.5.1. Respiratory Control

Under most physiological conditions, synthesis of ATP is *tightly coupled* to respiration and to operation of the ETS. ATP synthesis is absolutely dependent upon the flow of electrons, and electrons do not normally flow through the ETS unless ADP is being phosphorylated to ATP (see Section 12.5.3 for an exception). We call this regulatory phenomenon **respiratory control,** and its essence lies in the intracellular level of ADP. Active, energy-consuming cells use up ATP and accumulate ADP. The resultant high level of ADP stimulates respiration and enhances the activity of the ETS. Conversely, in resting and well-nourished cells, ATP accumulates at the expense of ADP. Depletion of ADP limits respiration and decreases the activity of the ETS.

Respiratory control constitutes one of the major overall control mechanisms in metabolism. It represents the link between cellular requirements for ATP and the rate at which foodstuffs are oxidized via the ETS.

12.5.2. Energy Charge

Both hydrolysis and synthesis of ATP involve, in addition to ATP, either ADP or AMP. These three adenine nucleotides are interconvertible by means of **adenylate kinase,** which catalyzes a *disproportionation reaction* in which a single reactant yields two different products:

$$2ADP \rightleftarrows ATP + AMP$$

A cell's capacity to carry out ATP-driven reactions depends on the relative concentrations of AMP, ADP, and ATP. A mathematical measure of this capacity was pro-

posed by Daniel Atkinson and is called the **energy charge:**

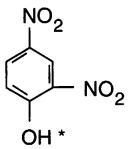
$$\text{Energy charge} = \frac{[\text{ATP}] + \tfrac{1}{2}[\text{ADP}]}{[\text{ATP}] + [\text{ADP}] + [\text{AMP}]} \qquad (12.6)$$

The numerator of this expression consists of the two energy-rich forms of adenine nucleotides. We multiply the ADP concentration by 0.5 because ADP has one-half the number of energy-rich bonds present in ATP. The denominator represents the total concentration of all adenine nucleotides. Values of the energy charge vary from 0 (all AMP) to 1.0 (all ATP). Most normal cells operate in an energy charge range of 0.8–0.9.

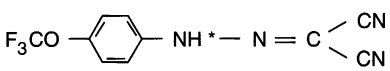
The energy charge exerts its control in metabolism through allosteric regulation of specific enzymes by AMP, ADP, and ATP. You saw previously that high concentrations of ATP inhibit energy-generating pathways such as glycolysis and the citric acid cycle. High concentrations of ADP or AMP, on the other hand, tend to stimulate such pathways. Thus, a large energy charge reflects high intracellular levels of ATP and inhibits ATP-generating pathways. A small energy charge indicates high intracellular levels of ADP and AMP and stimulates ATP-generating pathways.

12.5.3. Uncouplers of Oxidative Phosphorylation

A number of compounds are known that can disconnect electron transport from ATP synthesis. Such **uncouplers** permit electron transport to proceed but prevent ATP synthesis. Figure 12.15 shows the structure of two uncouplers, *2,4-dinitrophenol* and carbonylcyanide-*p*-trifluoromethoxyphenylhydrazone. You can understand their mechanism of action on the basis of chemiosmotic coupling. Both uncouplers are nonpolar and therefore readily pass through the inner mitochondrial membrane. They are also weak acids and exist as anions at physiological pH. The anionic form of the uncoupler binds protons on the acidic (intermembrane space) side of the membrane. It then diffuses across the membrane and releases the protons on the alkaline (matrix) side. In so doing, the uncoupler leads to dissipation of the electrochemical gradient, resulting in loss of the capacity to synthesize ATP.



2,4-Dinitrophenol



Carbonylcyanide-*p*-trifluoromethoxyphenylhydrazone

Figure 12.15. Some uncouplers of oxidative phosphorylation. The asterisk designates an ionizable proton.

Some antibiotics, such as valinomycin and gramicidin A (see Figures 6.34 and 6.35), also function as uncouplers. You can explain their action by chemiosmotic coupling as well. Recall that these antibiotics act as *ionophores* and mediate transport of cations across the mitochondrial membrane. Movement of cations across the membrane diminishes the transmembrane potential. Movement of protons decreases the pH gradient as well. One or both of these factors lead to dissipation of the electrochemical gradient and a resultant loss of the capacity to synthesize ATP.

Some uncoupling of oxidative phosphorylation occurs naturally in newborn mammals that lack hair (for example, humans), hibernating animals, and those adapted to cold weather. In these instances, uncoupling of ATP synthesis is desirable because dissipation of the electrochemical gradient generated by the ETS leads to heat production. The process occurs under hormonal control and in special tissues, called *brown fat* or *brown adipose tissue,* located in the neck and upper back. The name derives from the fact that the tissues are rich in cytochromes. Mitochondria of brown fat have become specialized to generate heat from the oxidation of fatty acids.

12.5.4. Incomplete Reduction of Oxygen

As you saw, the complete reduction of oxygen by cytochrome oxidase requires four electrons (Eq. 12.5a). However, other reactions of oxidative metabolism frequently produce partially reduced forms of oxygen. Any source of electrons, such as the thiol group of cysteine or the reduced form of vitamin C, can readily reduce oxygen to form oxygen radicals. One-, two-, and three-electron reductions of O_2 yield the *superoxide anion* ($O_2\cdot^-$), *hydrogen peroxide* (H_2O_2), and the *hydroxyl radical* ($OH\cdot$), respectively:

$$O_2 + e^- \rightarrow O_2\cdot^-$$
$$O_2 + 2e^- + 2H^+ \rightarrow H_2O_2$$
$$O_2 + 3e^- + 3H^+ \rightarrow H_2O + OH\cdot$$

Incompletely reduced forms of oxygen are extremely reactive; they constitute powerful oxidizing agents that possess high toxicity for living systems. The hydroxyl radical, in particular, represents the most potent oxidizing agent known and the most active mutagen produced by ionizing radiation.

These highly active forms of oxygen must be converted to less reactive ones if the organism is to survive.

Several major self-defense mechanisms protect aerobic cells from the ravages of incompletely reduced oxygen. These involve the enzymes *superoxide dismutase, catalase,* and *peroxidase.*

The primary mode for detoxifying the superoxide anion involves its conversion to hydrogen peroxide by action of superoxide dismutase. This enzyme catalyzes a *dismutation reaction*—a reaction in which two identical substrates have different fates. In the superoxide dismutase reaction, one superoxide anion undergoes oxidation while the other undergoes reduction:

$$O_2\cdot^- + O_2\cdot^- + 2H^+ \rightarrow H_2O_2 + O_2$$

The superoxide anion can also give rise to hydrogen peroxide via a second pathway. Protonation of $O_2\cdot^-$ yields the *hydroperoxyl radical,* $HO_2\cdot$, the conjugate acid of the superoxide anion:

$$O_2\cdot^- + H^+ \rightarrow HO_2\cdot$$

Two hydroperoxyl radicals combine spontaneously to form hydrogen peroxide:

$$HO_2\cdot + HO_2\cdot \rightarrow H_2O_2 + O_2$$

Hydrogen peroxide formed from the superoxide anion, or produced by other metabolic reactions, is detoxified enzymatically. Catalase catalyzes the decomposition of hydrogen peroxide without requiring a second substrate

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

while peroxidase destroys hydrogen peroxide with the aid of an added electron donor (AH_2)

$$AH_2 + H_2O_2 \rightarrow 2H_2O + A$$

Glutathione peroxidase represents an important peroxidase and one of a small number of enzymes that contain selenium. The enzyme occurs in erythrocytes, where it catalyzes the decomposition of H_2O_2 coupled to the oxidation of glutathione (GSH):

$$2GSH + H_2O_2 \rightarrow GSSG + 2H_2O$$

Glutathione peroxidase performs an essential function in protecting erythrocytes against the accumulation of peroxides. The enzyme glutathione reductase subsequently catalyzes the regeneration of GSH from GSSG.

The major mechanism for formation of the hydroxyl radical requires both a superoxide anion and hydrogen peroxide:

$$O_2\cdot^- + H_2O_2 \rightarrow OH\cdot + OH^- + O_2$$

Accordingly, effective scavenging of $O_2\cdot^-$ and H_2O_2 not only removes these two harmful oxidizing agents, but also prevents their giving rise to the even more dangerous hydroxyl radical. However, in the absence of a 100% scavenging efficiency, and as a result of other reactions, some hydroxyl radicals may form in living systems. In that event, they are likely to participate in three main reactions:

—oxidation of metal ions (M) to higher oxidation states:

$$OH\cdot + M^{n+} \longrightarrow (M\,OH)^{n+} \xrightarrow{H^+} M^{n+1} + H_2O$$

—abstraction of a hydrogen from a C-H bond to produce water and an organic radical

—addition to a double bond to form secondary radicals

Because of the hazards that incompletely reduced forms of oxygen pose for humans, they have been implicated by some in the development of cancer and other ailments. Consequently, the claim has been made that dietary supplements of antioxidant vitamins (A, C, and E) are beneficial in the battle against these diseases.

12.6. BALANCE SHEET OF CARBOHYDRATE CATABOLISM

Having covered the electron transport system, the final stage of aerobic metabolism, we can now evaluate the overall energetics of carbohydrate catabolism. To do so, we need to consider the combined operation of glycolysis, the citric acid cycle, and the electron transport system.

The maximum energy yield from glycolysis, under anaerobic conditions, is a net of 2 ATP formed per molecule of glucose catabolized, for the sequence glucose to pyruvate or lactate (Section 10.3). The maximum energy yield from the citric acid cycle/ETS, functioning aerobically, is 12 ATP formed per molecule of acetyl CoA catabolized (Section 11.5). A net energy yield of *2 ATP* from the catabolism of a *six-carbon compound* in glycolysis is much smaller than a yield of *12 ATP* from the catabolism of a *two-carbon compound* in the citric acid cycle. As pointed out earlier, the reason for the difference in energy yield lies in the extent to which the metabolite undergoes oxidation. Conversion of glucose to pyruvate or lactate represents a small degree of oxidation, but conversion of the acetyl group of acetyl CoA to CO_2 and H_2O represents complete oxidation.

Now consider the total energy yield of carbohydrate

catabolism under optimal circumstances in which glycolysis proceeds under aerobic conditions. Pyruvate, produced by glycolysis, forms acetyl CoA. Catabolism of acetyl CoA via the citric acid cycle and its link to the ETS results in ATP synthesis. Additionally, NADH produced in glycolysis is oxidized via the electron transport system and also leads to production of ATP. Combined operation of glycolysis, the citric acid cycle, and the ETS represents full catabolism of carbohydrates and results in *complete oxidation of glucose to CO_2 and H_2O.*

12.6.1. Theoretical Energy Yields

Pyruvate, produced in the cytosol, represents the end product of glycolysis under aerobic conditions. Catabolism of glucose to pyruvate yields a net of 2 ATP. Pyruvate is transported across the mitochondrial membrane and converted *inside the mitochondria* to acetyl CoA by the pyruvate dehydrogenase complex. Acetyl CoA enters the citric acid cycle, located *inside the mitochondria,* and leads to production of 12 ATP/acetyl CoA. Conversion of pyruvate to acetyl CoA by the pyruvate dehydrogenase complex (Section 11.2) results in formation of NADH, which is oxidized via the ETS, located *inside the mitochondria,* yielding 3 ATP/NADH. Since glucose → 2 pyruvate, a total of 2 acetyl CoA (24 ATP) and 2 NADH (6 ATP) are produced per molecule of glucose.

Additionally, 2 NADH are produced per molecule of glucose in the glyceraldehyde 3-phosphate dehydrogenase reaction, the first step of stage II in glycolysis. Assume for a moment that the glycolytic NADH is readily oxidized via the ETS, yielding 3 ATP/NADH, or a total of 6 ATP/glucose (see below for a correction). In that case, the maximum energy yield of carbohydrate metabolism would be as follows:

Glucose → pyruvate (glycolysis)	2 ATP
2 (NADH, H^+) from glycolysis (via ETS)	6 ATP
2 (NADH, H^+) from pyrvate → acetyl CoA (via ETS)	6 ATP
2 Acetyl CoA (citric acid cycle)	24 ATP
Total	38 ATP

Thus, 38 molecules (moles) of ATP are formed per molecule (mole) of glucose catabolized. Because $\Delta G^{\circ\prime} = -2870\ \text{kJ mol}^{-1}$ for glucose oxidation and $\Delta G^{\circ\prime} = -30.5\ \text{kJ mol}^{-1}$ for ATP hydrolysis, carbohydrate catabolism would have an efficiency of

$$\frac{(38)\ (30.5\ \text{kJ mol}^{-1})}{2870\ \text{kJ mol}^{-1}} \times 100 = 40\%$$

similar to that of the ETS alone (Section 12.3). By comparison, the maximum efficiency of a steam engine, operating between the boiling point of water and room temperature, is only about 22%. Keep in mind, though, that these calculations are based on $\Delta G^{\circ\prime}$ values.

12.6.2. Shuttle Systems

A yield of 38 ATP/glucose represents a *maximum* value. The actual number of ATP produced may be smaller because of a *compartmentation* problem that we ignored in our calculations. Glycolysis occurs in the cytosol, but the ETS is located inside the mitochondria, in the inner membrane. As stressed above, conversion of pyruvate to acetyl CoA, catabolism of acetyl CoA via the citric acid cycle, and oxidation of NADH produced by the pyruvate dehydrogenase complex all occur inside the mitochondria and result in ATP production as outlined. By contrast, the NADH generated by glycolysis cannot be oxidized directly by the ETS as we assumed. For NADH to be oxidized via the ETS, it must first be transported from the cytosol into the mitochondria. Yet transport is impossible, because the membranes of animal mitochondria are *impermeable* to NADH and NAD^+. How then can glycolytic NADH be used?

To circumvent this compartmentation problem in animal tissues (yeast mitochondria are permeable to NADH and NAD^+), a number of **shuttle mechanisms** have evolved that accomplish oxidation of NADH in an indirect way. Figure 12.16 outlines one such mechanism, called the **glycerol phosphate shuttle.** The shuttle occurs in certain muscle and nerve cells, and its operation involves two compounds, *dihydroxyacetone phosphate,* a glycolytic intermediate, and *glycerol 3-phosphate.*

In this mechanism, NADH generated in glycolysis is oxidized in the *cytosol* by action of an *NAD^+-linked dehydrogenase.* Oxidation of NADH is coupled to reduction of dihydroxyacetone phosphate to glycerol 3-phosphate. The latter diffuses to the *inner mitochondrial membrane,* where it undergoes oxidation to dihydroxyacetone phosphate by action of an *FAD-linked dehydrogenase.* The active site of the mitochondrial dehydrogenase is on the outer face of the inner mitochondrial membrane. As glycerol 3-phosphate undergoes oxidation, FAD is reduced to $FADH_2$.

The net result of these reactions amounts to a *reduction of FAD to $FADH_2$ inside the mitochondria* at the expense of an *oxidation of NADH to NAD^+ in the cytosol.* In effect, electrons are transported across the mitochondrial membrane while carrier concentrations remain constant. $FADH_2$, formed inside the mitochondria, enters the ETS directly but yields only 2 ATP (P/O = 2) compared to the 3 ATP (P/O = 3) produced by NADH. Thus, the cell pays a price for operation of this shuttle; instead of deriving 3 ATP/glycolytic NADH, it only derives 2 ATP.

A different shuttle, the **malate–aspartate shuttle,** functions in liver and heart tissue (Figure 12.17). It may appear from the figure that this shuttle operates without

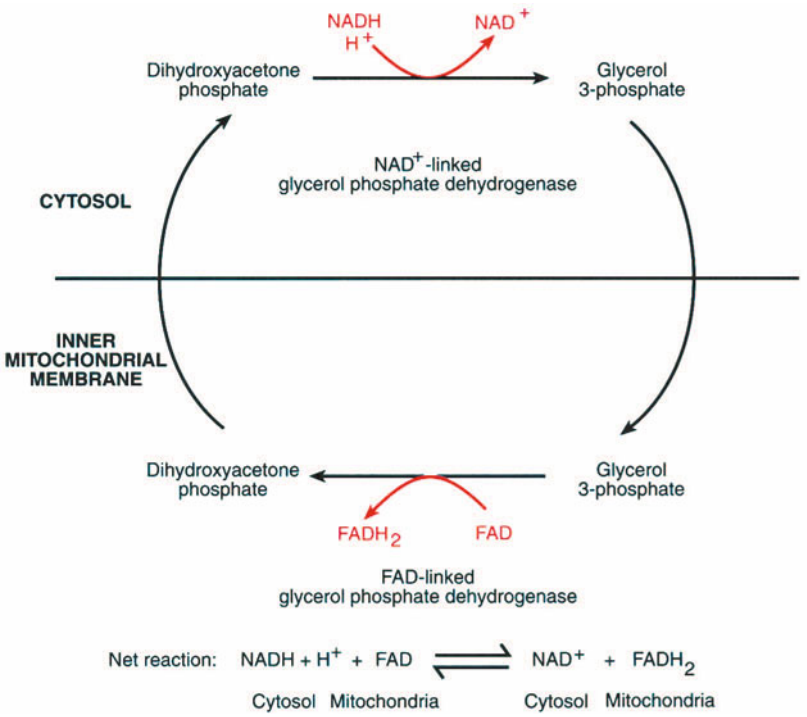


Figure 12.16. The glycerol phosphate shuttle.

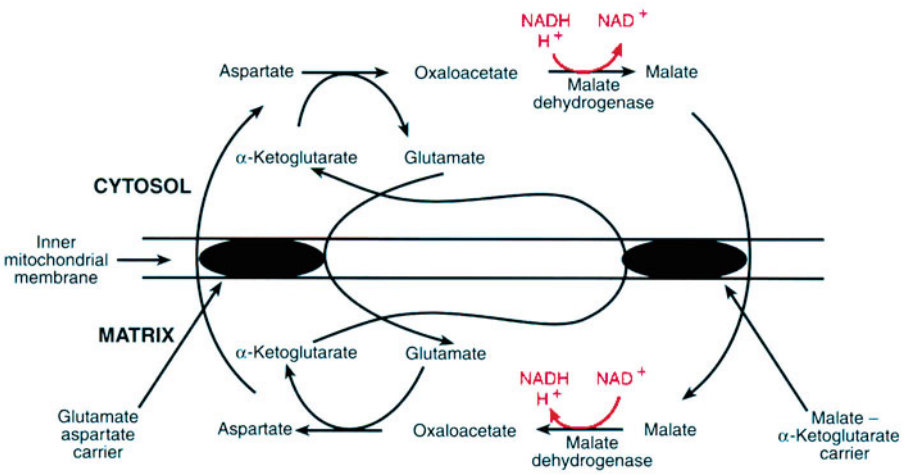


Figure 12.17. The malate–aspartate shuttle.

loss of efficiency, producing mitochondrial NADH at the expense of cytosolic NADH. This is incorrect. Operation of the shuttle involves a proton motive force estimated to be equivalent to about 0.5 ATP. Thus, the true ATP production of the shuttle amounts to about 2.5 ATP per NADH (P/O = 2.5).

12.6.3. Actual Energy Yields

In view of the need for some type of shuttle mechanism, the true energy yield of carbohydrate catabolism is less than the theoretical yield calculated above and varies depending on the shuttle involved. Assuming operation of the *glycerol phosphate shuttle*, we calculate the maximum energy yield of carbohydrate catabolism under aerobic conditions as follows:

Glucose → pyruvate (glycolysis)	2 ATP
2 (NADH, H^+) from glycolysis (via glycerol phosphate shuttle and ETS)	4 ATP
2 (NADH, H^+) from pyruvate → acetyl CoA(via ETS)	6 ATP
2 Acetyl CoA (citric acid cycle)	24 ATP
Total	36 ATP

Hence, the efficiency of energy conservation is:

$$\frac{(36)\ (30.5\ \text{kJ mol}^{-1})}{2870\ \text{kJ mol}^{-1}} \times 100 = 38\%$$

Likewise, we calculate the maximum energy yield, assuming operation of the malate–aspartate shuttle (P/O = 2.5), as follows:

Glucose → pyruvate (glycolysis)	2 ATP
2 (NADH, H^+) from glycolysis (via malate-asparate shuttle and ETS)	5 ATP
2 (NADH, H^+) from pyruvate → acetyl CoA(via ETS)	6 ATP
2 Acetyl CoA (citric acid cycle)	24 ATP
Total	37 ATP

Hence, the efficiency of energy conservation is:

$$\frac{(37)\ (30.5\ \text{kJ mol}^{-1})}{2870\ \text{kJ mol}^{-1}} \times 100 = 39\%$$

SUMMARY

We define the biochemical standard reduction potential ($E^{\circ\prime}$) as the potential of a reaction (at pH 7) in which initial reactant and product concentrations are 1.0*M* each. We define the biochemical actual reduction potential (E') as the potential of a reaction (at pH 7) in which the initial concentrations of reactants and products are not 1.0*M* each. $E^{\circ\prime}$ is related to the equilibrium constant of the reaction, and we compute E' by means of the Nernst equation. Changes in reduction potentials are equivalent to changes in free energy.

A redox reaction comprises two half-reactions. The half-reaction having the smaller reduction potential represents the stronger reducing agent, provides electrons to the other half-reaction, and proceeds as an oxidation. We deduce the direction of an overall redox reaction from $E^{\circ\prime}$ or E' values of the component two half-reactions. E' values alone determine the directions of overall reactions under intracellular conditions.

In cellular respiration, metabolites are oxidized by removal of hydride ions and protons. The electrons are then passed along a series of electron carriers. The mitochondrial electron transport system (ETS) contains five types of electron carriers: pyridine-linked dehydrogenases, flavoproteins, ubiquinones, cytochromes, and iron–sulfur proteins. Some carriers have the capacity for both one- and two-electron transfers; other carriers only transfer either one or two electrons. Two pathways of electron transport occur; in both, a metabolite serves as the initial electron donor, and molecular oxygen functions as the ultimate electron acceptor.

The sequence of electron carriers in the ETS was elucidated by isolating macromolecular complexes catalyzing portions of the entire sequence and by using artificial electron acceptors and inhibitors. The change in free energy resulting from the change in reduction potential

along the ETS drives ATP synthesis. Oxidative phosphorylation refers to this coupling of ATP synthesis and electron transport. According to the chemiosmotic coupling hypothesis, an electrochemical gradient functions as a common intermediate for coupling the two processes.

Maximally, three molecules of ATP are synthesized per atom of oxygen consumed, yielding a P/O ratio of 3. P/O ratios less than 3 result when metabolites are linked to the ETS at other than the initial point. ATP synthase, a complex membrane-bound enzyme, accomplishes the actual synthesis of ATP. The electrochemical gradient produced by proton pumping of three respiratory complexes drives ATP synthesis. Uncouplers of oxidative phosphorylation permit electron transport to occur but prevent ATP synthesis. Incompletely reduced species of oxygen form during oxidative metabolism and are eliminated by specific enzymatic mechanisms. ATP yield from complete oxidation of carbohydrates depends on the shuttle involved in the oxidation of glycolytic NADH. When the glycerol phosphate shuttle is operative, a maximum number of 36 ATP is produced per molecule of glucose oxidized to CO_2 and H_2O.

SELECTED READINGS

Abrahams, J. P., Leslie, A. G. W., Lutter, R., and Walker, J. E., Structure at 2.8Å resolution of F_1-ATPase from bovine heart mitochondria, *Nature (London)* 370:621–628 (1994).

Bensasson, R. V., Land, E. J., and Truscott, T. G., *Excited States and Free Radicals in Biology and Medicine,* Oxford University Press, Oxford (1993).

Fetter, J. R., *et al.,* Possible proton relay pathways in cytochrome *c* oxidase, *Proc. Natl. Acad Sci. USA* 92:1604–1608 (1995).

Fridovich, I., Superoxide radical and superoxide dismutases, *Annu. Rev. Biochem.* 64:97–112 (1995).

Gray, H. B., and Winkler, J. R., Electron transfer in proteins, *Annu. Rev. Biochem.* 65:537–561 (1996).

Mitchell, P., Keilin's respiratory chain concept and its chemiosmotic consequences, *Science* 206:1148–1159 (1979).

Musser, S. M., and Stowell, M. H. B., Cytochrome *c* oxidase: Chemistry of a molecular machine, *Adv. Enzymol. Relat. Areas Mol. Biol.* 71:79–208 (1995).

Pedersen, P. L., and Amzel, L. M., ATP synthases: Structure, reaction center, mechanism, and regulation of one of nature's most unique machines, *J. Biol. Chem.* 268:9937–9940 (1993).

Racker, E., From Pasteur to Mitchell: A hundred years of bioenergetics, *Fed. Proc.* 39:210–215 (1980).

Rees, D. C., and Farrelly, D., Biological electron transfer, in *The Enzymes,* 3rd ed. (D. S. Sigman and P. D. Boyer, eds.), Vol. 19, pp. 38–97, Academic Press, New York (1990).

Stenesh, J., *Core Topics in Biochemistry,* Cogno Press, Kalamazoo, MI (1993).

Trumpower, B. L., and Gennis, R. B., Energy transduction by cytochrome complexes in mitochondrial and bacterial respiration: The enzymology of coupling electron transfer reactions to transmembrane proton translocation, *Annu. Rev. Biochem.* 63:675–716 (1994).

REVIEW QUESTIONS

A. Define each of the following terms:

Oxidative phosphorylation	Uncoupler
Respiratory complex	Shuttle mechanism
Q-cycle	Respiratory control
Iron–sulfur protein	Coenzyme Q

B. Differentiate between the two terms in each of the following pairs:

$E^{\circ\prime}/E'$	Cytochrome/cytochrome oxidase
Glycerol phosphate shuttle/ malate–aspartate shuttle	Respiratory complex/ respiratory chain

C. (1) Outline the chemiosmotic coupling hypothesis.
(2) What is the effect of adding an inhibitor or an artificial electron acceptor to the electron transport system? Does the P/O ratio change in the presence of such a compound? How can you determine the point of action of the compound?
(3) Explain how efficiency of energy conservation, in terms of ATP synthesis in oxidative phosphorylation, can be calculated by using either free energy changes or changes in redox potentials. Why should an ideal catabolic pathway consist of a large number of steps?
(4) How do the reduction potentials (E') of the electron carriers change along the ETS? Why must this be so?
(5) What steps comprise the four-electron reduction of O_2 catalyzed by cytochrome oxidase? What is the significance of the cytochrome oxidase reaction?

PROBLEMS

12.1. What effects does a relatively low concentration of 2,4-dinitrophenol have on the rate of electron transport (oxygen consumption) and on the P/O ratio?

12.2. A researcher maintains a solution of cytochrome *c* at a potential of 0.50 V, a temperature of 25°C, and a pH of 7.0. What percent of cytochrome *c* will be in the reduced form?

12.3. Write the overall reaction that will occur when the following two half-reactions are coupled under biochemical standard conditions:

$$\text{Fumarate}^{2-} + 2H^+ + 2e^- \rightleftharpoons \text{succinate}^{2-} \qquad E^{\circ\prime} = +0.03 \text{ Volts}$$

$$\text{Dehydroascorbic acid} + 2H^+ + 2e^- \rightleftharpoons \text{ascorbic acid} \qquad E^{\circ\prime} = +0.06 \text{ Volts}$$

Calculate the value of the overall $\Delta E^{\circ\prime}$, $\Delta G^{\circ\prime}$, and K'_{bio}.

12.4. What is the total biochemical standard free energy change ($\Delta G^{\circ\prime}$) when 1.50 g of succinate (MW = 117) undergoes reaction according to the overall equation of the previous problem?

12.5. Assume that you couple the two half-reactions of Problem 12.3 at pH 7.0, using the following initial concentrations: [fumarate] = $5.0 \times 10^{-1} M$; [succinate] = $2.0 \times 10^{-2} M$; [dehydroascorbic acid] = $1.0 \times 10^{-4} M$; [ascorbic acid] = $3.0 \times 10^{-1} M$. (a) What overall reaction will take place under these conditions? (b) Calculate the change in reduction potential ($\Delta E'$) and the change in free energy ($\Delta G'$) for the overall reaction.

12.6.* What [succinate]/[fumarate] ratio is required in order to drive the reaction in the opposite direction to that deduced for Problem 12.5 when the initial concentration of dehydroascorbic acid is $5.0 \times 10^{-4} M$ and that of ascorbic acid is $2.0 \times 10^{-1} M$? (Hint: Set $E'_{succinate} = E'_{ascorbate}$.)

12.7. Calculate the [lactate]/[pyruvate] ratio when the half-reaction describing this system has an E' value of zero (see Table 12.1).

12.8.* Cyanide is a powerful poison because it binds avidly to the Fe^{3+} of cytochrome oxidase. Hemoglobin (Fe^{2+}) has a relatively low affinity for cyanide, but methemoglobin (Fe^{3+}) binds cyanide strongly. Sublethal cyanide poisoning may be reversed by immediate administration of nitrites, strong oxidizing agents capable of oxidizing the Fe^{2+} in hemoglobin to the Fe^{3+} state. Why is this treatment effective?

12.9.* What [succinate]/[fumarate] ratio would be required to enable succinate dehydrogenase to use NAD^+ as a coenzyme instead of its normal FAD under biochemical standard conditions?

12.10. Why is it that some soft drinks, with ice, can help keep you warm even on a very cold day? Why is this not the case with all soft drinks (diet drinks)?

12.11. What is the energy charge for a system in which adenine nucleotides are present entirely as ADP?

12.12. Which cell is likely to have a higher energy charge, one at rest or one actively metabolizing? Why?

12.13. Calculate the efficiency of energy trapping, in the form of ATP synthesis from ADP, when succinate is oxidized by $FADH_2$ via the second pathway of electron transport. Base your calculation on the $E^{\circ\prime}$ values of the pathway. The $\Delta G^{\circ\prime}$ of ATP hydrolysis is -30.5 kJ mol^{-1}.

12.14. Refer to Figure 12.8 and Table 12.1 and calculate the efficiency of energy trapping, in the form of ATP synthesis from ADP, for respiratory complexes I, III, and IV. In other words, consider the following sequences under biochemical standard conditions:

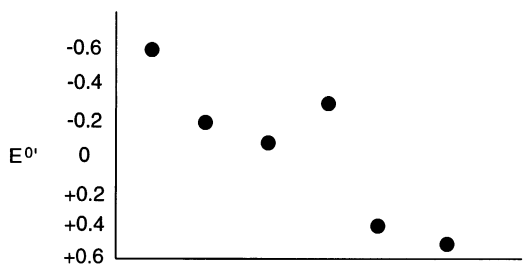
(a). Isocitrate ⟶ CoQ
(b). CoQ ⟶ cytochrome *c*
(c). Cytochrome *c* ⟶ O_2

12.15.* Succinate dehydrogenase and cytochrome oxidase are incubated, in the presence of oxygen, with cytochrome *c*, succinate, and coenzyme Q. What overall oxidation–reduction reaction would you expect to take place under these conditions?

12.16.* Calculate the *net* yield of ATP molecules per molecule of maltose for the complete oxidation of maltose to CO_2 and H_2O under aerobic conditions, assuming operation of: (a) the glycerol phosphate shuttle; (b) the malate–aspartate shuttle. Hint:

$$\text{Maltose} + H_2O \xrightarrow{\text{maltase}} \text{D-glucose} + \text{D-glucose}$$

12.17. Electron carriers of a prokaryotic ETS have $E^{\circ\prime}$ values as shown. What is the maximum number of ATP molecules that could be synthesized, per molecule of metabolite, when ATP synthesis is coupled to the operation of this ETS under biochemical standard conditions? What is the likely number of ATP molecules actually obtained? The $\Delta G^{\circ\prime}$ for the hydrolysis of ATP is -30.5 kJ mol^{-1}.



12.18. If the system described in the previous problem is known to operate at an efficiency of energy conservation of 30%, how many ATP molecules would be synthesized per molecule of metabolite?

12.19. Which of the following oxidation–reduction reactions will proceed as written under biochemical standard reaction conditions? (See Table 12.1.) Hb = hemoglobin.

(a) Ferredoxin (Fe^{2+}) + Hb (Fe^{3+}) → ferredoxin (Fe^{3+}) + Hb (Fe^{2+})
(b) Myoglobin (Fe^{2+}) + NAD^+ → myoglobin (Fe^{3+}) + NADH + H^+
(c) NAD^+ + pyruvate$^-$ → NADH + H^+ + lactate$^-$

12.20. The $E^{\circ\prime}$ of covalently bound FAD is +0.05 V. The $E^{\circ\prime}$ values of NAD^+ and succinate are listed in Table 12.1. On this basis, explain why the coenzyme of succinate dehydrogenase is FAD rather than NAD^+.

12.21.* The compound 2,4-dinitrophenol was at one time prescribed as a weight-reducing drug. People soon realized that this constitutes an *extremely dangerous* method of weight control, and the compound was no longer used after some deaths occurred. (a) Why was this compound chosen as a weight-reducing drug? (b) Why did some deaths occur in using it? (c) Ingestion of 2,4-dinitrophenol also led to an increase in body temperature and to profuse sweating. What is the explanation for these side effects?

12.22.* Adding dicyclohexylcarbodiimide (DCCD) to actively respiring mitochondria leads to a decrease in both the rate of the ETS (oxygen consumption) and the rate of ATP production (ATP synthesis). Adding 2,4-dinitrophenol restores the oxygen consumption to its normal level, but ATP synthesis remains inhibited. How can the effect of DCCD be explained? Why does addition of 2,4-dinitrophenol lead to the observed results?

12.23.* Oxidative phosphorylation of a mitochondrial sample led to a $1.2 \times 10^{-4} M$ increase in the concentration of ATP. It also led to a decrease in the concentration of NADH; this amounted to a decrease of 0.25 in the absorbance at 340 nm (light path = 1.0 cm). Given that the extinction coefficient (ϵ) of NADH is $6.2 \times 10^3\ M^{-1}\ cm^{-1}$, calculate the number of ATP molecules synthesized per molecule of NADH oxidized. Note that $A = \epsilon c l$, where A is absorbance, ϵ is the extinction coefficient, c is the concentration, and l is the length of the light path (see Appendix C).

12.24. The compounds 2,6-dichlorophenol indophenol and tetramethyl-*p*-phenylenediamine (TMPD) are two artificial electron acceptors used in studies of the ETS. Their $E^{\circ\prime}$ values are +0.22 and +0.26, respectively. Predict the points in the ETS where they are likely to act under biochemical standard conditions.

12.25. What is the expected P/O ratio when the ETS is: (a) inhibited at cytochrome *c*; (b) inhibited at coenzyme Q; (c) provided with ascorbic acid and tetramethyl-*p*-phenylenediamine (TMPD)?

12.26.* A male adult requires a minimum of about 7500 kJ per day (Table 8.6). Assuming that all of this energy derives from ATP hydrolysis ($\Delta G^{\circ\prime} = -30.5$ kJ mol^{-1}), calculate the number of grams of ATP (MW = 507) that must be hydrolyzed per day. Given that the actual amount of ATP in the individual is about 50 g, compute the approximate number of times that all of the body's ATP must be hydrolyzed and resynthesized.

12.27.* Calculate the *net* yield of ATP molecules per molecule of metabolite (underlined) for the following cases (refer to Chapters 10 and 11). The reactions occur under aerobic conditions and involve participation of the glycerol phosphate shuttle.

(a) Oxidation of glyceraldehyde 3-phosphate to acetyl CoA
(b) Complete oxidation of fructose 6-phosphate to CO_2 and H_2O
(c) Complete oxidation of 3-phosphoglycerate to CO_2 and H_2O.

12.28.* What would be the theoretical energy yield, in terms of ATP molecules per molecule of glucose, if glucose is phosphorylated by hexokinase, the glucose 6-phosphate is catabolized via the pentose phosphate pathway, and the NADPH of the pentose phosphate pathway is oxidized via the ETS, yielding the same number of molecules of ATP as are produced by NADH? (a) Assume that NADPH is formed inside the mitochondria; (b) assume that NADPH is formed in the cytosol and the glycerol phosphate shuttle applies.

Lipid Metabolism

13

Lipid metabolism, like that of carbohydrates, can be divided into five broad areas—*digestion, transport, storage, degradation,* and *biosynthesis*. Since **fats** (acylglycerols) constitute most of an organism's lipids and are the major dietary form of lipids, this chapter will focus primarily on the metabolism of fats.

Fat digestion occurs in the aqueous environment of the intestine and requires the action of water-soluble **lipases,** enzymes that catalyze the hydrolysis of fats. Digestion also requires solubilization of the nonpolar fats. **Bile salts,** present at high concentrations in the bile (which empties into the intestine), solubilize fats by emulsifying them. Following their digestion (Section 8.4), most fats enter the lymphatic circulation as **chylomicrons** that subsequently enter the bloodstream. Short-chain fatty acids are absorbed directly from the intestine into the circulatory system.

Transport of fats by the blood and intracellular fluids also requires their solubilization. Formation of *lipoproteins* provides water-soluble macromolecular aggregates owing to the presence of proteins and amphipathic lipids. Formation of fatty acid/serum albumin complexes solubilizes free fatty acids. **Serum albumin** is a monomeric protein (MW = 66,000) and the most abundant of all plasma proteins. It occurs at a concentration of about 4 g/100 ml of plasma and constitutes about 50% of total plasma protein. Serum albumin can bind up to 10 fatty acid molecules per molecule of protein.

In addition to being incorporated into lipoproteins and albumin complexes, lipids are transported as **ketone bodies,** products of fatty acid catabolism that serve as an energy source under some conditions. Fatty acid metabolism comprises two major pathways: their stepwise degradation by means of *β-oxidation* and their stepwise synthesis by means of *fatty acid biosynthesis.*

13.1. STORAGE OF FATS

13.1.1. Depot Fat

Lipids, particularly fats, constitute the main storage form of energy in animals. Because lipids have the highest caloric value of all nutrients (37.7 kJ/g or 9.0 kcal/g), using them for storage provides a decided advantage. For animals to store an energetically equivalent amount of either carbohydrate or protein would necessitate deposition of considerably larger quantities of substance. Also, carbohydrate and protein stores would be even larger because the polar character of these substances would result in inclusion of water in the deposits.

We refer to stored fat as **depot fat** or **adipose tissue,** and dietary lipids must be converted to it prior to storage. The composition of stored fat is always characteristic of the organism, regardless of the source of dietary lipids. We can distinguish human, sheep, pig, and other animal fats by the relative proportions and compositions of their *mono-, di-,* and *triacylglycerols.* To produce an organism's unique mixture of acylglycerols from dietary fats requires a variety of chemical reactions, including changes in fatty acid chain length and saturation. Transformations of dietary fats, termed "working over," occur predominantly in the liver.

Storage and transport of fats are interrelated processes (Figure 13.1). When fat has to be catabolized to produce usable energy in the form of ATP, fatty acids are removed from adipose tissue and transported to the liver for degradation. We refer to the release of fatty acids from depot fat as **mobilization.** The process requires the action of enzymes. *Lipases* and *phospholipases* catalyze hydrolysis of ester bonds in acylglycerols and glycerophospholipids, respectively (Figure 13.2). Released fatty acids undergo catabolism in the mitochondria, and glycerol is subject to degradation in the cytosol.

A hormonally controlled enzyme cascade (Figure 13.3), similar to the cascades in regulation of carbohydrate metabolism (see Figures 10.26 and 10.30), regulates fatty acid mobilization. Once free fatty acids have been released inside *adipocytes* (cells of adipose tissue), they diffuse across cell membranes, become bound to serum albumin, and are transported to the tissues.

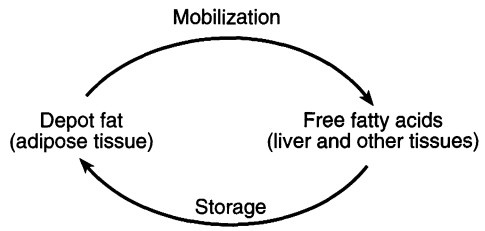


Figure 13.1. The link between mobilization and storage of fatty acids.

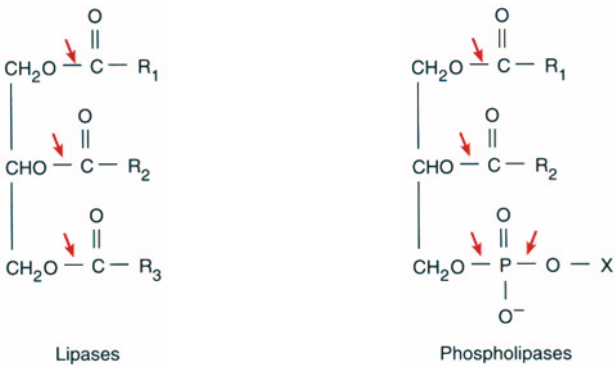


Figure 13.2. Points of attack of lipases and phospholipases.

13.1.2. Fatty Liver

Excessive mobilization of fatty acids may lead to development of **fatty liver,** a liver that has been infiltrated by fat cells so that portions have become nonfunctional fatty tissue. Fatty livers may occur in diabetics suffering from insulin deficiency. Diabetics cannot metabolize glucose properly and must use other nutrients as sources of energy. Typically, this results in an overreliance on fat metabolism, including excessive fatty acid mobilization and excessive fatty acid metabolism in the liver.

Fatty liver may also result from exposure to chemicals such as carbon tetrachloride and pyridine. These compounds destroy liver cells and lead to their replacement by fatty tissue. Dietary deficiencies of *choline* and *methionine,* called **lipotropic agents,** may likewise result in formation of fatty liver because of an effect on lipid transport.

Recall that choline occurs as a component of *phosphatidyl choline,* one of the *phospholipids.* Its synthesis involves the carbon skeleton of serine and requires a transfer of three methyl groups from *S*-adenosylmethionine (Figure 13.4). ***S*-Adenosylmethionine (SAM),** which forms by reaction of methionine and ATP, serves as a donor of methyl groups for many biological methylations. We can describe phosphatidyl choline biosynthesis by a schematic sequence, which indicates the metabolic fate of the methyl group:

Methionine → *S*-adenosylmethionine → choline → phosphatidyl choline

Inadequate supplies of methionine and choline in the diet lead to insufficient synthesis of phosphatidyl choline in particular and of phospholipids in general. De-

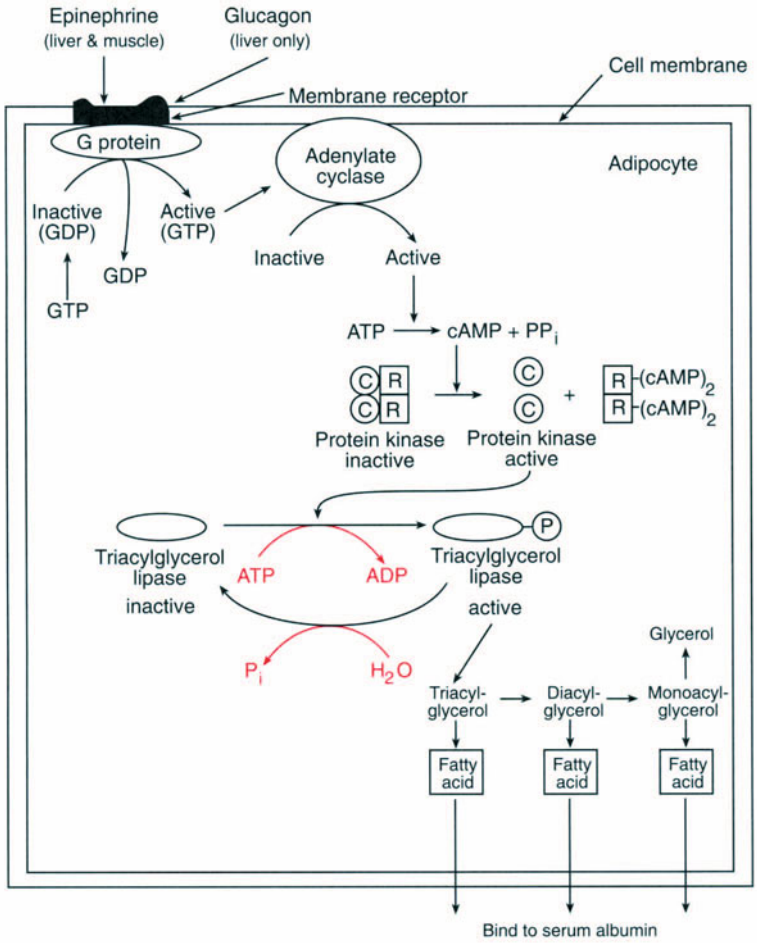


Figure 13.3. The hormonally controlled enzyme cascade regulating fatty acid mobilization. Triacylglycerol lipase catalyzes the hydrolysis of fatty acids linked to C(1) and C(3) of glycerol, yielding mono- and diacylglycerols that are then acted upon by mono- and diacylglycerol lipases.

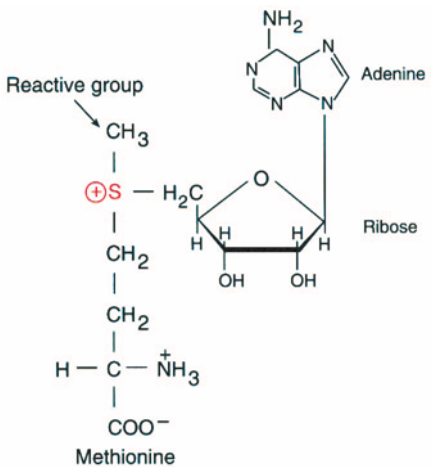


Figure 13.4. *S*-Adenosylmethionine (SAM).

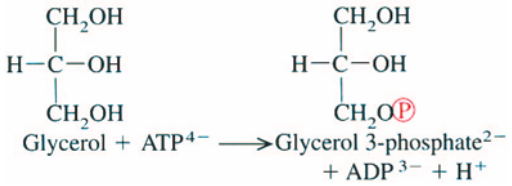
creased phospholipid synthesis results in a deficiency of lipoproteins. Because lipoproteins constitute a major form in which lipids are transported from the liver, decreased lipoprotein concentrations may lead to excessive lipid accumulation in the liver, ultimately producing fatty liver.

13.2. FAT CATABOLISM

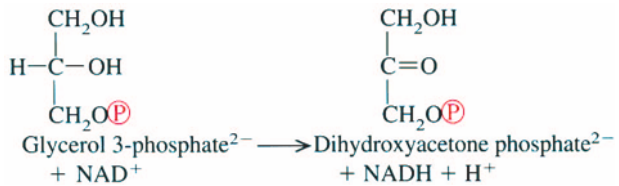
13.2.1. Degradation of Glycerol

Fat catabolism begins with the hydrolysis of fats to glycerol and fatty acids. These two structural components are metabolized separately. Glycerol catabolism occurs in liver cytosol and involves conversion to a glycolytic inter-

mediate. First, *glycerokinase* catalyzes a phosphorylation of glycerol to glycerol 3-phosphate:



Glycerol 3-phosphate dehydrogenase then catalyzes the oxidation of glycerol 3-phosphate to dihydroxyacetone phosphate. This reaction constitutes part of the *glycerol phosphate shuttle.*



Dihydroxyacetone phosphate enters glycolysis, where it is converted to glyceraldehyde 3-phosphate, an intermediate of both the glycolytic and gluconeogenic pathways. Glycolysis leads to pyruvate and amino acids, and gluconeogenesis leads to glucose. Thus, glycerol, a component of *fats,* can be converted to both *proteins* and *carbohydrates.* Such interconversions are characteristic of the interrelationships in metabolism and illustrate the principle that one type of nutrient may readily be converted to several others.

Glycerol catabolism can provide energy. NADH, produced in the glyeraldehyde 3-phosphate dehydrogenase reaction of glycolysis, can be oxidized to NAD^+ by the electron transport system via specific shuttles that link the compartments in which these pathways occur.

13.2.2. Knoop's Hypothesis

A clue to the mechanism of fatty acid catabolism came from some ingenious experiments carried out by F. Knoop at the turn of the century (1904). In those days, radioactive isotopes were not yet in use, radioactivity having only recently been discovered, and labeled compounds were not available. Yet Knoop designed the first "tracer" experiment in biochemistry. He did this by tagging the methyl carbon of both even- and odd-numbered fatty acids with a phenyl group. He then fed these "labeled" fatty acids to dogs and analyzed the urine excreted. Under these conditions, the phenyl group is not metabolized and is excreted in the form of specific organic compounds.

Knoop found that dogs fed even-numbered fatty acid derivatives excreted the phenyl group in the form of *phenylaceturic acid;* those fed odd-numbered fatty acid derivatives excreted the phenyl group in the form of *hippuric acid* (Figure 13.5). Phenylaceturic acid and hippuric acid were known to be formed, respectively, from *phenylacetic acid* and *benzoic acid* by coupling each with a molecule of *glycine.* Based on these experiments, Knoop proposed that fatty acids are degraded in a stepwise fashion, by successive cleavage of two-carbon fragments (in the form of acetate), beginning with the carboxyl end of the molecule.

Coupling of phenylacetic acid and benzoic acid to glycine and excretion of the products illustrates an organ-

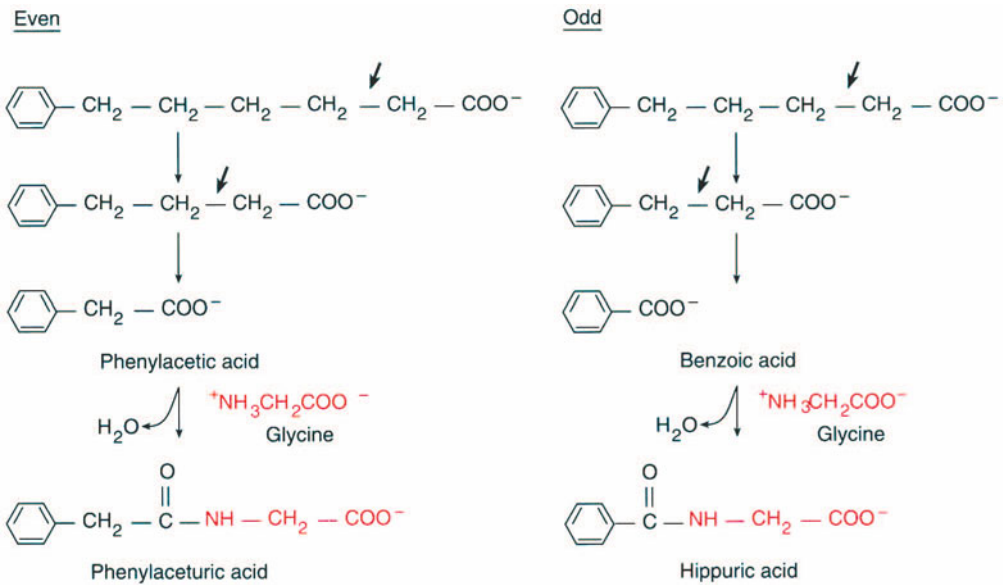


Figure 13.5. Illustration of Knoop's experiments on the oxidation of phenyl-labeled fatty acids. Arrows indicate the presumed points of cleavage.

ism's capacity to divest itself of foreign toxic compounds by *detoxification.* In this process, enzymatic reactions convert such compounds to less harmful substances or couple them to naturally occurring metabolites. The products of these reactions are then excreted. For example, use of benzoic acid as a food preservative poses little danger because benzoic acid is readily detoxified to hippuric acid, followed by excretion.

13.3. BETA OXIDATION OF FATTY ACIDS

Subsequent work corroborated Knoop's brilliant deductions. Fatty acid degradation, called **β-oxidation,** does occur by successive cleavage of two-carbon fragments from the carboxyl end. However, the modern view differs from Knoop's hypothesis in three ways:

1. The two-carbon fragment is removed as *acetyl CoA,* not as acetate.
2. All intermediates in the reaction sequence are bound to coenzyme A.
3. Initiation of the degradation requires ATP hydrolysis.

β-Oxidation occurs principally in the mitochondria and consists of five enzymatic steps:

1. Activation: Thiokinase (acyl CoA synthase)
2. Dehydrogenation: Acyl CoA dehydrogenase
3. Hydration: Enoyl CoA hydratase
4. Dehydrogenation: L-3-Hydroxyacyl CoA dehydrogenase
5. Cleavage: Thiolase (β-ketothiolase)

Fatty acid activation takes place in the endoplasmic reticulum or the outer mitochondrial membrane. The remaining four reactions occur in the mitochondrial matrix and are linked to fatty acid activation by means of a specific carrier system. Figure 13.6 shows the entire reaction sequence.

13.3.1. Individual Reactions

13.3.1A. Thiokinase. The first step in the β-oxidation of fatty acids involves a conversion of the fatty acid to a chemically more reactive form. The reaction, called **fatty acid activation,** is catalyzed by **thiokinase** or **acyl CoA synthase.** Three types of thiokinases occur, specific for short-, medium-, and long-chain fatty acids, re-

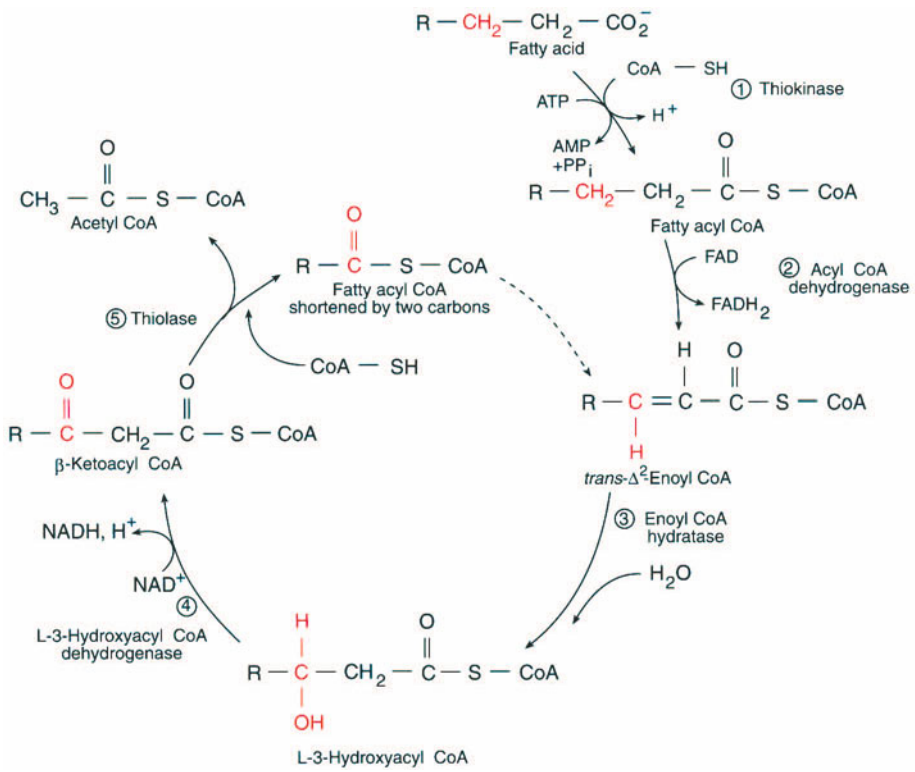


Figure 13.6. The β-oxidation cycle of saturated fatty acids. One turn of the cycle produces a fatty acyl CoA, shortened by two carbons from the one that entered the cycle. The two carbons are split off as acetyl CoA.

spectively. Thiokinases act on both saturated and unsaturated fatty acids. Fatty acid activation requires cleavage of an energy-rich bond in ATP. We can write the coupled reactions as follows:

$$\underset{}{\overset{(2)}{\text{Fatty acid}^- + \text{ATP}^{4-} + \text{H}^+}} \rightleftharpoons \underset{}{\overset{(1)}{\text{fatty acyl-AMP}^-} + \overset{(1)}{\text{PP}_i^{3-}}}$$

$$\underset{}{\overset{(1)}{\text{Fatty acyl-AMP}^-} + \text{HS-CoA}^{4-}} \rightleftharpoons \underset{}{\overset{(1)}{\text{fatty acyl CoA}^{4-}} + \text{AMP}^{2-} + \text{H}^+}$$

Overall reaction:

$$\underset{}{\overset{(2)}{\text{Fatty acid}^- + \text{ATP}^{4-}}\ \text{HS-CoA}^{4-} \rightleftharpoons} \underset{}{\overset{(1)}{\text{Fatty acyl CoA}^{4-}} + \overset{(1)}{\text{AMP}^{2-}} + \text{PP}_i^{3-}} \qquad \Delta G^{\circ\prime} = -0.8 \text{ kJ mol}^{-1}$$

The common intermediate, **fatty acyl adenylate (fatty acyl-AMP),** consists of a fatty acid molecule attached to AMP via a mixed acid anhydride linkage (Figure 13.7). Because of the anhydride structure, the fatty acid–AMP link constitutes an energy-rich bond. Accordingly, fatty acyl adenylates are *energy-rich compounds.* Hydrolysis of the fatty acid–AMP bond results in a reaction with a highly negative biochemical standard free energy change ($\Delta G^{\circ\prime}$).

In the above three equations, we have indicated the number of energy-rich bonds per compound in parentheses. Note that each reaction has an equal number of energy-rich bonds for the reactants and the products. Hence all three reactions are readily reversible, with biochemical standard free energy changes close to zero. The overall reaction is driven toward formation of fatty acyl CoA because of a fourth, related reaction involving *pyrophosphatase.* This enzyme catalyzes the strongly exergonic hydrolysis of pyrophosphate (PP_i) to inorganic phosphate (P_i):

$$\text{PP}_i^{3-} + \text{H}_2\text{O} \rightarrow 2\text{P}_i^{2-} + \text{H}^+ \qquad \Delta G^{\circ\prime} = -33.1 \text{ kJ mol}^{-1}$$

By coupling PP_i hydrolysis to fatty acid activation, the overall reaction proceeds spontaneously with a free energy change of $\Delta G^{\circ\prime} = -33.9$ kJ mol^{-1}. Using stearic acid (18 carbons, saturated) to illustrate β-oxidation, we have for the fatty acid activation step:

$$\underset{\text{Fatty acid}}{\overset{\beta\quad\alpha}{\text{CH}_3\text{—(CH}_2)_{14}\text{—CH}_2\text{—CH}_2\text{—COO}^-}} + \text{ATP}^{4-} + \text{HS—CoA}^{4-} \rightleftharpoons$$

$$\underset{\text{Fatty acyl CoA}}{\overset{\beta\quad\alpha}{\text{CH}_3\text{—(CH}_2)_{14}\text{—CH}_2\text{—CH}_2\text{—}\overset{\overset{\text{O}}{\|}}{\text{C}}\text{—S—CoA}^{4-}}} + \text{AMP}^{2-} + \text{PP}_i^{3-}$$

13.3.1B. Carnitine Carrier System. Because fatty acid activation occurs in the endoplasmic retic-

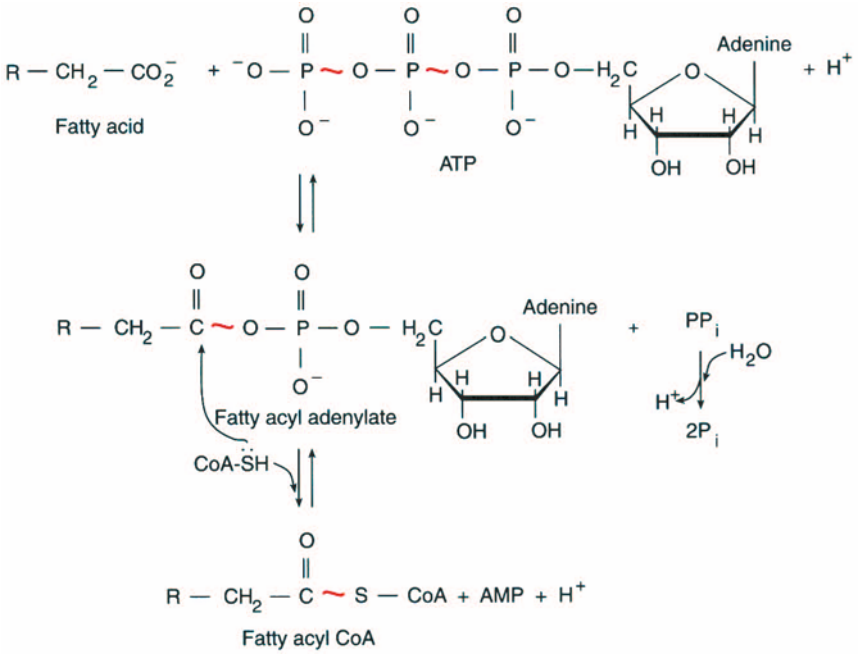


Figure 13.7. Fatty acid activation. Pyrophosphatase-catalyzed hydrolysis of PP_i drives the reaction to completion.

ulum or the outer mitochondrial membrane, the reaction product must be transported from there *into* the mitochondrial matrix, where subsequent steps of β-oxidation take place. Because fatty acyl CoA cannot cross the inner mitochondrial membrane directly, its movement requires a transport system. The specific carrier of fatty acyl CoA is **carnitine,** a low-molecular-weight compound derived from lysine (Figure 13.8).

The transport mechanism, diagrammed in Figure 13.9, consists of four steps:

1. Cytoplasmic activation of a fatty acid to a fatty acyl CoA, as described above.
2. Passage of fatty acyl CoA through the outer mitochondrial membrane and into the intermembrane space. There the fatty acyl group is transferred to carnitine, forming *acyl carnitine,* and coenzyme A is released back to its extramitochondrial pool.
3. Movement of acyl carnitine from the intermembrane space, across the inner membrane, into the matrix. In the matrix, the fatty acid is transferred to a molecule of coenzyme A from the mitochondrial pool.
4. Free carnitine exits the matrix through the inner mitochondrial membrane.

Transfer of the fatty acid from acyl carnitine to coenzyme A proceeds without additional ATP expenditure because the acyl carnitine bond is sufficiently energy-rich.

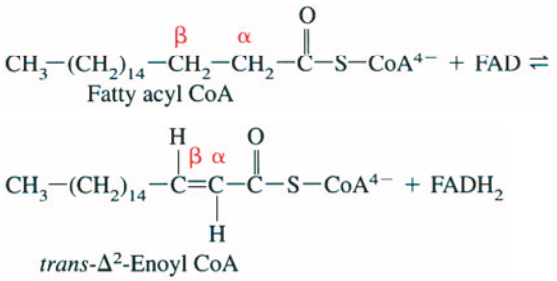
Operation of this transport mechanism means that cells maintain two separate pools of coenzyme A, in the cytoplasm and inside the mitochondria, that participate in different metabolic processes. Mitochondrial coenzyme A serves in catabolism of fatty acids (β-oxidation), while cytoplasmic coenzyme A functions in fatty acid biosynthesis. Cells maintain similar separate cytoplasmic and mitochondrial pools of ATP and NAD^+.

The **carnitine carrier system** provides a control point of β-oxidation. Malonyl CoA is an allosteric inhibitor of carnitine acyltransferase I and prevents transfer of fatty acyl CoA into mitochondria. Because of malonyl CoA's role as an intermediate in fatty acid biosynthesis, it exerts a dual effect: it stimulates fatty acid synthesis and inhibits fatty acid oxidation.

Other than the effect of malonyl CoA, no specific enzyme control point has yet been identified in β-oxidation. The only other known control of fatty acid catabolism consists of the hormonal effects on fatty acid mobilization (Figure 13.3).

13.3.1C. Acyl CoA Dehydrogenase.

The second step of β-oxidation involves an oxidation (*dehydrogenation*) of fatty acyl CoA, catalyzed by **acyl CoA dehydrogenase.** Three types of acyl CoA dehydrogenases occur, specific for short-, medium-, and long-chain fatty acids, respectively. All three enzyme types are flavoproteins and carry tightly bound FAD. The product formed is an α,β-unsaturated fatty acyl CoA. Acyl CoA dehydrogenase produces only the *trans* isomer, designated *trans-Δ^2-enoyl CoA:*



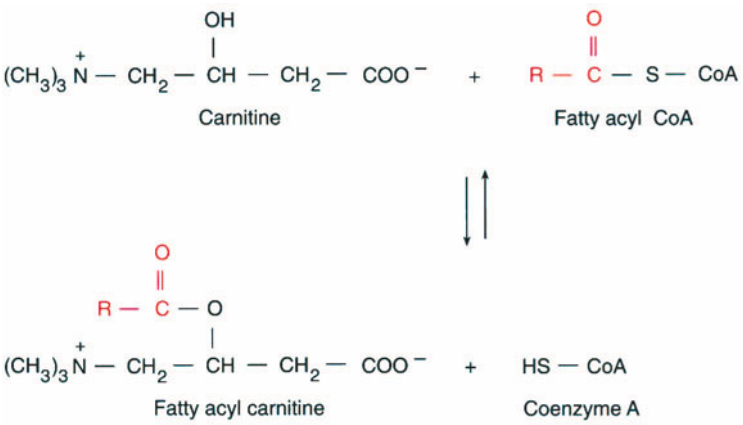


Figure 13.8. Carnitine and its reversible reaction with fatty acyl CoA.

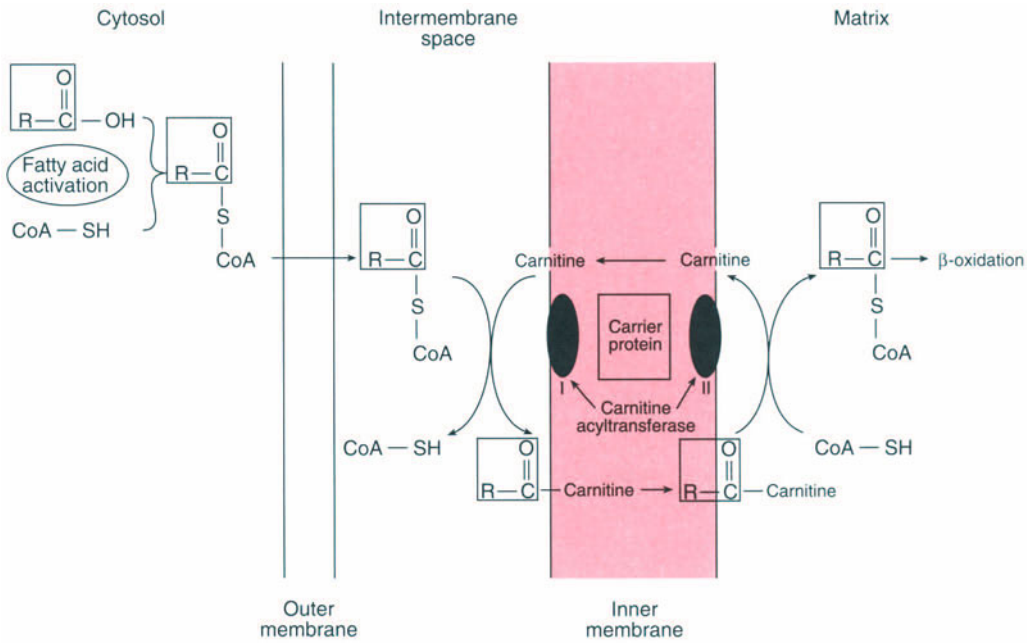
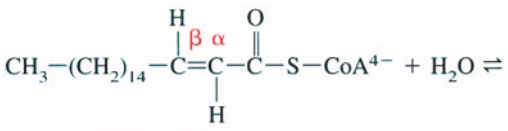
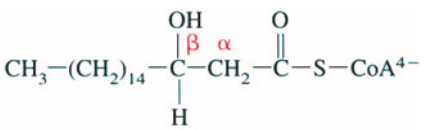


Figure 13.9. The carnitine carrier system. Translocation is mediated by a carrier protein. Carnitine acyltransferases I and II are located, respectively, on the external and internal surfaces of the inner mitochondrial membrane.

13.3.1D. Enoyl CoA Hydratase. The third reaction of β-oxidation consists of a *hydration,* catalyzed by **enoyl CoA hydratase.** The hydratase requires the *trans* isomer produced in the previous step and catalyzes addition of water across the α,β-double bond. Hydratase action results in formation of a β-alcohol fatty acyl CoA. Enoyl CoA hydratase is stereospecific so that only the L-isomer, *L-3-hydroxyacyl CoA,* is produced:

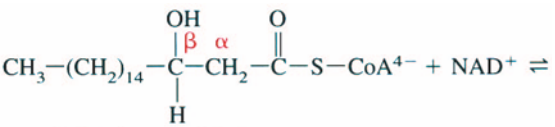


trans-Δ^2-Enoyl CoA

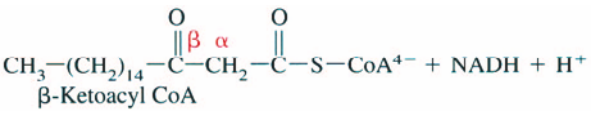


L-3-Hydroxyacyl CoA

13.3.1E. L-3-Hydroxyacyl CoA Dehydrogenase. A second oxidation, or *dehydrogenation,* follows the hydration step. This reaction is catalyzed by a pyridine-linked dehydrogenase, **L-3-hydroxyacyl CoA dehydrogenase,** that uses NAD^+ as coenzyme. The enzyme has an absolute specificity for L-3-hydroxyacyl CoA and forms *β-ketoacyl CoA:*

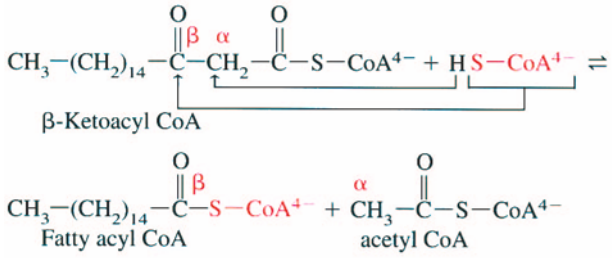


L-3-Hydroxyacyl CoA



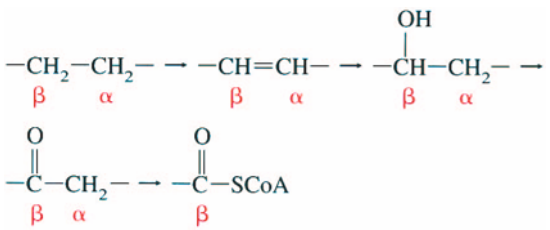
β-Ketoacyl CoA

13.3.1F. Thiolase. The last reaction of β-oxidation consists of a *cleavage* of the fatty acid molecule at its β-carbon. We call the reaction **thiolysis** by analogy with other cleavage reactions involving water (hydrolysis) or phosphoric acid (phosphorolysis). Thiolytic cleavage is catalyzed by **thiolase** (or **β-ketothiolase**) and consists of a nucleophilic attack by the thiol sulfur of coenzyme A on the electron-deficient carbon of the keto group to form a shortened fatty acyl CoA and a molecule of acetyl CoA. The fatty acid residue in the fatty acyl CoA represents the original fatty acid *shortened by two carbons;* the two carbons removed form the acetyl group of acetyl CoA:



Following thiolysis, the shortened fatty acyl CoA recycles through the last four steps of β-oxidation—from acyl CoA dehydrogenase through thiolase. The second thiolytic cleavage yields another fatty acyl CoA, in which the original fatty acid has now been shortened by four carbons, as well as another molecule of acetyl CoA. These steps are repeated, each time shortening the original fatty acid by two carbons, *removed as acetyl CoA*. The ATP-requiring activation step is needed only once to form the coenzyme A derivative of the original fatty acid. Because each turn of the cycle yields a fatty acyl CoA molecule, each turn produces an already activated shortened fatty acid that can immediately reenter the cyclic reaction sequence.

We refer to the degradative pathway as β-oxidation (also called *β-oxidation cycle* or *spiral*) because of the progressive oxidation that occurs at the β-carbon of the fatty acid. The β-carbon changes from the reduced carbon of a methylene group (CH_2) to the oxidized carbon of a keto group (CO):



The overall reaction for the complete degradation of stearic acid (18 carbons) is

$$\text{Stearic acid}^- + 8\text{FAD} + 8\text{NAD}^+ + 8\text{H}_2\text{O} + 9\text{CoA}{-}\text{SH}^{4-} + \text{ATP}^{4-} \rightarrow 9\text{Acetyl CoA}^{4-} + 8\text{FADH}_2 + 8\text{NADH} + 8\text{H}^+ + \text{AMP}^{2-} + \text{PP}_i^{3-}$$

Once acetyl CoA forms by β-oxidation, it can enter the citric acid cycle. Catabolism of acetyl CoA via the citric acid cycle/electron transport system results in complete oxidation of the acetyl group to CO_2 and H_2O. Hence, a combination of β-oxidation and the citric acid cycle/ETS leads to *complete oxidation of fatty acids to carbon dioxide and water.*

13.3.2. Energetics

β-Oxidation of fatty acids leads to the production of large amounts of energy in the form of ATP. The energy derives from $FADH_2$ and NADH, produced during β-oxidation and subsequently oxidized via the ETS. Recall that β-oxidation takes place in the mitochondrial matrix and that the ETS is located in the inner mitochondrial membrane. Hence, $FADH_2$ and NADH formed during β-oxidation can be reoxidized directly via the ETS; no shuttle is required. Oxidation of NADH yields 3 ATP (P/O = 3). $FADH_2$ is oxidized through the intermediacy of an *electron-transfer flavoprotein (ETF),* an inner membrane flavoprotein. From ETF, electrons flow through an iron–sulfur protein (Fe-S) to coenzyme Q, resulting in the following electron carrier sequence:

$$FADH_2 \rightarrow \text{ETF} \rightarrow \text{Fe-S} \rightarrow \text{CoQ}$$

Because $FADH_2$ enters the ETS at the level of CoQ, oxidation of $FADH_2$ yields only 2 ATP (P/O = 2), rather than the 3 ATP obtained with NADH (P/O = 3).

Consider the β-oxidation of palmitic acid, a 16-carbon saturated fatty acid. Palmitic acid is broken down completely to acetyl CoA. In that process, it passes through the β-oxidation cycle seven times, yielding eight molecules of acetyl CoA. Since the catabolism of one acetyl CoA via the citric acid cycle/ETS yields 12 ATP, we have the following:

8 Acetyl CoA (citric acid cycle)	8×12	= 96 ATP
7 $FADH_2$ (ETS)	7×2	= 14 ATP
7 ($NADH, H^+$) (ETS)	7×3	= 21 ATP
		131 ATP

From this value we must subtract the 2 ATP equivalents expended in β-oxidation. One ATP equivalent is used directly for formation of fatty acyl CoA (ATP → AMP + PP_i). The other ATP equivalent is used indirectly, since driving fatty acid activation to completion requires hydrolysis of the energy-rich bond in PP_i ($PP_i \rightarrow 2P_i$). Thus, in terms of *net yield of energy-rich bonds,* β-oxidation of palmitic acid produces 129 such bonds.

You can see that β-oxidation of one fatty acid molecule produces a very large number of ATP molecules. This explains why fats represent such a good source of energy. Complete oxidation of fatty acids to CO_2 and H_2O yields greater amounts of energy than the corresponding oxidation of carbohydrates because fatty acids are more highly reduced than carbohydrates. In fatty acids, almost all of the carbons are methylene carbons (CH_2) whereas most of the carbons in carbohydrates carry an OH group (CHOH). Thus, stearic acid

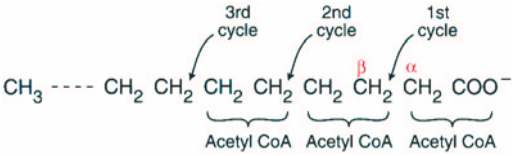


Figure 13.10. Stepwise shortening of fatty acids during β-oxidation.

($C_{18}H_{36}O_2$) has a C:H:O ratio of 1:2:0.11 while the corresponding ratio for glucose ($C_6H_{12}O_6$) is 1:2:1. Glucose has nine times more oxygen per carbon than stearic acid. The more reduced the state of a compound, the greater is the number of hydride ions and protons abstracted from it in the process of complete oxidation. The larger number of hydride ions and protons, when channeled into the ETS, generate a larger amount of energy in the form of ATP.

Because even a small amount of fat yields considerable amounts of energy, fats constitute ideal substances for energy storage. By the same token, elimination of even a small amount of fat (weight loss) requires a great deal of exercise (energy expenditure). A change of dietary habits will, of course, also help in this regard.

Using the above yield of ATP, we can calculate the efficiency of β-oxidation. Biochemical standard free energy changes ($\Delta G^{\circ\prime}$) for complete oxidation of palmitic acid to CO_2 and H_2O and for hydrolysis of ATP are -9791 kJ mol^{-1} and -30.5 kJ mol^{-1}, respectively. On this basis, the efficiency equals

$$\frac{(129 \times 30.5 \text{ kJ mol}^{-1})}{9791 \text{ kJ mol}^{-1}} \times 100 = 40\%$$

a value similar to the efficiencies we calculated for other metabolic pathways. Remember, however, that a calculation of the actual intracellular efficiency must be based on $\Delta G'$ values.

13.3.3. Even- and Odd-Numbered Fatty Acids

Degradation by means of β-oxidation applies to both even- and odd-numbered fatty acids (Figure 13.10). Even-numbered fatty acids are degraded *completely to acetyl CoA*. An even-numbered fatty acid having *n* carbons yields *n*/2 molecules of acetyl CoA but passes through the cycle only (*n*/2) − 1 times. For example, stearic acid (18 carbons) yields nine acetyl CoA but passes through the cycle only eight times.

Odd-numbered fatty acids also yield large quantities of acetyl CoA when subjected to β-oxidation. However, the final thiolytic cleavage produces acetyl CoA and *propionyl CoA* (CH_3-CH_2-CO-S-CoA). An odd-numbered fatty acid having n carbons yields (n − 3)/2 acetyl CoA plus propionyl CoA, and passes through the cycle (n − 3)/2 times. Propionyl CoA undergoes a three-step conversion to succinyl CoA (Figure 13.11), an intermediate of the citric acid cycle. The first enzyme in this reaction sequence requires biotin as coenzyme, and the third enzyme requires 5′-deoxyadenosylcobalamin, a derivative of vitamin B_{12} (cyanocobalamin), as coenzyme.

13.3.4. Unsaturated Fatty Acids

β-Oxidation of unsaturated fatty acids requires the action of as many as three auxiliary enzymes, depending on the fatty acid and the organism (Figure 13.12). Two problems must be addressed.

First, most naturally occurring unsaturated fatty acids contain only *cis* double bonds. In order to serve as substrate for enoyl CoA hydratase, such *cis* double bonds must be converted to the *trans* configuration, a reaction catalyzed by **enoyl CoA isomerase.**

Second, a double bond may have to be eliminated in order to convert the molecule to a better substrate for enoyl CoA hydratase. The enzyme **dienoyl CoA reductase** catalyzes reduction of a double to a single bond and

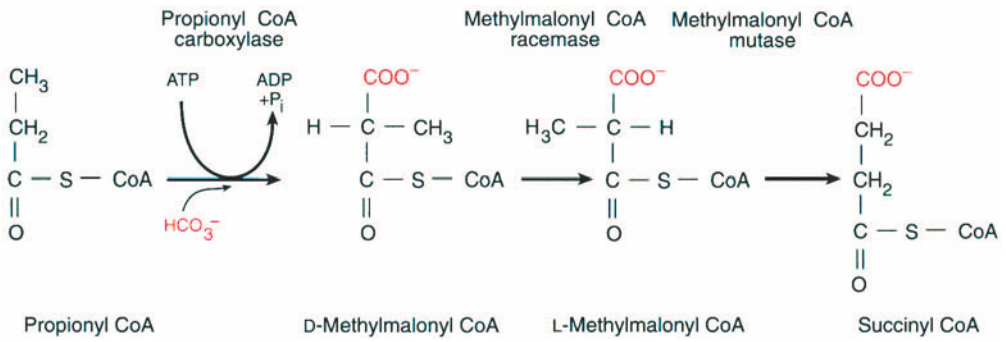


Figure 13.11. Conversion of propionyl CoA to succinyl CoA.

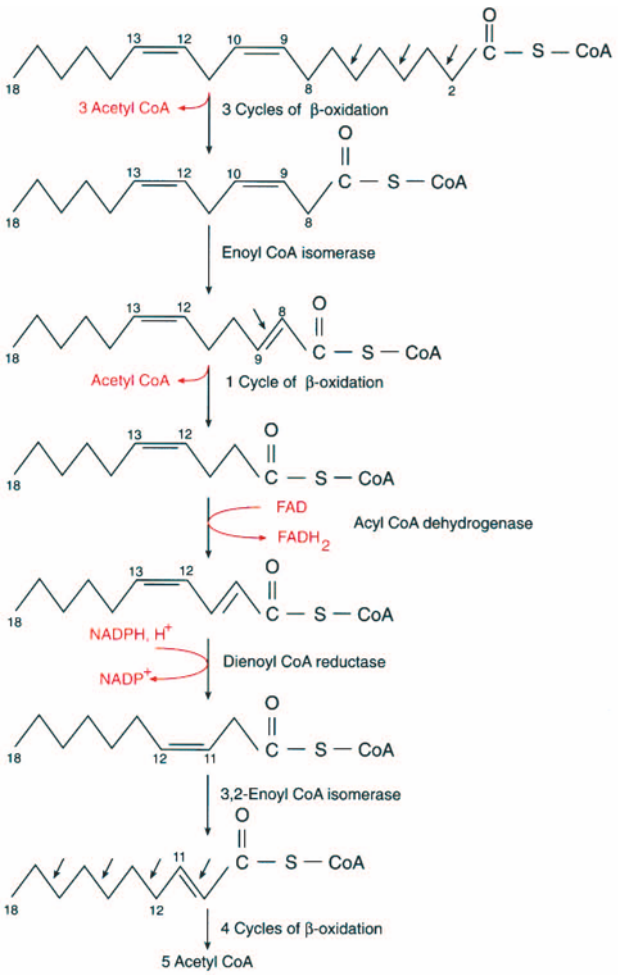


Figure 13.12. β-Oxidation of an unsaturated fatty acid (linoleic acid) in mammals. The original numbering of carbon atoms has been maintained throughout for clarity.

uses NADPH as coenzyme. A third auxiliary enzyme, **3,2-enoyl CoA isomerase,** is required in mammals but not in *Escherichia coli.*

13.4. KETONE BODIES

13.4.1. Metabolic Fates of Acetyl CoA

Next to ATP, acetyl CoA probably represents the most important low-molecular-weight biomolecule of metabolism. Its four major metabolic fates exemplify its pivotal role (Figure 13.13):

1. It represents the form in which carbohydrates, lipids, and some amino acids enter the citric acid cycle. Catabolism of these nutrients via the combined action of the citric acid cycle and the electron transport system results in their complete oxidation to CO_2 and H_2O.
2. It constitutes the source of all the carbons for the biosynthesis of cholesterol. Because cholesterol is the parent compound of steroids, acetyl CoA serves as the precursor of all steroids.
3. It functions as the precursor for the biosynthesis of fatty acids.
4. It serves as the precursor for the biosynthesis of *acetone, acetoacetate,* and *β-hydroxybutyrate,* three compounds known as *ketone bodies.*

13.4.2. Properties of Ketone Bodies

The term *ketone bodies* is somewhat misleading. Acetone, acetoacetate, and β-hydroxybutyrate constitute chemical compounds, not "bodies." Moreover, only two of the compounds are actually ketones; the third—β-hydroxybutyrate—is not a ketone. All three ketone bodies derive from acetyl CoA and form primarily in the liver.

Ketone bodies represent low-molecular-weight, water-soluble forms of lipid-based energy since they can serve as substrates for generating energy in place of glucose in muscle and brain tissue. When functioning in this capacity, ketone bodies undergo conversion to acetyl CoA, which enters the citric acid cycle. Some tissues, notably the heart, commonly derive large amounts of their energy requirements from ketone bodies. Other tissues, especially the brain, can also use ketone bodies as a source of energy under certain conditions.

Brain cells form tight junctions in capillaries that prevent passive entry into the brain of water-soluble substances. This selective filtration system separates the brain from the general circulation and is known as the **blood–brain barrier.** Water-soluble compounds pass this barrier only if they can be moved across by specific transport systems, and fatty acids cannot pass the blood–brain barrier. Glucose is the normal fuel of the brain. However, upon starvation, the brain can adapt to use ketone bodies as a source of energy. With prolonged starvation, acetoacetate can provide 75% of the energy needs of the brain.

You could call synthesis of ketone bodies an "overflow pathway." Such a pathway normally has minor significance but becomes important when concentrations of specific substances increase and metabolites "spill over" into the secondary pathway. In the case of ketone bodies, the pathway becomes accentuated *when the concentration of acetyl CoA builds up,* which occurs when its formation exceeds its consumption. As the overflow pathway begins to be used, concentrations of ketone bodies in the blood

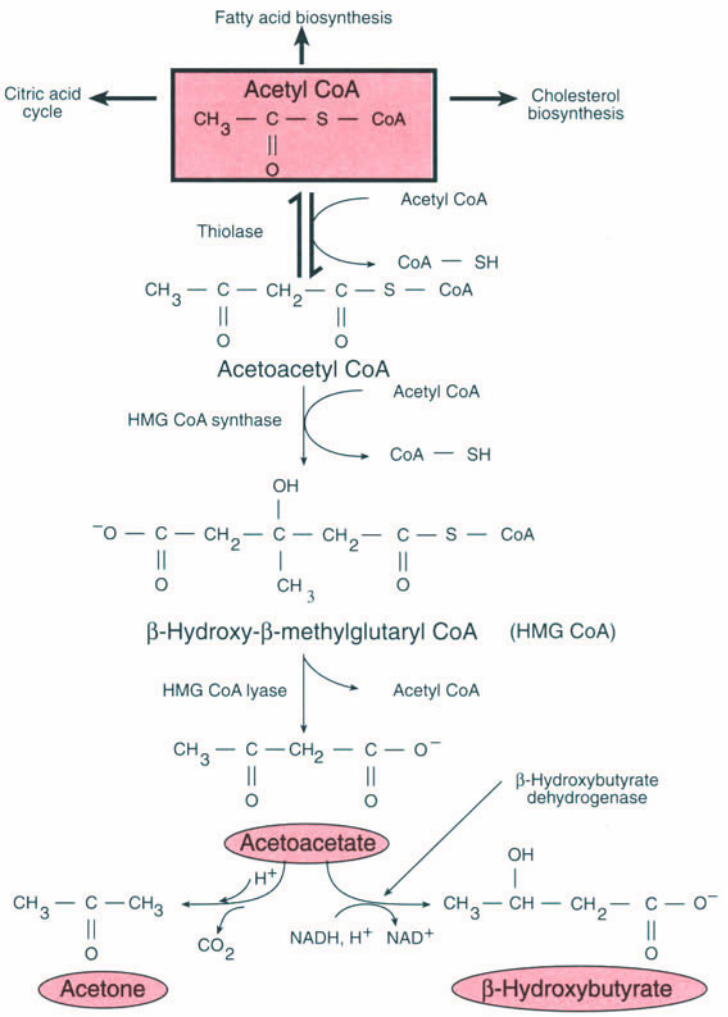


Figure 13.13. Formation of ketone bodies from acetyl CoA.

begin to increase (*ketonemia*). If the condition persists, large quantities of ketone bodies appear in the urine (*ketonuria*). The generalized condition, called **ketosis,** is of clinical concern; it must be treated to avoid serious pathological complications.

13.4.3. Ketosis and Its Implications

Several conditions can lead to accumulation of acetyl CoA and development of ketosis. These conditions include starvation, diabetes (insulin deficiency), and a carbohydrate-deficient diet.

During starvation, the body has to rely on energy stores of glycogen and adipose tissue. Glycogen stores are small and rapidly exhausted. At that point, extensive mobilization of fatty acids from adipose tissue takes place. Subsequent oxidation of the fatty acids generates large amounts of acetyl CoA. Recall that citric acid cycle intermediates, such as oxaloacetate, have fixed levels of concentration. Therefore, only a limited amount of oxaloacetate is available for combination with acetyl CoA to form citrate in the first reaction of the citric acid cycle. Because there is insufficient oxaloacetate to combine with all of the acetyl CoA formed, excess acetyl CoA begins to accumulate. Thus, extensive reliance on lipid catabolism during starvation leads to an overabundance of acetyl CoA.

Both diabetes caused by insulin deficiency and a carbohydrate-deficient diet have the same effect as starvation. In both instances, the body compensates for a deficiency of carbohydrate metabolism by increased lipid metabolism, and acetyl CoA may accumulate faster than it is consumed. In all three cases, *faulty carbohydrate metabolism* becomes expressed as an *abnormality of lipid metabolism.*

Why is excessive formation of ketone bodies of serious clinical concern? The answer lies in an acid–base reaction that involves ketone bodies. Sodium bicarbonate ($NaHCO_3$), termed the **alkaline reserve** of the body, constitutes one of a small number of compounds in metabolism that have strong basic character and the capacity to raise pH. Basicity of $NaHCO_3$ results from both the Na^+ content and the production of OH^- by hydrolysis:

$$HCO_3^- + H_2O \rightarrow H_2CO_3 + OH^-$$
$$\downarrow$$
$$CO_2 + H_2O$$

In contrast, metabolism abounds in acidic compounds that can lower pH. Examples include the many weak acids you encountered as intermediates in the citric acid cycle and glycolysis. Consequently, sodium bicarbonate has a crucial role in maintaining physiological pH values.

When ketone bodies accumulate, acetoacetic acid and β-hydroxybutyric acid react with sodium bicarbonate, forming salts:

$$\underset{\text{Acetoacetic acid}}{CH_3{-}CO{-}CH_2{-}COOH} + NaHCO_3 \rightarrow$$
$$\underset{\text{Sodium acetoacetate}}{CH_3{-}CO{-}CH_2{-}COO^-Na^+} + CO_2 + H_2O$$

The salts formed by these reactions (e.g., sodium acetoacetate) are excreted. Not only has sodium bicarbonate been destroyed, but its sodium ions are also lost from the body. Both effects combine to impair seriously the body's capacity to counteract pH-lowering effects of acidic metabolites. As a result, the overall pH tends to drop, causing *acidosis.* Additionally, the salts formed must be dissolved so that they can be excreted in the urine. Hence, development of acidosis is accompanied by loss of water or *dehydration.*

Symptoms of ketosis become more pronounced as time goes on. In advanced stages, acetone may actually be detected in the breath of afflicted individuals. If not treated, ketosis may lead to a progression of symptoms from nausea to depression of the central nervous system, severe dehydration, deep coma, and ultimately death. Clearly, ketosis cannot be allowed to persist and must be treated promptly.

13.5. FATTY ACID BIOSYNTHESIS

Fatty acid synthesis occurs when the organism needs to store excess energy derived from the diet. The process takes place in the cytosol and involves progressive lengthening of a fatty acid chain by two carbons derived from acetyl CoA. The fatty acids subsequently combine with glycerol and become stored as acylglycerols until needed.

Acetyl CoA is generated inside the mitochondria: the pyruvate dehydrogenase complex catalyzes the conversion of pyruvate to acetyl CoA, and β-oxidation produces acetyl CoA in the thiolase reaction. Before this mitochondria-generated acetyl CoA can participate in fatty acid synthesis, it needs to be transported from the mitochondria to the cytosol. However, the inner mitochondrial membrane is essentially impermeable to acetyl CoA. Accordingly, a shuttle mechanism functions in lieu of direct transport. The pertinent mechanism, called the **tricarboxylate transport system** (Figure 13.14), uses citrate as carrier of the acetyl group. Inside the mitochondria, citrate forms from acetyl CoA and oxaloacetate in the first reaction of the citric acid cycle. In the cytosol, cleavage of citrate, catalyzed by *citrate lyase,* regenerates both acetyl CoA and oxaloacetate:

$$Citrate^{3-} + ATP^{4-} + CoA{-}SH^{4-} \rightarrow$$
$$Acetyl\ CoA^{4-} + ADP^{3-} + P_i^{2-} + oxaloacetate^{2-}$$

Oxaloacetate leads to malate or pyruvate, either one of which can be transported back into the mitochondria. Under some conditions, the shuttle results in production of NADPH that can be used for reductive reactions of fatty acid synthesis. If malate returns to the mitochondria, no NADPH forms, but if malate is converted to pyruvate by *malic enzyme,* NADPH forms in the cytosol. In the latter case, one molecule of NADPH is produced per molecule of acetyl CoA transported, a substantial fraction of the total required for fatty acid synthesis. For example, synthesis of palmitate requires 8 molecules of acetyl CoA and 14 molecules of NADPH. Transport of 8 molecules of acetyl CoA from the mitochondria to the cytosol generates 8 molecules of NADPH, leaving 6 molecules of NADPH to be obtained from other sources. The remaining 6 molecules of NADPH can be furnished by the pentose phosphate pathway in the liver or by malic enzyme in adipose tissue.

13.5.1. Acetyl CoA Carboxylase

Synthesis of fatty acids begins with formation of **malonyl CoA** from acetyl CoA. The reaction, catalyzed by **acetyl CoA carboxylase,** constitutes the *committed step* of fatty acid synthesis.

The acetyl CoA carboxylase of prokaryotes, such as *E. coli,* comprises a complex of three separate proteins. One protein, called **biotin carboxyl carrier protein**

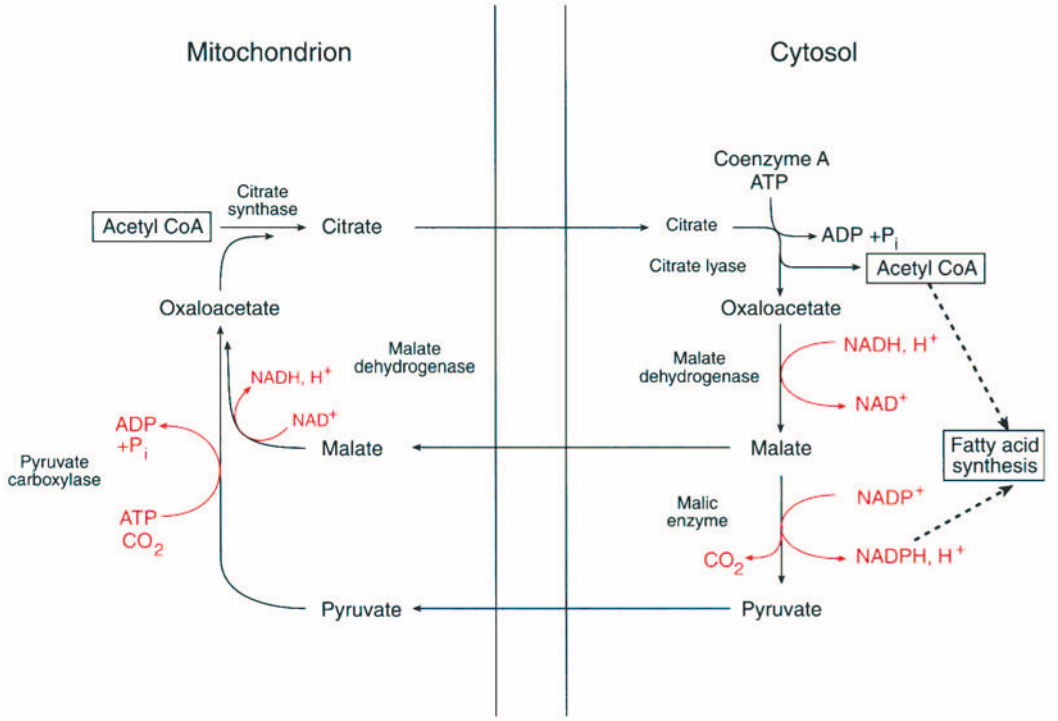


Figure 13.14. The tricarboxylate transport system.

(**BCCP**), serves as a carrier of *biotin.* Biotin is linked covalently to an ϵ-amino group of a lysine residue in the protein, forming *biocytin* (see Figure 11.14). The remaining two proteins of acetyl CoA carboxylase are enzymes, **biotin carboxylase** and **transcarboxylase.** In *E. coli,* the carrier protein and the two enzymes have molecular weights of 23,000, 98,000 (two subunits of 49,000 each), and 130,000, respectively.

Acetyl CoA carboxylase of eukaryotes is a dimer of two identical subunits (MW = 260,000 each). Each subunit contains one molecule of biotin. Both enzymatic activities (biotin carboxylase and transcarboxylase) and the biotin-binding function are located on a single polypeptide chain. Biotin carboxylase catalyzes the following reaction:

$$\begin{array}{c}\text{BCCP-biotin} + \text{ATP}^{4-} + \text{CO}_2 + \text{H}_2\text{O} \\ \downarrow \\ \text{BCCP-carboxybiotin}^- + \text{ADP}^{3-} + \text{P}_i^{2-} + 2\text{H}^+\end{array}$$

Transcarboxylase catalyzes a transfer of the activated carboxyl group from BCCP-carboxybiotin to acetyl CoA (Figure 13.15), yielding malonyl CoA and BCCP-biotin:

$$\begin{array}{c}\text{BCCP-carboxybiotin}^- + \text{acetyl CoA}^{4-} \rightarrow \\ \text{Malonyl CoA}^{5-} + \text{BCCP-biotin}\end{array}$$

13.5.2. Fatty Acid Synthase

The remaining reactions of fatty acid synthesis, beginning with acetyl CoA and malonyl CoA, are catalyzed by the **fatty acid synthase complex.** In animal cells, fatty acid synthase consists of seven enzymatic activities and an acyl carrier protein.

Acyl carrier protein (ACP) is a small protein that functions in fatty acid synthesis much as coenzyme A does in fatty acid degradation. In β-oxidation, fatty acid derivatives become linked to the SH group of a *phosphopantetheine group* that forms part of the structure of coenzyme A. In fatty acid synthesis, fatty acid derivatives become linked to the SH group of a phosphopantetheine group that is covalently attached to an acyl carrier protein (Figure 13.16). ACP, with its attached phosphopantetheine group, carries the acyl group in fatty acid synthesis much as coenzyme A carries it in fatty acid degradation. Scientists believe that the long phosphopantetheine group serves as a "swinging arm" (like the lipoyllysine arm of the pyruvate dehydrogenase complex) that moves the substrate from one catalytic site to another on the enzyme complex (Figure 13.17).

Acyl carrier protein was first isolated from *E. coli.* The 4′-phosphopantetheine group is esterified to the hydroxyl group of serine 36 of the protein. The protein itself

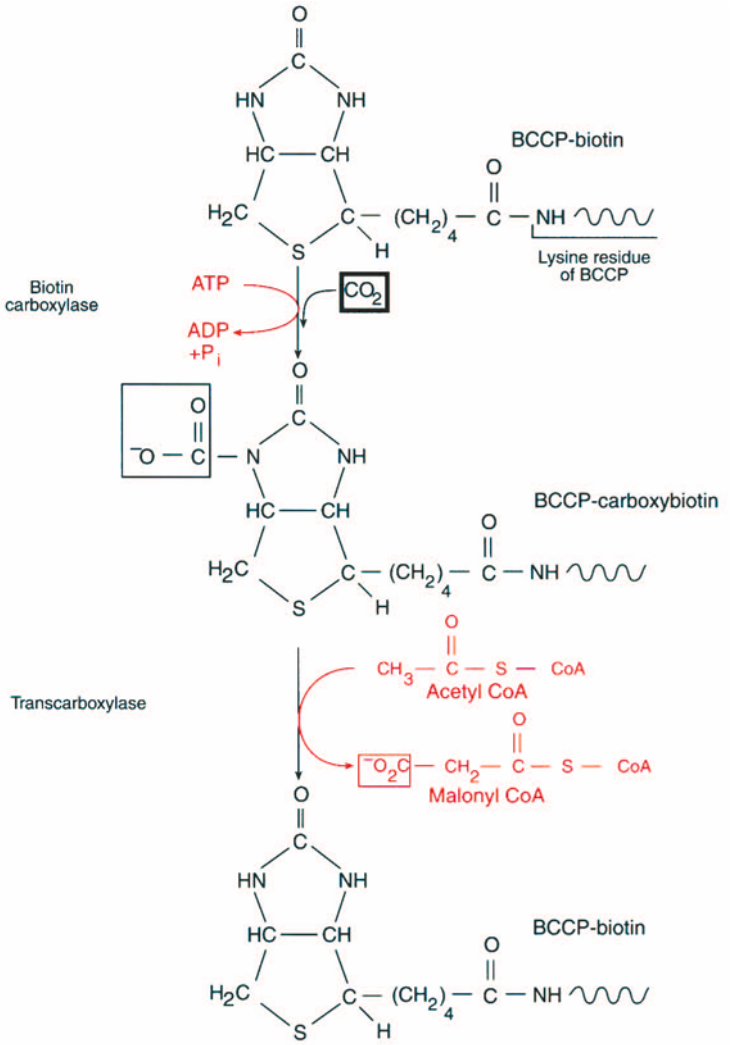
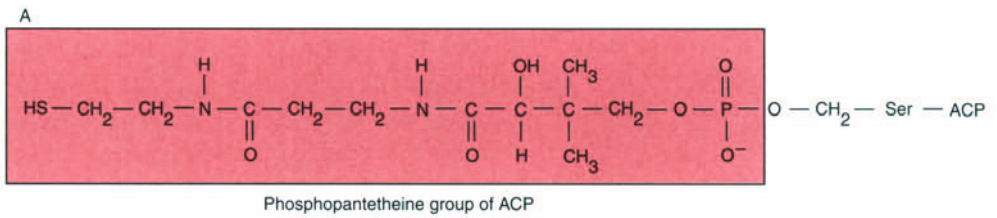


Figure 13.15. Formation of malonyl CoA by acetyl CoA carboxylase.



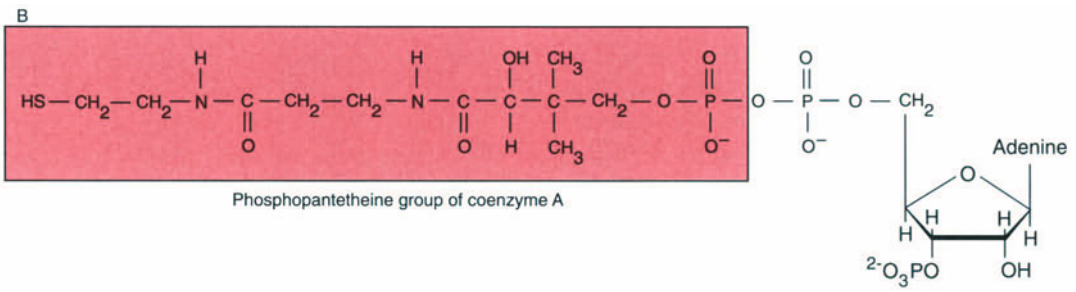


Figure 13.16. The phosphopantetheine group: (A) as a prosthetic group of acyl carrier protein (ACP), where it is esterified to the hydroxyl group of a serine residue; (B) as a structural component of coenzyme A.

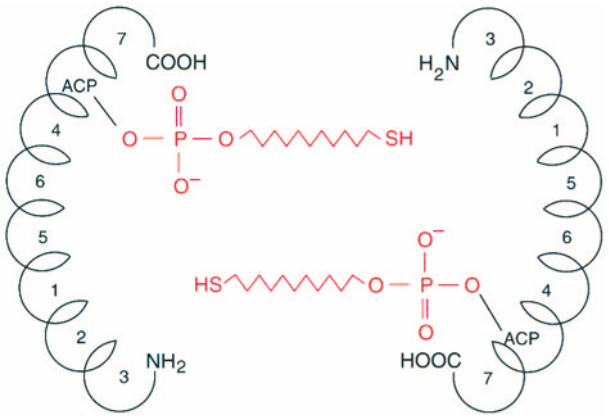


Figure 13.17. The dimeric structure of eukaryotic fatty acid synthase. Each subunit contains an acyl carrier protein (ACP) and catalytic sites of seven enzymatic activities: 1, Acetyl CoA:ACP transacylase; 2, malonyl CoA:ACP transacylase; 3, β-ketoacyl-ACP synthase; 4, β-ketoacyl-ACP reductase; 5, β-hydroxyacyl-ACP dehydratase; 6, enoyl-ACP reductase; 7, palmitoyl-ACP thioesterase. [Adapted, with permission, from S. J. Wakil, J. K. Stoops, and V. C. Joshi, *Annu. Rev. Biochem.* 52:537–579 (1983). © 1983 by Annual Reviews, Inc.]

consists of 77 amino acid residues (MW = 10,000). Assembly of ACP and the enzymes of fatty acid synthesis depends on the type of organism involved. In *E. coli* and plants, fatty acid synthase comprises a multienzyme system, an aggregate of *seven separate polypeptide chains.* One polypeptide chain constitutes ACP, and the remaining *six* represent enzymes.

In yeast, fatty acid synthase also consists of ACP and *six* enzymes except that these are located on *two multifunctional polypeptide chains.* One chain (MW = 185,000) contains the ACP function and two enzymatic activities. The other chain (MW = 175,000) contains the remaining four enzymatic activities. Six dimers associate to form a very large complex (MW ≈ 2.4×10^6).

In animals, fatty acid synthase consists of ACP and *seven* enzymes, all of which are located on *a single multifunctional polypeptide chain* (Figure 13.17). Contiguous regions of the polypeptide chain fold in unique fashion to generate different enzymatic activities and the ACP function. The enzyme is a dimer, containing two identical subunits (MW ≈ 260,000 each) with an antiparallel head-to-tail arrangement. The additional enzymatic activiy of animal fatty acid synthase, called **palmitoyl-ACP thioesterase,** catalyzes the hydrolysis of the final palmitoyl-ACP to palmitate and ACP; it becomes functional only after a fatty acid chain of 16 carbons has been synthesized. Other organisms lack palmitoyl-ACP thioesterase and use palmitoyl-ACP directly.

13.5.3. Individual Reactions

Fatty acid synthesis in animals consists of the following seven reactions, the first six of which are shown in Figure 13.18:

1. Priming: Acetyl CoA:ACP transacylase
2. Loading: Malonyl CoA:ACP transacylase
3. Condensation: β-Ketoacyl-ACP synthase
4. Reduction: β-Ketoacyl-ACP reductase
5. Dehydration: β-Hydroxyacyl-ACP dehydratase
6. Reduction: Enoyl-ACP reductase
7. Release: Palmitoyl-ACP thioesterase

13.5.3A. Acetyl CoA:ACP Transacylase.

Acetyl CoA:ACP transacylase catalyzes a "priming" reaction whereby acetyl CoA becomes transferred, first to ACP and then to an SH group of β-ketoacyl-ACP synthase. We describe the two-step priming process as follows (in mammals, the acetyl-ACP intermediate does not form):

$$\text{Acetyl CoA}^{4-} + \text{HS—ACP}^{-} \rightarrow$$

$$\text{CH}_3\text{—C(=O)—S—ACP}^{-} + \text{CoA—SH}^{4-}$$

Acetyl-ACP (13.1a)

$$\text{CH}_3\text{—C(=O)—S—ACP}^{-} + \text{HS—synthase} \rightarrow$$

$$\text{CH}_3\text{—C(=O)—S—synthase} + \text{HS—ACP}^{-}$$

Acetyl-synthase (13.1b)

13.5.3B. Malonyl CoA:ACP Transacylase.

Malonyl CoA:ACP transacylase catalyzes a reaction whereby the second substrate of β-ketoacyl-ACP synthase is readied by "loading" malonate onto ACP. In this reaction, the free SH group of ACP attacks the carbonyl group of malonyl CoA to form malonyl-ACP. The priming and loading reactions prepare the two substrates for subsequent condensation.

$$^{-}\text{O}_2\text{C—CH}_2\text{—C(=O)—S—CoA}^{4-} + \text{HS—ACP}^{-} \rightarrow$$

Malonyl CoA

$$^{-}\text{O}_2\text{C—CH}_2\text{—C(=O)—S—ACP}^{-} + \text{CoA—SH}^{4-}$$

Malonyl-ACP (13.2)

13.5.3C. β-Ketoacyl-ACP Synthase.

The next reaction involves a *condensation* between the acetyl

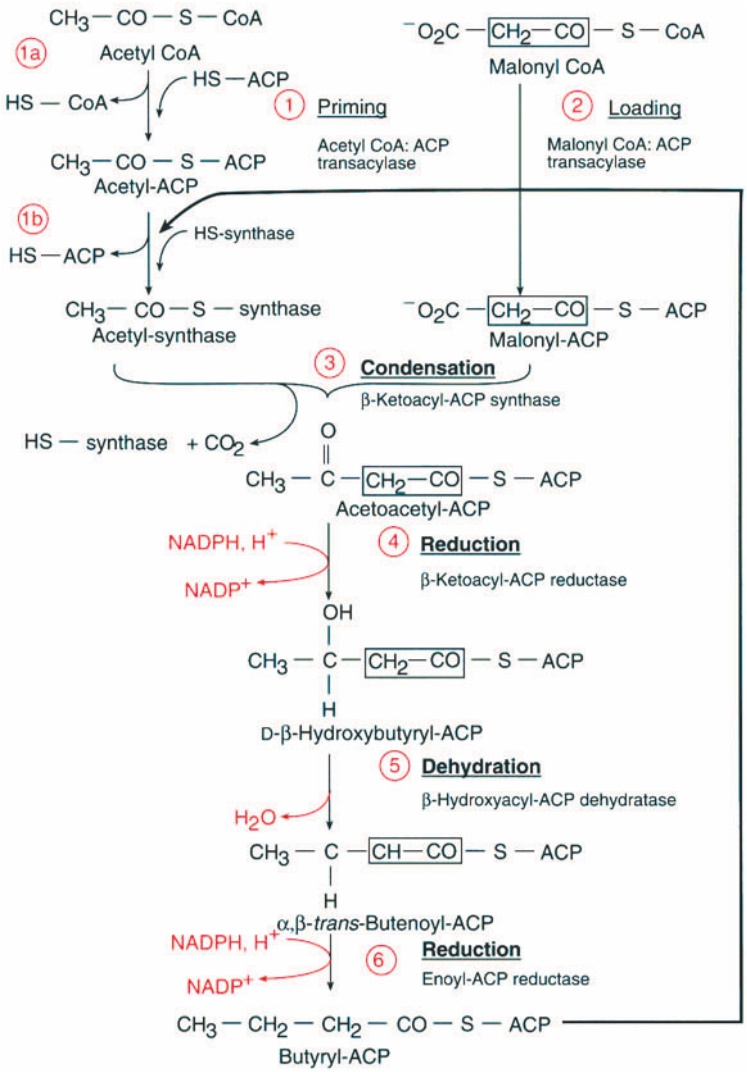


Figure 13.18. Fatty acid synthesis in *E. coli.* In the first round of elongation, an acetyl group becomes extended by a two-carbon fragment (boxed) derived from malonyl CoA. In the second round, butyryl-ACP substitutes for acetyl-ACP in reaction 1b. Subsequent steps lead to hexanoyl-ACP, which then substitutes for acetyl-ACP in reaction 1b, and so on. In mammals, acetyl-ACP is not formed in reaction 1, and elongation stops with palmitoyl-ACP, from which palmitate is then released by palmitoyl thioesterase (reaction 7, not shown).

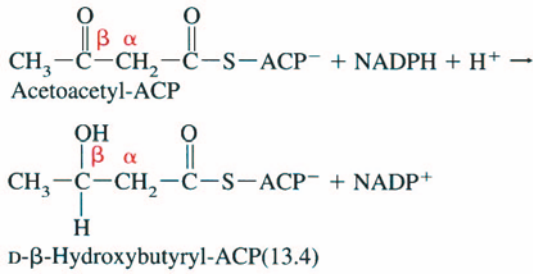
group (linked to an SH group on the enzyme) and the malonyl group (linked to ACP). Decarboxylation activates the methylene carbon of malonyl CoA and makes it a better nucleophile for attacking the carbonyl carbon of the acetyl group. Loss of CO_2 also helps to make this reaction thermodynamically favorable and irreversible. The condensation is catalyzed by **β-ketoacyl-ACP synthase**, and the product formed consists of an acetoacetyl group linked to ACP.

$$\underset{\text{Acetyl-synthase}}{CH_3{-}\overset{O}{\overset{\|}{C}}{-}S{-}\text{synthase}} + {}^{-}O_2C{-}CH_2{-}\underset{\text{Malonyl-ACP}}{\overset{O}{\overset{\|}{C}}{-}S{-}ACP^{-}} + H^{+} \rightarrow$$

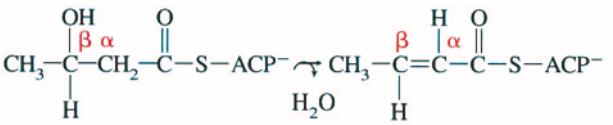
$$\underset{\text{Acetoacetyl-ACP}}{CH_3{-}\overset{O}{\overset{\|}{C}}{}^{\beta}{-}{}^{\alpha}CH_2{-}\overset{O}{\overset{\|}{C}}{-}S{-}ACP^{-}} + HS{-}\text{synthase} + CO_2 \qquad (13.3)$$

13.5.3D. β-Ketoacyl-ACP Reductase.

The reaction catalyzed by **β-ketoacyl-ACP reductase** constitutes the first *reduction* step of fatty acid synthesis. Note that reduction, like oxidation in β-oxidation, takes place at the β-carbon of the molecule. NADPH serves as reducing agent, and the product has the D-configuration.

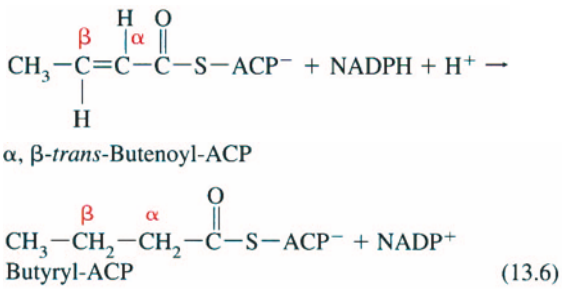


13.5.3E. β-Hydroxyacyl-ACP Dehydratase. The next step represents a *dehydration* reaction, catalyzed by **β-hydroxyacyl-ACP dehydratase.** The reaction yields an α, β-unsaturated compound with a *trans* configuration:



D-β-Hydroxybutyryl-ACP α, β-*trans*-butenoyl-ACP (13.5)

13.5.3F. Enoyl-ACP Reductase. This step represents the end of the first round of fatty acid synthesis, whereby a four-carbon fatty acid forms from two two-carbon fragments. The step involves a second *reduction,* again at the β-carbon and mediated by NADPH. The enzyme **enoyl-ACP reductase** catalyzes this reaction, which yields a four-carbon fatty acid attached to ACP (butyryl-ACP):



(13.6)

At this point, the acetyl group with which the system was originally primed has been elongated by a two-carbon fragment derived from malonate. Butyryl-ACP now becomes the substrate for a second round of elongation. It substitutes for acetyl-ACP, and its acyl group is transferred from butyryl-ACP to the SH group of β-ketoacyl-ACP synthase (reaction 13.1b). The synthase then catalyzes a condensation between the butyryl group (linked to an SH group of the enzyme) and a malonyl group (linked to ACP) according to reaction (13.3).

The condensation yields a six-carbon product, attached to ACP. The condensation product is converted to hexanoyl-ACP, a six-carbon fatty acid attached to ACP, by means of reactions (13.4) through (13.6). In a third round of elongation [reactions (13.1b), (13.3), and (13.4)–(13.6)], another two-carbon fragment is added, forming an eight-carbon fatty acid attached to ACP, and so on.

13.5.3G. Palmitoyl-ACP Thioesterase. During each round of elongation, fatty acid chains are elongated by two carbons. In animal cells, elongation stops with synthesis of a 16-carbon fatty acid in the form of *palmitoyl-ACP. Palmitoyl-ACP thioesterase* then catalyzes a reaction whereby palmitate is *released* from the fatty acid synthase complex. The enzyme catalyzes the hydrolysis of palmitoyl-ACP to palmitate and ACP:

$$\text{Palmitoyl-ACP}^- + H_2O \rightarrow \text{Palmitate}^- + \text{HS—ACP}^- + H^+ \qquad (13.7)$$

Elongation of palmitate to form longer chain fatty acids, as well as introduction of double bonds, must be carried out by other enzyme systems (covered below).

Because the first round of synthesis generates a four-carbon, rather than a two-carbon, fatty acyl-ACP, formation of palmitate (16 carbons) requires that reactions (13.3) through (13.6) be carried out only seven times, using seven malonyl CoA molecules. Synthesis of palmitoyl-ACP yields seven molecules of water, but hydrolysis of palmitoyl-ACP to palmitate and ACP requires one molecule of water, so that the net number of water molecules formed is six. Accordingly, the synthesis of palmitate, beginning with acetyl CoA and malonyl CoA, is described by the following equation:

$$\text{Acetyl CoA}^{4-} + 7\text{malonyl CoA}^{5-} + 14\text{NADPH} + 20H^+ \downarrow \text{Palmitate}^- + 7CO_2 + 14\text{NADP}^+ + 8\text{CoA—SH}^{4-} + 6H_2O$$

To form the seven molecules of malonyl CoA used in this equation requires an input of seven molecules of acetyl CoA:

$$7\text{Acetyl CoA}^{4-} + 7CO_2 + 7\text{ATP}^{4-} + 7H_2O \downarrow 7\text{Malonyl CoA}^{5-} + 7\text{ADP}^{3-} + 7P_i^{2-} + 14H^+$$

Combining the two equations provides the overall stoichiometry for *palmitate synthesis beginning with acetyl CoA:*

$$8\text{Acetyl CoA}^{4-} + 7\text{ATP}^{4-} + 14\text{NADPH} + 6H^+ \downarrow \text{Palmitate}^- + 14\text{NADP}^+ + 8\text{CoA—SH}^{4-} + 6H_2O + 7\text{ADP}^{3-} + 7P_i^{2-}$$

NADPH and H^+ for the reductive reactions come

from the malic enzyme reaction (see Figure 13.14) and the pentose phosphate pathway.

13.5.4. Comparison with β-Oxidation

We can summarize similarities and differences between the pathways of fatty acid synthesis and fatty acid degradation (β-oxidation) as follows:

- The two pathways differ in their compartmentation. Fatty acid synthesis takes place in the cytosol, whereas fatty acid degradation takes place in the mitochondria.
- Both pathways involve intermediates linked to a carrier. In fatty acid synthesis, the carrier is ACP; in fatty acid degradation, it is coenzyme A.
- Four reactions in one pathway are chemically the reverse of four reactions in the other pathway but use different enzymes and cofactors. These are the condensation, reduction, dehydration, and reduction steps in fatty acid synthesis and the oxidation, hydration, oxidation, and cleavage steps in fatty acid degradation.
- Both pathways require a transport mechanism linking the mitochondrial and cytosolic compartments. The tricarboxylate transport system for acetyl CoA functions in fatty acid synthesis, and the carnitine carrier system for fatty acyl CoA functions in fatty acid degradation.
- Both pathways feature progressive alterations of hydrocarbon chains of fatty acids. In fatty acid synthesis, chains become extended by successive additions of two-carbon fragments derived from acetyl CoA but condensed in the form of malonyl CoA. In fatty acid degradation, chains become shortened by successive removal of two-carbon fragments in the form of acetyl CoA.
- Fatty acids are synthesized from the methyl to the carboxyl end of the molecule (the COOH end is synthesized last). Fatty acids are degraded in the opposite direction (the COOH end is removed first).
- The hydroxyacyl intermediate has a D-configuration in fatty acid synthesis but an L-configuration in fatty acid degradation.
- Fatty acid synthesis constitutes a reductive pathway requiring NADPH; fatty acid degradation constitutes an oxidative pathway requiring FAD and NAD^+.
- Even though the "spiral" or "cyclic" portion (the repeated steps) of each pathway requires addition or removal of a two-carbon fragment, this portion is traversed only seven times for either synthesis or degradation of a 16-carbon fatty acid.
- In animals, enzymes of fatty acid synthesis are located on a single polypeptide chain as part of the fatty acid synthase complex. To what degree corresponding enzymes of fatty acid degradation are associated is still unsettled.

13.5.5. Elongases and Desaturases

You saw that in animals fatty acid synthesis stops with the formation of palmitate, a 16-carbon fatty acid. To produce fatty acids having longer carbon chains requires additional reactions, using palmitate as starting material. These reactions are catalyzed by specific **elongases** and occur both in the mitochondria and in the endoplasmic reticulum. In mitochondria, elongation involves a reversal of β-oxidation, with some modifications. In the endoplasmic reticulum, elongation resembles ordinary fatty acid synthesis except that the fatty acid occurs as a coenzyme A derivative rather than as an ACP derivative.

Unsaturated fatty acids form from saturated ones by action of **desaturases.** These enzymes contain nonheme iron, and in eukaryotes they catalyze the following general reaction:

$$\begin{array}{c} CH_3-(CH_2)_x-CH_2-CH_2-(CH_2)_y-CO-S-CoA \\ + \ NADH + H^+ + O_2 \\ \downarrow \\ CH_3-(CH_2)_x-CH=CH-(CH_2)_y-CO-S-CoA \\ + \ NAD^+ + 2H_2O \end{array}$$

In mammals, introducing double bonds requires a mini electron transport system composed of two enzymes and a cytochrome (Figure 13.19). Mammals lack the enzymes capable of introducing double bonds at carbons beyond C(9). For this reason, linoleic acid and linolenic acid cannot be synthesized by mammals and must be obtained through the diet; they constitute *essential fatty acids.*

Odd-numbered fatty acids are synthesized by priming fatty acid synthase with propionyl CoA rather than with acetyl CoA.

13.5.6. Regulation

The reaction catalyzed by acetyl CoA carboxylase represents a major control point of fatty acid biosynthesis. In eukaryotes, the monomeric form of this enzyme (composed of two subunits) is inactive. However, the monomer polymerizes to form long filamentous structures that are active. Activity of acetyl CoA carboxylase is controlled by shifts in the equilibrium between its two forms:

$$\begin{array}{ccc} \text{Monomer} & \rightleftharpoons & \text{polymer} \\ \text{(inactive)} & & \text{(active)} \end{array}$$

Citrate shifts the equilibrium toward the active polymer and stimulates fatty acid synthesis, but researchers

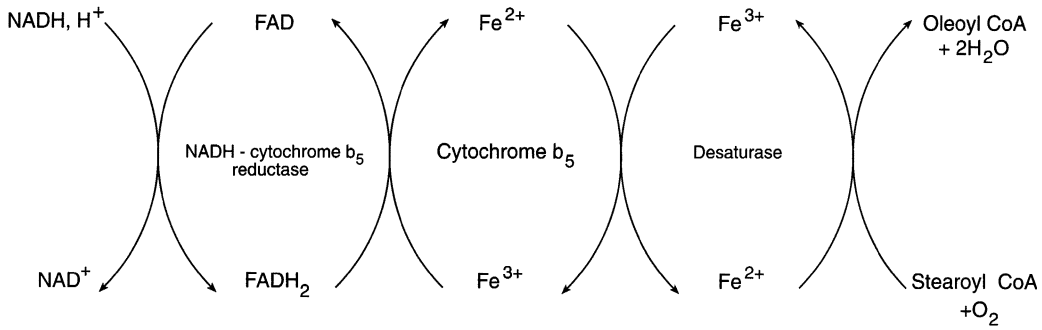


Figure 13.19. The mini electron transport system of fatty acid desaturases in mammals, located on the cytosolic face of the endoplasmic reticulum. Formation of two water molecules constitutes a four-electron reaction; two electrons come from NADH, and two come from the fatty acid bond being reduced.

have not yet decided whether this constitutes a physiological role of citrate. Palmitoyl CoA shifts the equilibrium toward the inactive monomer and inhibits synthesis. Since palmitoyl CoA is the product of fatty acid synthesis, its effect constitutes feedback inhibition.

The equilibrium is also affected by phosphorylation/dephosphorylation of the enzyme in response to certain hormones. Glucagon and epinephrine promote phosphorylation by stimulating cAMP-dependent protein kinase. Phosphorylation shifts the equilibrium toward the inactive monomer (Figure 13.20) and results in inhibition of fatty acid synthesis. Insulin has the opposite effect. Insulin decreases the level of cAMP and promotes dephosphorylation and formation of the active polymer, thus stimulating fatty acid synthesis.

In prokaryotes, acetyl CoA carboxylase is regulated through intracellular levels of guanine nucleotides. Guanosine 5′-triphosphate (GTP) serves as the major source of energy for protein synthesis, where its hydrolysis produces GDP. Thus, in the course of protein synthesis, intracellular levels of GTP and GDP undergo change. The extent of protein synthesis varies, depending on the state of growth of the cells, so that fatty acid synthesis becomes regulated by the growth requirements of cells.

13.6. BIOSYNTHESIS OF OTHER LIPIDS

Free fatty acids do not occur in cells to any great extent. Most fatty acids are present in esterified form as *triacylglycerols* or *glycerophospholipids*. The biosynthesis of these two types of complex lipids occurs primarily on the endoplasmic reticulum of liver cells or fat cells.

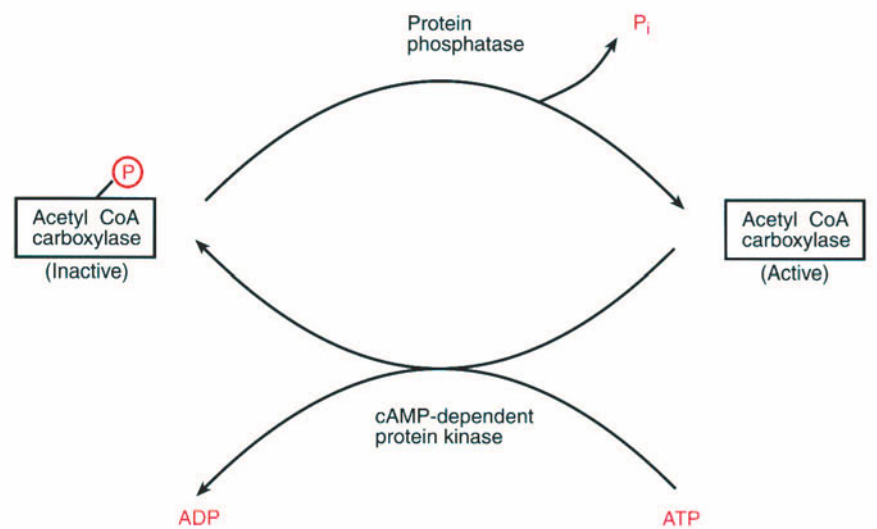


Figure 13.20. Equilibrium between inactive (monomeric) and active (polymeric) acetyl CoA carboxylase. Hormonally controlled phosphorylation and dephosphorylation shift the equilibrium.

13.6.1. Acylglycerols

Acylglycerols are synthesized from two precursors, *fatty acyl CoA* and *glycerol 3-phosphate*. Fatty acyl CoA is the product of fatty acid activation. Glycerol 3-phosphate can be formed either from the glycolytic intermediate dihydroxyacetone phosphate or by phosphorylation of glycerol, a product of the degradation of acylglycerols:

$$\text{Dihydroxyacetone phosphate}^{2-} + \text{NADH} + \text{H}^+ \rightarrow \text{glycerol 3-phosphate}^{2-} + \text{NAD}^+$$

$$\text{Glycerol} + \text{ATP}^{4-} \rightarrow \text{glycerol 3-phosphate}^{2-} + \text{ADP}^{3-} + \text{H}^+$$

Mono- and diacylglycerols form from glycerol 3-phosphate by succesive esterifications with fatty acyl CoA (Figure 13.21). Triacylglycerols form from 1,2-diacylglycerol 3-phosphate (phosphatidic acid) in a two-step process. First, the phosphate group is removed by hydrolysis, yielding 1,2-diacylglycerol. The diacylglycerol then reacts with another molecule of fatty acyl CoA.

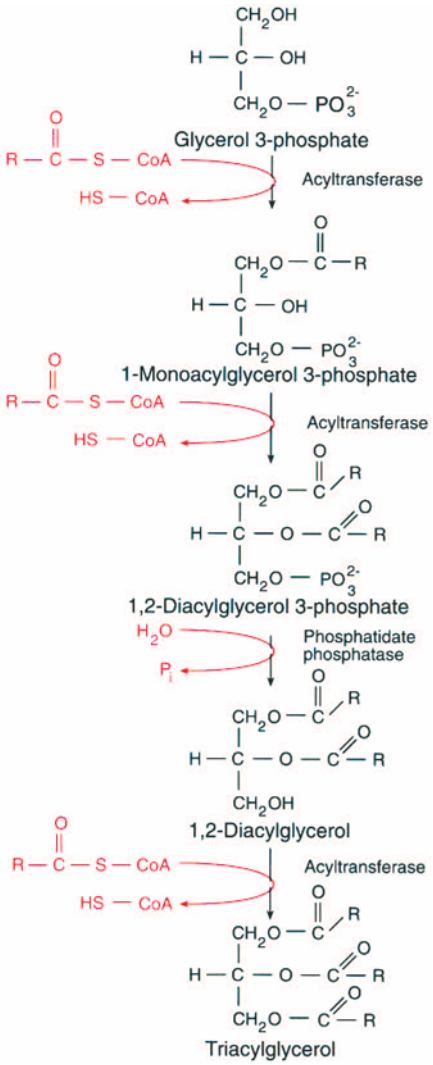


Figure 13.21. Biosynthesis of acylglycerols from glycerol 3-phosphate. Different fatty acyl groups may be transferred at the three acyltransferase steps.

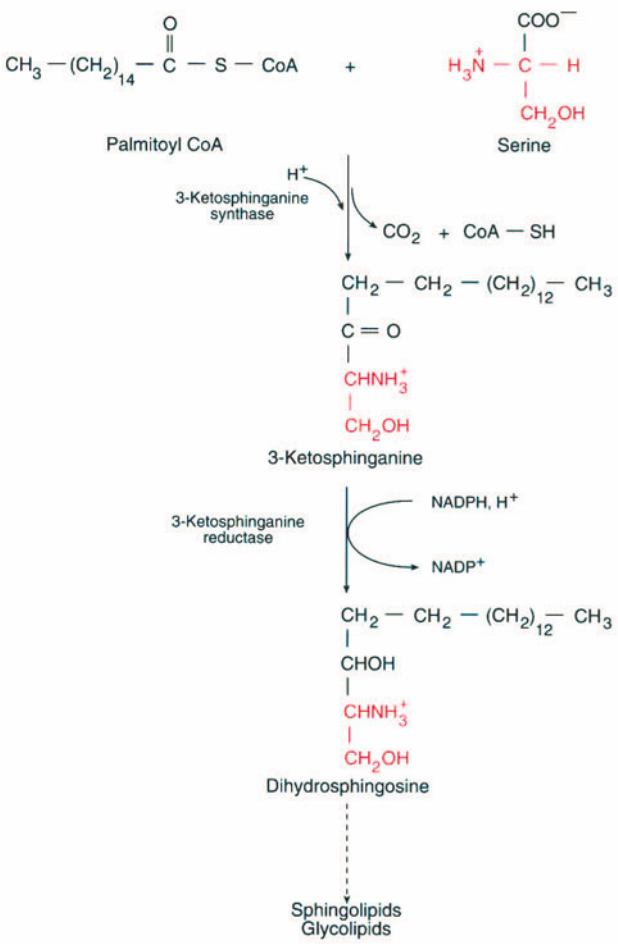


Figure 13.22. Biosynthesis of dihydrosphingosine, the precursor of sphingolipids and glycolipids.

13.6.2. Phospholipids

13.6.2A. Glycerophospholipids. Glycerophospholipids represent important constituents of biological membranes. Pathways for their biosynthesis vary with the type of organism and differ in their details. Some synthetic reactions start with 1,2-diacylglycerol, and others with 1,2-diacylglycerol 3-phosphate. Frequently, CDP (5′-cytidine diphosphate) serves as a carrier for a lipid component much as UDP serves as a carrier for carbohydrates. Some pathways for the biosynthesis of phosphatidyl choline and phosphatidyl ethanolamine require CDP-choline and CDP-ethanolamine, respective-

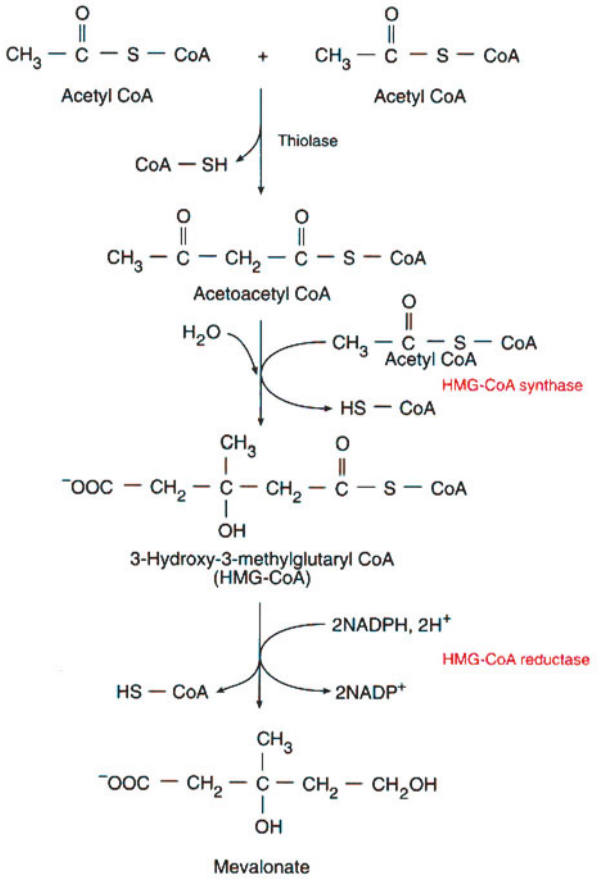
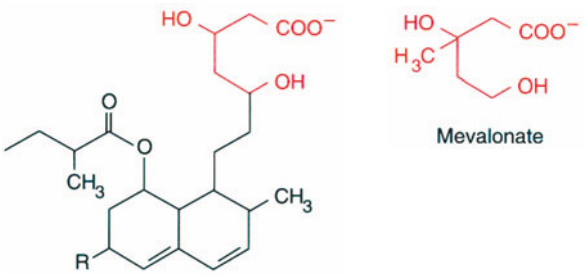


Figure 13.23. Stage I of cholesterol biosynthesis: Formation of mevalonate from three molecules of acetyl CoA.



R = H Compactin
R = CH_3 Lovastatin

Figure 13.24. Examples of drugs used to lower blood cholesterol levels. Structural similarity to mevalonate accounts for the competitive inhibition of HMG-CoA reductase by these compounds.

ly. In these compounds, choline and ethanolamine are esterified via a hydroxyl group to the β-phosphate group of CDP.

13.6.2B. Sphingolipids. Sphingolipids also occur in plant and animal membranes. In mammals, they are particularly abundant in the brain and nervous tissue. Sphingolipids have sphingosine as their structural backbone, but their biosynthetic precursor is dihydrosphingosine, formed from palmitoyl CoA and serine (Figure 13.22). Dihydrosphingosine can be converted to *sphingomyelin, cerebrosides,* or *gangliosides.* A number of hereditary disorders of sphingolipid metabolism occur. These are known as **lipid storage diseases,** and three (gangliosidosis, Gaucher's disease, and Niemann–Pick disease) were listed in Table 8.3.

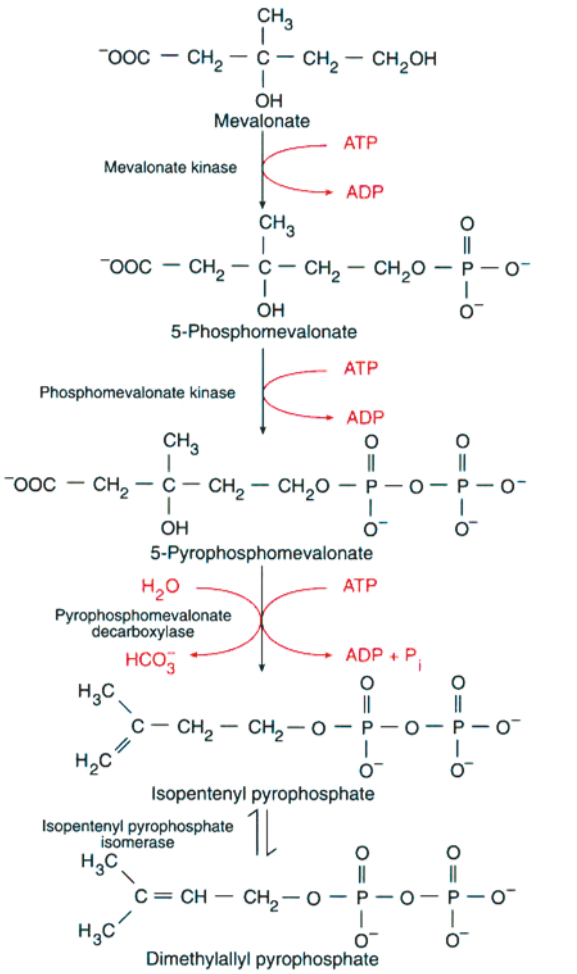


Figure 13.25. Stage II of cholesterol biosynthesis: ATP-dependent conversion of mevalonate to dimethylallyl pyrophosphate.

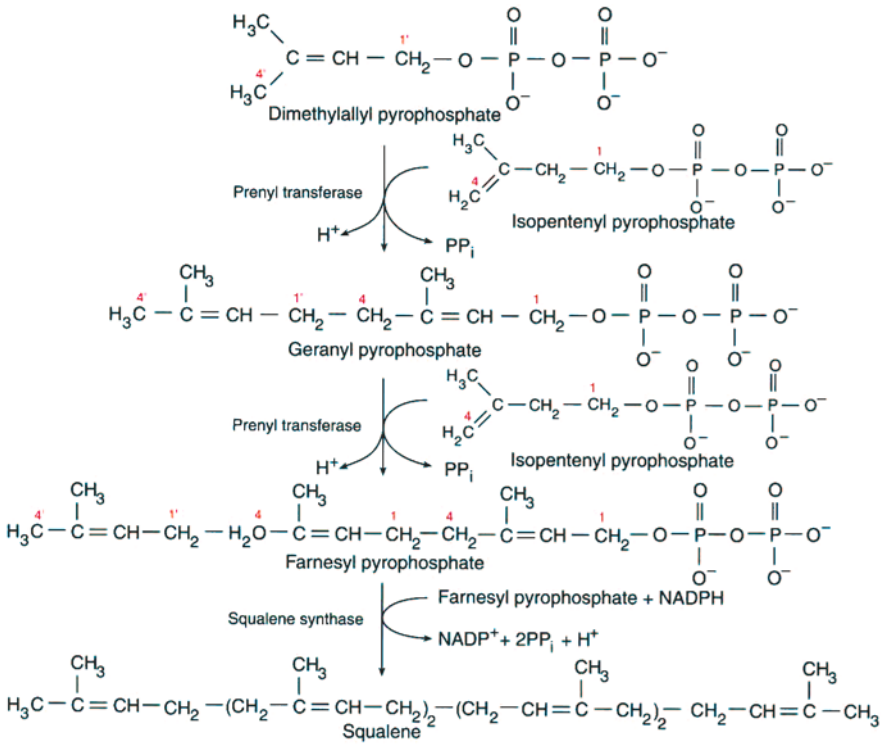


Figure 13.26. Stage III of cholesterol biosynthesis: Condensation reactions from dimethylallyl pyrophosphate to squalene. Loss of PP_i converts dimethylallyl pyrophosphate to an allylic carbonium ion [$(H_3C)_2C{=}CH\text{-}CH_2^+$] that is attacked by the electrons of the double bond in isopentenyl pyrophosphate, resulting in a head-to-tail (1′–4) condensation. The second condensation proceeds likewise, following conversion of geranyl pyrophosphate to an allylic carbonium ion by loss of PP_i.

13.6.3. Cholesterol

The structure of cholesterol was elucidated in the 1930s, but it was not clear how such a complex molecule could be assembled from smaller molecules. The pathway for cholesterol biosynthesis was only unraveled after radioactive tracers became available. In 1941, Rudolf Schoenheimer and Konrad Bloch (who was awarded the Nobel Prize in 1964) showed that labeled acetate was a precursor of cholesterol in rats and mice. In 1952, Bloch proposed a scheme for cholesterol biosynthesis in which the molecule was assembled in a number of stages. The scheme required formation of an unknown precursor in the early part of the pathway. This precursor was identified in 1956 as *mevalonic acid* and provided the missing link for the mechanism of cholesterol biosynthesis. Since then, the individual steps of this lengthy pathway have been defined in detail.

All of the carbons of cholesterol derive from acetyl CoA. Both the methyl and the carboxyl carbons of the acetyl group become incorporated into the steroid nucleus. The entire pathway can be divided into five stages:

$$\underset{C_2}{\text{Acetate}} \xrightarrow{\text{I}} \underset{C_6}{\text{mevalonate}} \xrightarrow{\text{II}} \underset{C_5}{\text{isoprene derivatives}} \xrightarrow{\text{III}} \underset{C_{30}}{\text{squalene}} \xrightarrow{\text{IV}}$$

$$\underset{C_{27}}{\text{lanosterol}} \xrightarrow{\text{V}} \underset{C_{27}}{\text{cholesterol}}$$

13.6.3A. Stage I. The first stage (Stage I) leads to synthesis of **mevalonate,** a six-carbon compound, from three molecules of acetyl CoA (Figure 13.23). The initial step of this stage produces *acetoacetyl CoA* from acetyl CoA by reversal of the thiolase reaction. The last step of Stage I involves the reduction of *3-hydroxy-3-methylglutaryl CoA (HMG-CoA)* to mevalonate. This reaction requires two molecules of NADPH and is catalyzed by **HMG-CoA reductase;** it constitutes both the committed and the rate-determining step of cholesterol biosynthesis. The reaction also represents the primary control site of cholesterol biosynthesis. The enzyme is subject to both covalent modification and allosteric control. HMG-CoA reductase exists in two interconvertible forms, a phosphorylated and a dephosphorylated form. Phosphorylation inactivates the enzyme and involves a cAMP-depen-

dent bicyclic enzyme cascade similar to those that function in the control of glycogen phosphorylase and glycogen synthase (see Figures 10.26 and 10.30). The enzyme cascade is hormonally controlled: insulin, which decreases [cAMP], stimulates cholesterol biosynthesis; glucagon, which increases [cAMP], inhibits it.

HMG-CoA reductase is inhibited by long-chain fatty acyl CoA molecules. The inhibition may be due to both a direct allosteric effect on HMG-CoA reductase and an indirect effect produced by an activation of the kinase that catalyzes the phosphorylation of the reductase. Cholesterol also affects the activity of HMG-CoA reductase. High concentrations of cholesterol lead to formation of cholesterol derivatives that inhibit the enzyme allosterically and lead to increased degradation and decreased synthesis of HMG-CoA reductase.

Because of its pivotal role in the regulation of cholesterol biosynthesis, HMG-CoA reductase has been a target of attempts to lower cholesterol levels by means of drug therapy. Two examples of such drugs are the fungal products *compactin* and *lovastatin* (Figure 13.24). These compounds bear a structural similarity to mevalonate and function as competitive inhibitors of HMG-CoA reductase. Lovastatin was approved in 1987 for the treatment of patients with very high cholesterol levels. Drug therapy represents one of two lines of therapy for lowering high levels of serum cholesterol. The other line involves dietary modifications to decrease the intake of dietary cholesterol.

13.6.3B. Stage II. Stage II of cholesterol biosynthesis involves a phosphorylation and decarboxylation of mevalonate. The product, *isopentenyl pyrophosphate,* is a derivative of *isoprene* (see Figure 6.16). Isopentenyl pyrophosphate isomerizes to a second pyrophosphate derivative of isoprene, *dimethylallyl pyrophosphate* (Figure 13.25). Stage II requires the expenditure of energy; three molecules of ATP are cleaved to ADP and P_i.

13.6.3C. Stage III. Stage III consists of a series of condensation reactions involving a total of six pyrophosphate derivatives of isoprene. First, dimethylallyl pyrophosphate condenses head-to-tail (1′–4) with isopentenyl pyrophosphate to yield *geranyl pyrophosphate* (C_{10}). A second head-to-tail condensation of geranyl pyrophosphate with isopentenyl pyrophosphate produces *farnesyl pyrophosphate* (C_{15}). Lastly, two molecule of farnesyl pyrophosphate condense head-to-head (1′–1) to form the 30-carbon, open-chain, unsaturated hydrocarbon *squalene* (Figure 13.26). Stage III requires one molecule of NADPH and leads to the release of four pyrophosphate groups (PP_i). The exergonic hydrolysis of pyrophosphate, catalyzed by pyrophosphatase, provides a driving force for cholesterol biosynthesis.

13.6.3D. Stage IV. Squalene is the immediate precursor of the sterols. Its two-step cyclization (Stage IV) leads to *lanosterol.* The first step, catalyzed by *squalene monooxygenase,* requires molecular oxygen and NADPH and forms a reactive intermediate, called *squalene epoxide.* The second step, catalyzed by *squalene cyclase,* produces the remarkable closure of the ring system (Figure 13.27). This step requires a *concerted* movement of electrons through four double bonds and migration of two methyl groups. The term concerted means that each part of the reaction is essential for any other part to take place; all parts occur simultaneously. The total number of enzymatic steps from mevalonate to lanosterol is about 10.

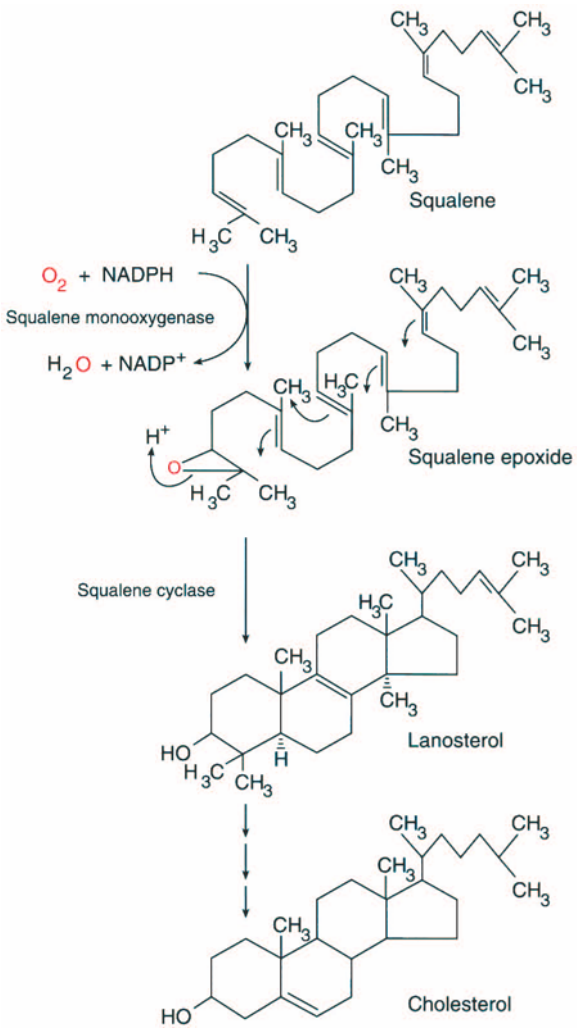


Figure 13.27. The last two stages of cholesterol biosynthesis. Stage IV: Cyclization of squalene to lanosterol by a two-step process. Stage V. Conversion of lanosterol to cholesterol by approximately 20 steps.

13.6.3E. Stage V. The last stage of cholesterol biosynthesis (Stage V) consists of the conversion of lanosterol to cholesterol and includes approximately 20 additional steps, most of which require both NADH (or NADPH) and molecular oxygen. You can see that cholesterol biosynthesis requires expenditure of energy in the form of ATP, is driven in part by hydrolysis of pyrophosphate, and requires a large amount of reducing power in the form of NADH and NADPH. The NADPH required for the reductive reactions comes from the malic enzyme reaction and the pentose phosphate pathway.

SUMMARY

Fats are stored in adipose tissue and transported via the blood as lipoproteins, as fatty acids bound to serum albumin, and as ketone bodies. A hormonally controlled enzyme cascade regulates the release of fatty acids from adipose tissue (mobilization). Execssive mobilization may occur in diabetics and may lead to development of fatty liver.

Fat catabolism begins with hydrolysis, producing glycerol and fatty acids. Glycerol is converted to dihydroxyacetone phosphate, an intermediate of glycolysis. Fatty acids are degraded by β-oxidation, in which chains are shortened by successive removal of two-carbon fragments in the form of acetyl CoA. The acetyl CoA enters the citric acid cycle, yielding NADH and $FADH_2$, which lead to production of ATP via oxidative phosphorylation. Fatty acid activation, located in the cytosol, converts a fatty acid to a fatty acyl CoA that is transported into the mitochondrial matrix by the carnitine carrier system. The remaining four reactions of β-oxidation take place in the matrix. Even-numbered fatty acids are degraded completely to acetyl CoA; odd-numbered fatty acids are degraded to acetyl CoA and propionyl CoA, which is converted to succinyl CoA, an intermediate of the citric acid cycle.

Ketone bodies include acetone, acetoacetate, and β-hydroxybutyrate. Their concentration builds up (ketosis) when acetyl CoA accumulates, as may occur in starvation or diabetes. Ketosis, if not treated, leads to destruction of the alkaline reserve of the body ($NaHCO_3$) and to severe dehydration.

In fatty acid synthesis, located in the cytosol, chains are elongated by two carbons at a time, with the carbons provided in the form of malonyl CoA that is produced from acetyl CoA. Acetyl CoA forms in the mitochondria and is transported to the cytosol by means of the tricarboxylate transport system. Fatty acid synthesis in animals is catalyzed by fatty acid synthase, a dimer of two multifunctional polypeptide chains. Each chain comprises seven enzymatic activities and an acyl carrier protein. Fatty acid synthesis requires both ATP and NADPH.

Acylglycerols are synthesized from fatty acyl CoA and glycerol 3-phosphate. Sphingolipids are produced from dihydrosphingosine. Cholesterol is synthesized entirely from acetyl CoA by a series of some 30 reactions.

SELECTED READINGS

Barron, J. T., Kopp, S. J., Tow, J., and Parrillo, J. E., Fatty acid, tricarboxylic acid cycle metabolites, and energy metabolism in vascular smooth muscle, *Am. J. Physiol.* 267:H764–H769 (1994).

Bieber, L. L., Carnitine, *Annu. Rev. Biochem.* 88:261–283 (1988).

Bradley, W. A., Gianturco, S. H., and Segrest, J. P. (eds.), Plasma lipoproteins, Part C, *Methods in Enzymology,* Vol. 263, Academic Press, San Diego (1996).

Carman, G. M., and Zeimetz, G. M., Regulation of phospholipid biosynthesis in the yeast *Saccharomyces cerevisiae, J. Biol. Chem.* 271:13293–13296 (1996).

Cohen, B. I., Mikami, T., Ayyad, N., Mikami, Y., and Mosbach, E. H., Dietary fat alters the distribution of cholesterol between vesicles and micelles in hamster bile, *Lipids* 30:299–305 (1995).
Hardie, D. G., Regulation of fatty acid synthesis via phosphorylation of acetyl-SCoA carboxylase, *Prog. Lipid. Res.* 28:117–146 (1989).
Kent, C., Eukaryotic phospholipid biosynthesis, *Annu. Rev. Biochem.* 64:315–344 (1995).
Moore, T. S., Jr., *Lipid Metabolism in Plants,* CRC Press, Boca Raton, Florida (1993).
Müller-Newen, G., Janssen, U., and Stoffel, W., Enoyl-CoA hydratase and isomerase form a superfamily with a common active-site glutamate residue, *Eur. J. Biochem.* 228:68–73 (1995).
Salvayre, R., Douste-Blazy, L., and Gatt, S. (eds.), *Lipid Storage Disorders: Biological and Medical Aspects,* Plenum, New York (1988).
Sevanian, A., *Lipid Peroxidation in Biological Systems,* AOCS Press, Champaign, Illinois (1988).
Tuomanen, E., Breaching the blood–brain barrier, *Sci. Am.* 268:80–85 (1993).
Vance, D. E., and Vance, J. E. (eds.), *Biochemistry of Lipids, Lipoproteins, and Membranes,* Elsevier, Amsterdam (1991).
Wakil, S. J., Fatty acid synthase, a proficient multifunctional enzyme, *Biochemistry* 28:4523–4530 (1989).

REVIEW QUESTIONS

A. Define each of the following terms:

Depot fat	Ketone bodies
Lipotropic agents	*S*-Adenosylmethionine
Fatty liver	Ketosis
Alkaline reserve	Blood–brain barrier
β-Oxidation	ACP

B. Differentiate between the two terms in each of the following pairs:

Thiokinase/thiolase	Elongase/desaturase
Tricarboxylate transport system/carnitine carrier system	β-Ketoacyl-ACP synthase/β-ketoacyl-ACP reductase
Acetyl CoA:ACP transacylase/malonyl CoA:ACP transacylase	Transcarboxylase/biotin carboxylase
Acyl CoA dehydrogenase/L-3-hydroxyacyl CoA dehydrogenase	Fatty acid mobilization/fatty acid activation

C. (1) Outline the major steps in β-oxidation of saturated, even-and odd-numbered fatty acids. Write formulas for the intermediates in the β-oxidation of myristic acid (14 carbons, saturated).
(2) How does the overall metabolism of a diabetic differ from that of a normal individual? Why are ketosis and fatty liver often associated with diabetes?
(3) Describe Knoop's experiments. What was the significance of these experiments and how were the results interpreted?
(4) What are the metabolic fates of acetyl CoA?
(5) Compare and contrast the pathways of fatty acid synthesis and fatty acid degradation.

PROBLEMS

13.1. How many molecules of acetyl CoA, $FADH_2$, and NADH are produced in β-oxidation per molecule of a saturated fatty acid that has: (a) 24 carbons; (b) 21 carbons?

13.2. Which carbons in cholesterol will become radioactively labeled when it is synthesized from acetyl CoA in which both carbons of the acetyl group are ^{14}C?

13.3. How many carbons are there in the acyl group of fatty acyl CoA formed from a 17-carbon fatty acid that has been subjected to three cycles of β-oxidation?

13.4. Energetically speaking, how many molecules of glucose can be converted to glyceraldehyde 3-phosphate for every molecule of palmitic acid (16 carbons) that is completely degraded to CO_2 and H_2O?

13.5. You are told that some animals have adapted to life in arid environments by using their lipid stores to produce metabolic water. How can this be?

13.6. Write a balanced equation for the overall β-oxidation to acetyl CoA of: (a) arachidic acid (see Table 6.1); (b) a 15-carbon, saturated fatty acid.

13.7. What is the similarity among the mechanisms of hydrolysis, phosphorolysis, and thiolysis?

13.8. Why is it advantageous that fatty acid mobilization is regulated by means of a hormonally controlled enzyme cascade?

13.9.* Calculate the net yield of ATP molecules per molecule of glycerol for the following catabolic sequences. Assume that the reactions occur under aerobic conditions with participation of the glycerol phosphate shuttle.

(a) Glycerol to pyruvate
(b) Glycerol to acetyl CoA
(c) Glycerol to CO_2 and H_2O

13.10.* Calculate the net yield of ATP molecules per molecule of acylglycerol for the following catabolic sequences. The names refer to the mono- and triacylglycerols, respectively, of palmitic and stearic acid. Assume aerobic conditions and participation of the glycerol phosphate shuttle.

(a) Monopalmitin to acetyl CoA
(b) Tristearin to CO_2 and H_2O

13.11.* Assume that 15% of the body mass of a 70.0-kg adult consists of adipose tissue composed entirely of tristearin (MW = 892). Tristearin is oxidized as indicated in the previous problem.

(a) What is the fuel reserve of this adipose tissue in terms of kilojoules?
(b) Using a daily energy requirement of 12,134 kJ (Table 8.6), calculate the number of days this adult could survive using only depot fat as a source of energy.
(c) What is the daily weight loss of the individual under these starvation conditions?

13.12.* According to a popular misconception, camels store water in their humps for long journeys through the desert. In actuality, these humps consist of large fat deposits that serve as a source of metabolic water. Assuming that such a deposit consists entirely of tristearin (the triacylglycerol of stearic acid; MW = 892), calculate the amount of water (in milliliters) that could be produced from β-oxidation of the stearic acid contained in 1.00 kg of fat. Ignore the water required for β-oxidation, and use 1.00 g/ml for the density of water.

13.13.* How much glycogen would an animal have to store in order to have an amount of stored energy equivalent to that contained in 1.00 g of tristearin (see Problem 13.10)? Given that 1.00 g of glycogen sequesters 2.80 g of bound water, what is the total weight that would have to be stored? (MW of any glucose residue = 162.) Assume aerobic conditions and operation of the glycerol phosphate shuttle.

13.14. Which of the following constitute energy-rich compounds? (a) Fatty acyl adenylate; (b) succinyl CoA; (c) pyrophosphate; (d) carnitine; (e) propionyl CoA; (f) acetoacetic acid; (g) acetyl CoA; (h) β-hydroxybutyric acid; (i) trans-Δ^2-enoyl CoA; (j) β-ketoacyl CoA; (k) acetone; (l) acetoacetyl CoA

13.15. A diabetic suffers from a severe case of ketosis. When acetyl CoA (labeled with ^{14}C in both carbons of the acetyl group) is administered to the diabetic, is it likely that her breath will contain labeled acetone? Why or why not?

13.16.* Consider an even-numbered, saturated fatty acid in which all of the methylene (CH_2) groups are doubly labeled (^{14}C and ^{3}H). When this fatty acid is catabolized under aerobic conditions, will labeled CO_2 be formed when (a) the first and (b) the last acetyl CoA produced from the fatty acid passes through one turn of the citric acid cycle? Will labeled water be produced under the same conditions?

13.17. Write balanced equations for the conversion of: (a) glycerol to pyruvate; (b) propionyl CoA to succinyl CoA; (c) acetyl CoA to acetoacetate.

13.18. A person subsists on a diet that is entirely devoid of carbohydrates. Would he be better off consuming odd- or even-numbered fatty acids? Why?

13.19. Some bacteria can grow using hydrocarbons as their sole nutrients. The hydrocarbons serve as a source of both carbon atoms and energy and are oxidized to their corresponding carboxylic acids. For example,

$$\underset{\text{octane}}{NAD^+ + CH_3(CH_2)_6CH_3} + O_2 \rightarrow \underset{\text{octanoic acid}}{CH_3(CH_2)_6COOH} + NADH + H^+$$

How might such organisms be used to clean up oil spills?

13.20.* What is the net yield of ATP per molecule of acetoacetate when this ketone body is used as a source of energy in the brain?

13.21. Palmitic acid is labeled with ^{14}C at the following specific sites: (a) C(1); (b) C(2); (c) C(7); (d) C(1, 4); (e) C(16). During which cycles in β-oxidation will each of these molecules yield labeled acetyl CoA?

13.22. Write a balanced equation for the complete oxidation of octanoic acid (Problem 13.19) to CO_2 and H_2O.

13.23. Which of the following is *not likely* to lead to the development of fatty liver?

(a) Excessive fatty acid mobilization
(b) A high-starch diet
(c) Diabetes
(d) Starvation
(e) Extensive exposure to pyridine

13.24. Consider the complete oxidation of glucose to CO_2 and H_2O via glycolysis/glycerol phosphate shuttle and the citric acid cycle/electron transport system. Under these conditions, how many moles of palmitate can be synthesized from acetyl CoA per mole of glucose catabolized? Base your calculation on the number of moles of ATP involved.

13.25. Calculate the number of moles of ATP (a) obtainable per mole of myristic acid (14 carbons, saturated) when it is degraded via β-oxidation to acetyl CoA and (b) required to synthesize a mole of myristic acid from acetyl CoA. How do you explain the difference between the two numbers?

13.26. A cell-free system for fatty acid synthesis is exposed to ^{14}C-labeled CO_2. Will the palmitate produced be labeled and, if so, at what position(s)?

13.27.* How many deuterium atoms (D, heavy isotope of hydrogen) are incorporated into palmitate, and what are their locations, when a fatty acid-synthesizing system is provided with the following? (a) Labeled acetyl CoA (CD_3—CO-S-CoA) and unlabeled malonyl CoA; (b) labeled malonyl CoA ($^-$OOC-CD_2-CO-S-CoA) and unlabeled acetyl CoA.

13.28. On the basis of this chapter, predict the effects on lipid metabolism that are likely to arise from insulin deficiency.

Amino Acid and Nucleotide Metabolism

14

Amino acid and nucleotide metabolism, like that of carbohydrates and lipids, can be divided into five major parts—*digestion, transport, storage, degradation,* and *biosynthesis*.

Digestion of proteins (see Section 8.4) produces free amino acids that are absorbed across the intestinal wall into the bloodstream. Proteins and amino acids are transported in biological fluids primarily as plasma proteins, lipoproteins, nucleoproteins, peptides, and free amino acids. In metabolism, nitrogenous compounds are degraded to, and synthesized from, intermediates of carbohydrate and lipid metabolism.

Animals do not store proteins as they do carbohydrates and lipids. The muscles of adult organisms are a functional tissue, not a storage form of energy. Likewise, nucleic acids and their components function in genetic information transfer and not as deposits of metabolic energy. If it becomes necessary for the body to draw on protein tissues, as during fasting or starvation, a "sparing effect" ensures use of carbohydrate and lipid deposits before protein tissues are attacked. Because of the absence of protein stores, humans cannot tolerate a protein-deficient diet, especially one lacking in essential amino acids, for prolonged times. Nutritional deficiency of protein produces *kwashiorkor* (an African word, meaning "weaning disease"), a disease that frequently strikes children who, after having been weaned, exist on a diet containing insufficient protein. The severity of kwashiorkor, particularly in growing children, is due not only to lack of dietary protein but also to breakdown of the body's own proteins.

14.1. NITROGEN UTILIZATION

All living organisms require nitrogen to synthesize building blocks like amino acids, purines, and pyrimidines and to synthesize other nitrogen-containing compounds. All of the nitrogen found in biological systems ultimately derives from gaseous N_2, which constitutes about 80% of the atmosphere. We can divide the overall process of nitrogen utilization by living organisms into three stages:

- Stage I: *Formation of ammonia.* Ammonia forms in living organisms via two major mechanisms. Some bacteria are capable of **nitrogen fixation,** involving reduction of nitrogen gas (N_2) to am-

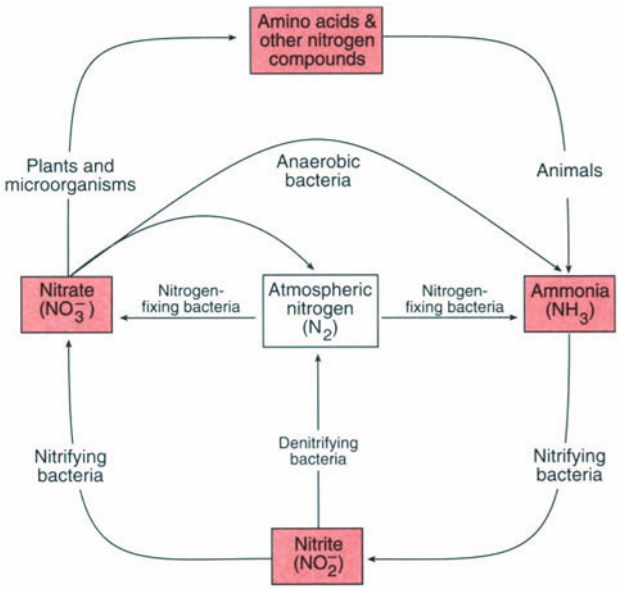


Figure 14.1. The nitrogen cycle.

monia (NH_3). Other bacteria, plants, and fungi carry out **nitrate assimilation,** in which they convert nitrite (NO_2^-) and nitrate (NO_3^-) to ammonia. The interconversions of nitrogen gas, nitrate, nitrite, and ammonia constitute part of the **nitrogen cycle** (Figure 14.1).

- Stage II: *Utilization of ammonia.* All living organisms can use ammonia and convert it to nitrogen-containing organic compounds. Conversion of ammonia to organic compounds, **ammonia fixation,** is accomplished by three reactions, one or more of which occur in all levels of organisms. Animals require ammonia as starting material for using nitrogen.
- Stage III: *Nitrogen transfer.* Nitrogen-containing compounds formed by ammonia fixation participate in metabolic reactions involving intermediates of carbohydrate and lipid metabolism. Such reactions (for example, *transamination*) generate a variety of other nitrogen-containing biomolecules.

14.1.1. Nitrogen Fixation

Some organisms can use atmospheric nitrogen directly and reduce it to ammonia. Nitrogen-fixing bacteria, sometimes in symbiosis with plants, carry out this conversion. Nitrogen fixation is catalyzed by **nitrogenase,** an enzyme complex composed of two parts. One part, *MoFe-protein,* is a tetrameric protein ($\alpha_2\beta_2$) containing both molybdenum and iron; it has a molecular weight of about 220,000. The second part, *Fe-protein,* is a dimer of two identical subunits, contains only iron, and has a molecular weight of about 64,000. Nitrogenase has several oxidation–reduction centers, and most of its iron occurs in the form of [4Fe-4S] complexes (see Figure 12.5). The overall reaction is:

$$N_2 \xrightarrow[6e^-,\ 6H^+]{\text{nitrogenase}} 2NH^3$$

Nitrogen fixation is initated by generating electrons through photosynthesis or from oxidation reactions. These electrons are transferred to *ferredoxin,* an electron carrier containing [4Fe-4S] complexes, which transfers the electrons to the Fe-protein component of nitrogenase (Figure 14.2).

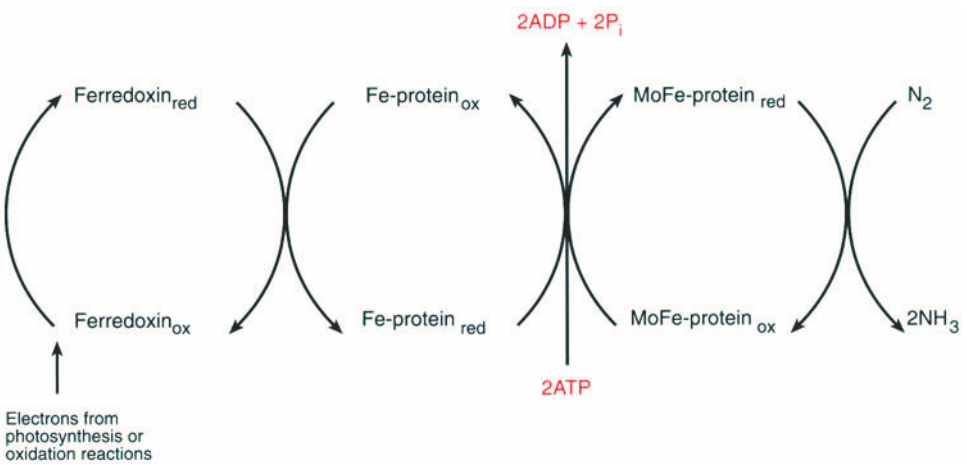


Figure 14.2. The electron transport system of nitrogenase.

Two molecules of ATP bind to reduced Fe-protein and are hydrolyzed as the electron passes from Fe-protein to MoFe-protein. Chemists think that ATP hydrolysis produces a conformational change in Fe-protein that results in the redox potential of Fe-protein becoming significantly more negative; $E^{\circ\prime}$ changes from -0.29 to -0.40 V. The change in $E^{\circ\prime}$ permits the electron to reduce N_2 to NH_3, a reaction that has an $E^{\circ\prime}$ of -0.34 V.

Actual reduction of nitrogen occurs on the MoFe-protein and proceeds in three discrete steps, each requiring two electrons:

$$\begin{aligned} N_2 + 2H^+ + 2e^- &\rightarrow HN{=}NH \\ &\quad \text{Diimine} \\ HN{=}NH + 2H^+ + 2e^- &\rightarrow H_2N{-}NH_2 \\ &\quad \text{Hydrazine} \\ H_2N{-}NH_2 + 2H^+ + 2e^- &\rightarrow 2NH_3 \\ &\quad \text{Ammonia} \end{aligned}$$

Because iron–sulfur proteins undergo oxidation–reduction reactions involving a single electron at a time, the electron transfer from ferredoxin to MoFe-protein must occur six times per N_2 and, therefore, requires hydrolysis of 12 molecules of ATP. This leads to the following overall stoichiometry:

$$N_2 + 6e^- + 6H^+ + 12ATP^{4-} + 12H_2O \rightarrow 2NH_3 + 12ADP^{3-} + 12P_i^{2-} + 12H^+$$

so that the *net* equation is

$$N_2 + 6e^- + 12ATP^{4-} + 12H_2O \rightarrow 2NH_3 + 12ADP^{3-} + 12P_i^{2-} + 6H^+$$

You can see that these reactions represent an energetically costly process. Actually, the energy price tag is even higher than indicated. Nitrogenase also reduces H^+ to H_2, which reacts with diimine to re-form the original substrate (N_2):

$$HN{=}NH + H_2 \rightarrow N_2 + 2H_2$$

When you combine this equation with that of diimine formation, you obtain an overall reaction

$$NH{=}NH + N_2 + 2H^+ + 2e^- \longrightarrow NH{=}NH + N_2 + H_2$$

that constitutes a *futile cycle* (see Figure 10.35). Although such cycles appear to achieve nothing—except dissipate free energy or possibly generate some heat—they may provide regulatory mechanisms if the two reactions are subject to different degrees of activation and inhibition. The nitrogenase futile cycle usually occurs to some extent. Because its operation involves a reduction by means of two electrons, four additional molecules of ATP must be hydrolyzed. If the cycle were to occur with *every* passage of N_2 through the pathway (which generally is not the case), the overall *net* equation for nitrogen fixation would become

$$N_2 + 8e^- + 16ATP^{4-} + 16H_2O \rightarrow 2NH_3 + H_2 + 16ADP^{3-} + 16P_i^{2-} + 8H^+$$

14.1.2. Nitrate Assimilation

Nitrate assimilation constitutes a second major ammonia-forming mechanism in living systems. In contrast to nitrogen fixation, which is limited to nitrogen-fixing bacteria and some plants, nitrate assimilation can be carried out by virtually all plants, fungi, and bacteria. Nitrate assimilation consists of a two-stage reduction of nitrate to ammonia. In the first stage, nitrate undergoes reduction to nitrite:

$$NO_3^- + 2H^+ + 2e^- \rightarrow NO_2^- + H_2O$$

This reaction is catalyzed by **nitrate reductase,** a large (MW $\approx$ 800,000) multisubunit enzyme that contains several electron carriers: FAD, molybdenum, and cytochrome$_{557}$ (contains a [4Fe-4S] complex). The electron donor is NADH in plants and NADPH in fungi and bacteria. The reduction comprises the following flow of electrons:

$$NAD(P)H \rightarrow FAD \rightarrow Cyt_{557} \rightarrow Mo \rightarrow NO_3^-$$

and results in the overall reaction

$$NO_3^- + NAD(P)H + H^+ \rightarrow NO_2^- + NAD(P)^+ + H_2O$$

The second stage of nitrate assimilation consists of three steps that accomplish the reduction of nitrite to ammonia. The same enzyme, **nitrite reductase,** catalyzes all three reactions. Nitrite reductase contains a number of electron carriers, including an iron–sulfur complex and an iron porphyrin, and leads to the following reaction sequence:

$$\begin{aligned} NO_2^- + 2H^+ + 2e^- &\rightarrow NO^- + H_2O \\ NO^- + 3H^+ + 2e^- &\rightarrow NH_2OH \\ NH_2OH + 2H^+ + 2e^- &\rightarrow NH_3 + H_2O \end{aligned}$$

Each step requires two electrons, donated by ferredoxin. NADPH represents the ultimate source of electrons. Three molecules of NADPH are required for every nitrite ion. The overall reaction is

$$NO_2^- + 3NADPH + 4H^+ \rightarrow NH_3 + 3NADP^+ + 2H_2O$$

14.1.3. Ammonia Fixation

Living organisms use three reactions to convert ammonia to organic compounds, which can then can be used in metabolism. In one reaction, catalyzed by **carbamoyl phosphate synthase,** ammonia, CO_2 (as HCO_3^-), and ATP serve as reactants for the synthesis of *carbamoyl phosphate:*

$$NH_4^+ + HCO_3^- + 2ATP^{4-} + H_2O \rightarrow$$
$$H_2N{-}C(=O){-}OPO_3^{2-}$$
$$\text{Carbamoyl phosphate}^{2-} + 2ADP^{3-} + P_i^{2-} + 2H^+$$

Carbamoyl phosphate is an important metabolite, not only because it represents a "fixed" form of ammonia, but also because it constitutes an energy-rich compound owing to its mixed anhydride structure. Synthesis of carbamoyl phosphate requires the expenditure of two molecules of ATP even though only one phosphate group is incorporated into the newly made compound (Figure 14.3).

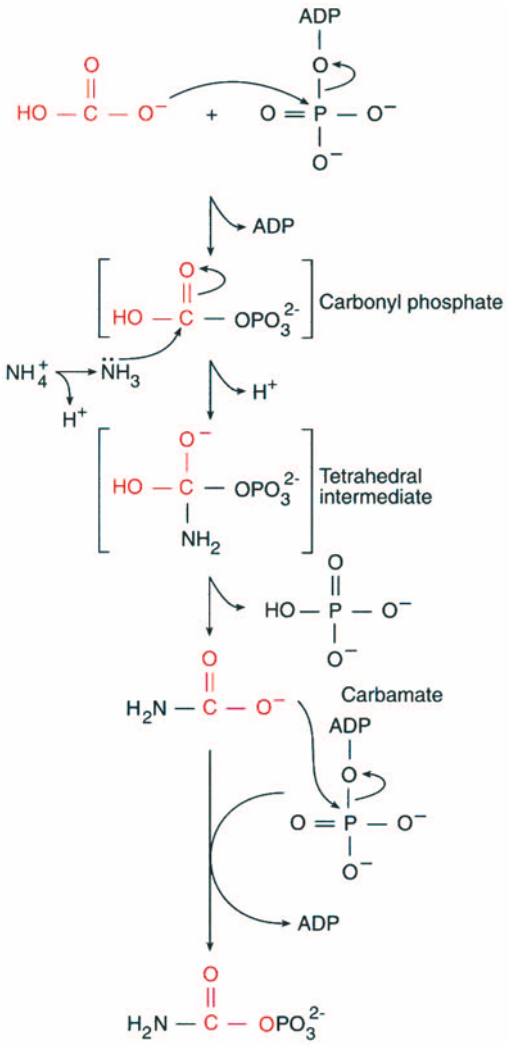


Figure 14.3. Synthesis of carbamoyl phosphate by carbamoyl phosphate synthase I. Activation of bicarbonate to carbonyl phosphate prepares the carbon for nucleophilic attack by ammonia. The mechanism's final step closely resembles the initial step.

Carbamoyl phosphate participates in arginine synthesis in the urea cycle (Section 14.3) and in pyrimidine biosynthesis (Section 14.4). In eukaryotes, synthesis of the carbamoyl phosphate used in the urea cycle is catalyzed by a mitochondrial enzyme (*carbamoyl phosphate synthase I*) that uses ammonia as nitrogen source. Synthesis of the carbamoyl phosphate used in pyrimidine biosynthesis is catalyzed by a different cytosolic enzyme (*carbamoyl phosphate synthase II*) that uses glutamine as nitrogen source. Prokaryotes have only one carbamoyl phosphate synthase, which catalyzes both arginine and pyrimidine synthesis and uses glutamine for both synthetic processes.

A second reaction for converting ammonia to nitrogen-containing compounds is catalyzed by **glutamate dehydrogenase** and reduces α-ketoglutarate to *glutamate.* Glutamate dehydrogenase is widespread in animal, plant, and microbial cells. In plants and animals, the enzyme is located in the mitochondria. Glutamate dehydrogenase of some species uses NADH as coenzyme, and that of other species uses NADPH. Some organisms can use either NADH or NADPH as coenzyme for glutamate dehydrogenase:

$$NH_4^+ + NAD(P)H + H^+ + {}^-OOC{-}C(=O){-}CH_2{-}CH_2{-}COO^- \text{ (α-Ketoglutarate}^{2-}) \rightleftharpoons$$
$$^-OOC{-}CH(NH_3^+){-}CH_2{-}CH_2{-}COO^- \text{ (Glutamate}^-) + NAD(P)^+ + H_2O$$

The third reaction of ammonia fixation involves a conversion of glutamate to *glutamine,* catalyzed by **glutamine synthase:**

$$^-OOC{-}CH(NH_3^+){-}CH_2{-}CH_2{-}COO^- \text{ (Glutamate}^-) + NH_4^+ + ATP^{4-} \rightarrow$$
$$^-OOC{-}CH(NH_3^+){-}CH_2{-}CH_2{-}C(=O){-}NH_2 \text{ (Glutamine)} + ADP^{3-} + P_i^{2-} + H^+$$

Glutamine occupies a pivotal position in nitrogen metabolism since it serves as a precursor for many metabolites. This makes glutamine synthase a critical enzyme. *E. coli* glutamine synthase has been studied extensively. It is an allosteric enzyme, composed of 12 identical subunits, and

subject to multiple regulation by allosteric effectors, feedback inhibition, and covalent modification. The covalent modification consists of an *adenylylation* whereby an AMP group of ATP (splitting out PP_i) becomes linked to a specific tyrosine residue in the enzyme. The bond forms by esterification of the phosphate of AMP and the phenolic hydroxyl of tyrosine. Adenylylation decreases glutamine synthase activity.

The mammalian enzyme differs from the bacterial enzyme in both structure and properties. It consists of eight identical subunits but also exists in the form of tetramers. Brain and liver glutamine synthase have been most thoroughly characterized. The liver enzyme is not subject to covalent modification. Glycine, serine, alanine, and carbamoyl phosphate inhibit the enzyme, and α-ketoglutarate activates it. In mammals, glutamine is synthesized primarily in muscle tissue and transported from there via the blood to other tissues. Glutamine concentration in the blood is about 0.6m*M*, higher than that of any other amino acid.

14.1.4. Essential Amino Acids

Based on their nutritional value, we classify amino acids as essential or nonessential ones (see Table 2.4). *Essential amino acids* cannot be synthesized by an organism or are not synthesized in sufficient quantity and must be obtained through the diet. Essential amino acids in humans represent a good cross section of the different types of amino acids that occur in nature—aromatic, aliphatic, sulfur-containing, and hydroxy amino acids. W. C. Rose identified them in the 1940s, using measurements of **nitrogen balance,** the difference between the amounts of nitrogen ingested and excreted:

$$\text{Nitrogen balance} = \text{N ingested} - \text{N excreted}$$

Normally, an animal's nitrogen balance equals zero; the amount ingested equals that excreted, and the animal is in **nitrogen equilibrium.** In this state, any added nitrogen intake (e.g., a high-protein meal) leads to an equivalent increase in nitrogen excretion so that the difference between intake and excretion remains zero.

A nitrogen balance may, however, be greater or smaller than zero. In growing organisms, dietary nitrogen is continuously assimilated, and muscle tissue is constantly produced. Hence, nitrogen ingestion exceeds nitrogen excretion, resulting in a positive nitrogen balance. Conversely, in organisms suffering from certain wasting diseases, muscle tissue breakdown continues unabated regardless of nitrogen intake. Accordingly, nitrogen excretion exceeds nitrogen ingestion, resulting in a negative nitrogen balance.

To study nitrogen balance, Rose designed a synthetic diet consisting of a mixture of purified amino acids, starch, sucrose, butter, inorganic salts, and vitamins. He put individuals on this diet and controlled the daily amount of nitrogen ingested. He measured daily nitrogen excretion and determined the nitrogen balance as a function of time (Figure 14.4).

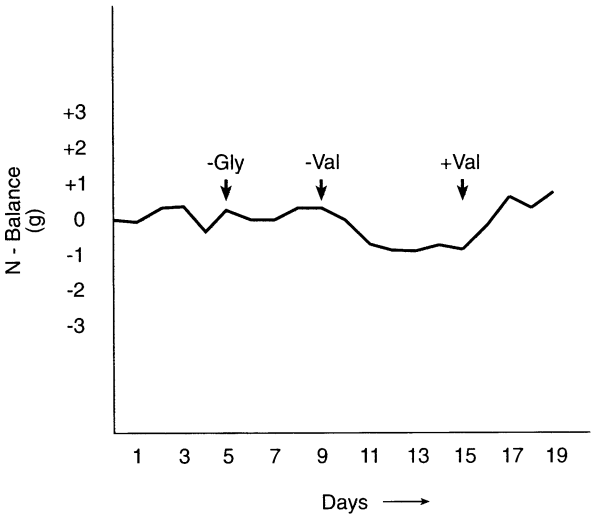


Figure 14.4. Determining nitrogen balance in humans. Omitting an essential amino acid, such as valine, from the diet leads to a negative nitrogen balance. Restoring the amino acid to the diet reestablishes a zero nitrogen balance.

So long as all 20 amino acids were present in the diet, the nitrogen balance was essentially zero. After several days, Rose omitted *one amino acid at a time* from the diet. Omission of some amino acids (e.g., glycine) had no effect on the nitrogen balance, and Rose judged these amino acids to be nonessential. Evidently, the body could compensate for lack of these amino acids in the diet by synthesizing them from other metabolites. Omission of other amino acids (e.g., valine) led to a negative nitrogen balance, and Rose judged these amino acids to be essential. Absence of such amino acids results in impaired protein synthesis but does not decrease protein degradation. Therefore, nitrogen excretion exceeds nitrogen intake, producing a negative nitrogen balance. Restoring an essential amino acid to the diet reestablishes nitrogen equilibrium. (Rose's studies also led to the identification of threonine as one of the 20 amino acids occurring in proteins.)

14.2. PATHWAYS OF AMINO ACID METABOLISM

Amino acid metabolism consists of many different reactions and a variety of pathways. Some pathways are

limited to a few amino acids, whereas others are more general in nature. **Transamination** represents a general pathway that applies to most amino acids and constitutes an important mechanism for interconversion of nitrogen-containing compounds.

14.2.1. Transamination

Transamination reactions are catalyzed by **transaminases** (**aminotransferases**) and involve the transfer of an amino group from an *amino acid* to a *keto acid.* The coenzyme of transaminases contains *vitamin B_6* as a structural component. Vitamin B_6 is widely distributed in nature and occurs in three forms, having an alcoholic ($-CH_2OH$), an amine ($-CH_2NH_2$), or an aldehyde ($-CHO$) group attached to the benzene ring (Figure 14.5). *Pyridoxal phosphate* (*PLP*) and *pyridoxamine phosphate* (*PMP*) constitute phosphorylated forms of vitamin B_6 and serve as coenzymes of transaminases. Thus, the vitamin and the coenzyme differ only by a phosphate group.

Figure 14.6 shows the mechanism of transamination. An incoming amino acid (AA_1) displaces *pyridoxal phosphate,* which is linked to a lysine residue of the enzyme in the form of a *Schiff base* (an *aldimine*). This *transimination* produces another aldimine (also a Schiff base), which tautomerizes to a *ketimine* (also a Schiff base) via a *quinonoid intermediate* (resonance-stabilized carbanion). Hydrolysis of the ketimine yields a keto acid (KA_1), derived from the original amino acid, and *pyridoxamine phosphate.* Thus, AA_1 has been converted to KA_1 while the coenzyme has been altered from the aldehyde to the amine form.

The entire sequence of steps then proceeds in the opposite direction, beginning with an incoming *keto acid* (KA_2) and *pyridoxamine phosphate.* The incoming keto acid differs from that produced in the forward sequence. The reverse sequence yields an *amino acid* (AA_2), derived from the incoming keto acid, and *pyridoxal phosphate.* The entering keto acid has been converted to an amino acid, while the coenzyme has been altered from the amine to the aldehyde form. The reverse sequence regenerates the original form of the coenzyme, pyridoxal phosphate.

Transamination brings about an interconversion of two amino acids and two keto acids. The coenzyme carries out the actual chemical reaction by having its vitamin component shuttle back and forth between the aldehyde and the amine forms.

Transamination reactions exemplify Stage III of nitrogen utilization. In this stage, compounds formed by ammonia fixation interact with intermediates of carbohy-

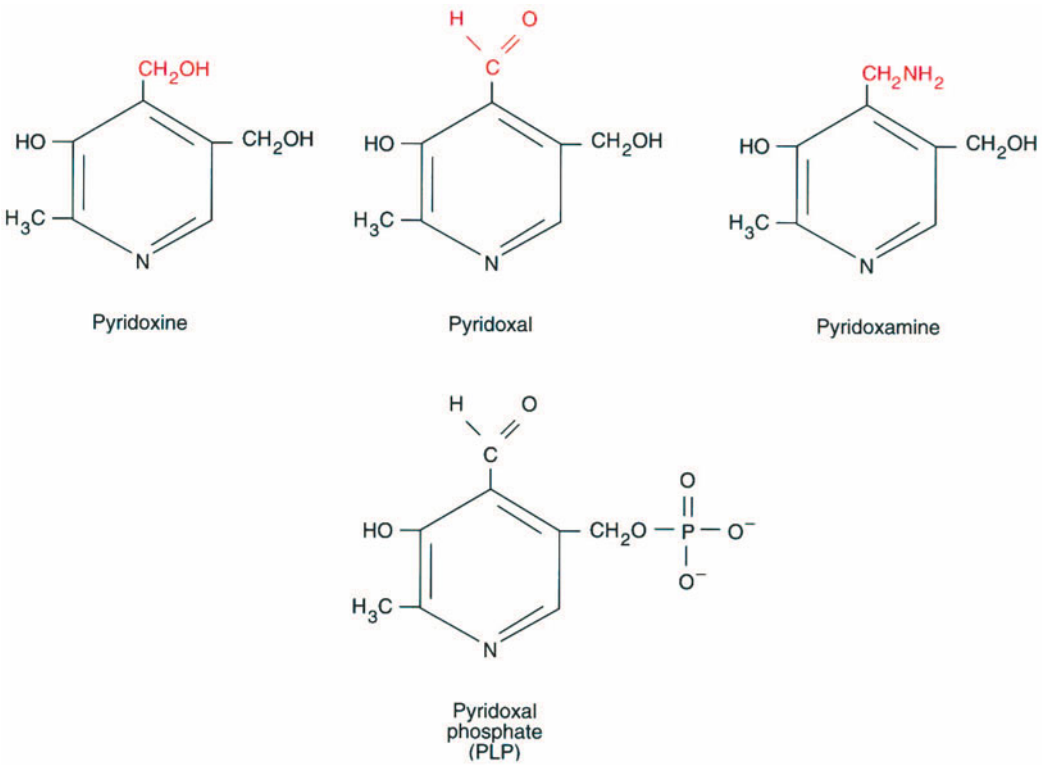


Figure 14.5. Vitamin B_6. The vitamin occurs as an alcohol, an aldehyde, and an amine. Phosphorylated forms are coenzymes.

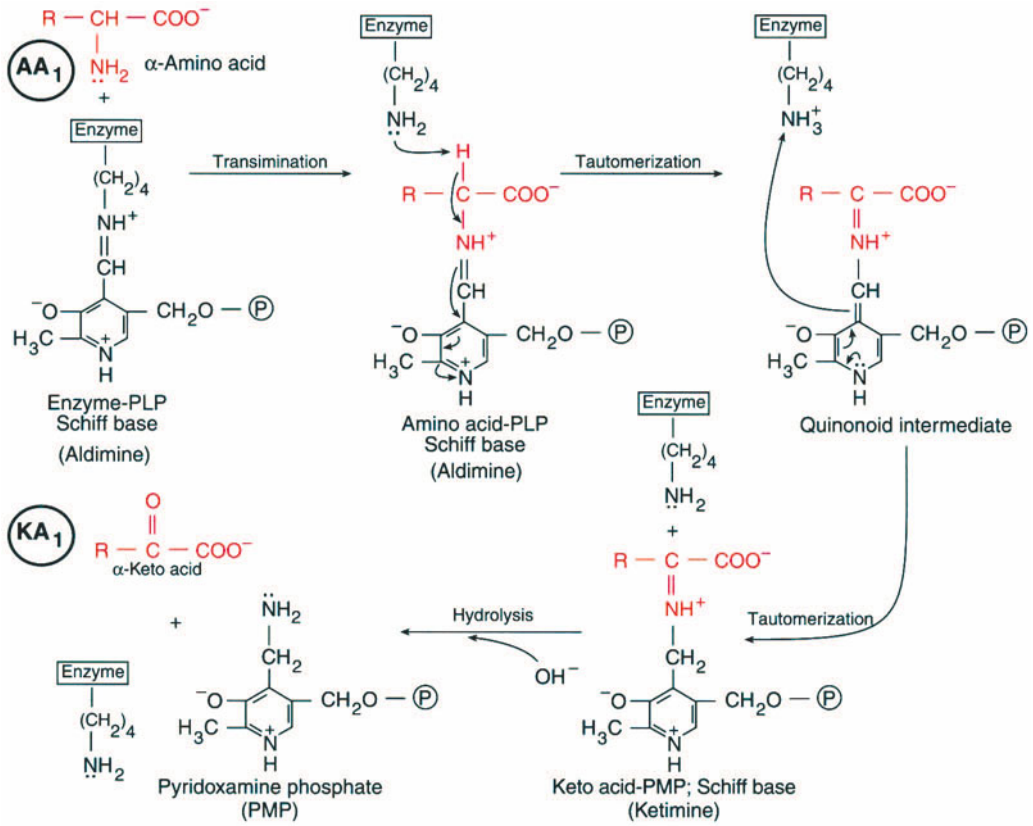


Figure 14.6. The mechanism of transamination. The reaction sequence shown then proceeds in the reverse direction, converting a different keto acid (KA_2) to the corresponding amino acid (AA_2). In the process, PLP is regenerated from PMP. The overall reaction is:

$$\begin{aligned} AA_1 + PLP &\rightleftharpoons KA_1 + PMP \\ KA_2 + PMP &\rightleftharpoons AA_2 + PLP \\ \hline AA_1 + KA_2 &\rightleftharpoons KA_1 + AA_2 \end{aligned}$$

drate and lipid metabolism. Transamination provides a means of using the amino group of glutamate, a key product of ammonia fixation, to form other amino acids by means of pyruvate and oxaloacetate, key intermediates of carbohydrate and lipid metabolism (Figure 14.7). The amino acids formed can subsequently be converted to other nitrogenous compounds. Transamination also provides simple routes for synthesis of nonessential amino acids. Alanine and aspartate are nonessential amino acids because they are readily synthesized via transamination from the common metabolites pyruvate and oxaloacetate. Lastly, transamination has a regulatory function. By varying rates of transamination, organisms can maintain a steady-state level of nonessential amino acids regardless of their concentrations in the diet.

Two transaminases, **serum glutamate-oxaloacetate transaminase** (**SGOT**) and **serum glutamate-pyruvate transaminase** (**SGPT**), were once of great clinical interest since their plasma levels were used to diagnose heart and liver damage. Both the heart and the liver normally contain high concentrations of SGOT and SGPT. In the event of a myocardial infarct (heart attack), a portion of the heart is damaged, causing SGOT and SGPT to leak into the bloodstream from the injured cells. By measuring the activity of these enzymes in the blood, physicians can estimate the extent of heart damage. Liver damage can be assessed similarly since injured liver cells also lead to leakage of SGOT and SGPT into the bloodstream.

In recent times, use of these enzymes has been supplanted by measurements of the isozymes of *lactate dehydrogenase* and *creatine kinase,* which have greater specificity and predictive accuracy. Creatine kinase catalyzes the phosphorylation of creatine to phosphocreatine.

14.2.2. Oxidative Deamination

With few exceptions, the first step in amino acid degradation involves a removal of the amino group by either

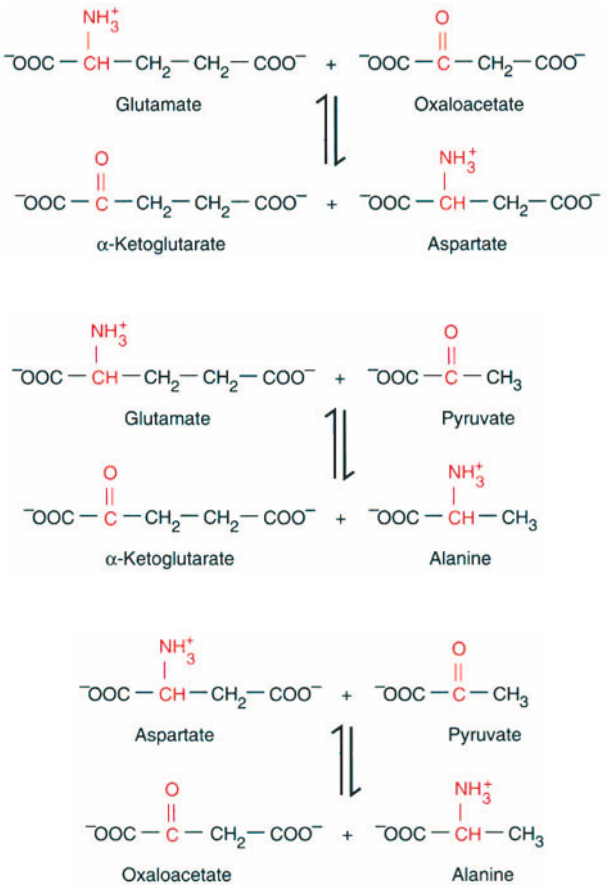


Figure 14.7. Examples of transamination reactions.

transamination or **deamination.** In deamination, the amino group is removed as ammonia. Deamination usually occurs coupled to an oxidation, resulting in conversion of an amino acid to a keto acid. This makes the overall reaction one of **oxidative deamination.** The ammonia enters the urea cycle, where it is converted to urea and excreted. The keto acid enters carbohydrate or lipid metabolism. Oxidative deamination occurs via two mechanisms using either glutamate dehydrogenase or amino acid oxidases.

14.2.2A. Glutamate Dehydrogenase.

Most amino acids can donate their amino groups and form glutamate via transamination. Amino acids undergoing transamination have their amino groups concentrated into glutamate, which can subsequently be oxidatively deaminated at an appreciable rate by glutamate dehydrogenase to provide ammonia for the urea cycle. Both the transaminase and glutamate dehydrogenase reactions are reversible.

Transaminase reaction:

$$^{-}OOC-\underset{\alpha\text{-Ketoglutarate}}{\overset{O}{\overset{\|}{C}}-CH_2-CH_2-COO^-} + \underset{\text{amino acid}}{R-\overset{NH_3^+}{\overset{|}{CH}}-COO^-} \rightleftharpoons \underset{\text{Glutamate}}{^{-}OOC-\overset{NH_3^+}{\overset{|}{CH}}-CH_2-CH_2-COO^-} + \underset{\text{keto acid}}{R-\overset{O}{\overset{\|}{C}}-COO^-}$$

Glutamate dehydrogenase reaction:

$$\underset{\text{Glutamate}}{^{-}OOC-\overset{NH_3^+}{\overset{|}{CH}}-CH_2-CH_2-COO^-} + NAD^+ + H_2O \rightleftharpoons \underset{\alpha\text{-Ketoglutarate}}{^{-}OOC-\overset{O}{\overset{\|}{C}}-CH_2-CH_2-COO^-} + NADH + H^+ + NH_4^+$$

The overall reaction is:

$$\text{Amino acid} + NAD^+ + H_2O \rightleftharpoons \text{keto acid} + NADH + H^+ + NH_4^+$$

14.2.2B. Amino Acid Oxidases.

Oxidative deamination of amino acids when catalyzed by amino acid oxidases proceeds in a single step. Amino acid oxidases occur predominantly in the liver and the kidney. *L-Amino acid oxidases* are specific for L-amino acids and have FMN as coenzyme. *D-Amino acid oxidases* are specific for D-amino acids and have FAD as coenzyme. The function of D-amino acid oxidases is enigmatic since D-amino acids are rare in animals, being associated primarily with bacterial cell walls. The general reaction with an amino acid oxidase takes the form:

$$\underset{\text{Amino acid}}{R-\overset{NH_3^+}{\overset{|}{CH}}-COO^-} + \underset{(FAD)}{FMN} + H_2O \rightarrow \underset{\text{Keto acid}}{R-\overset{O}{\overset{\|}{C}}-COO^-} + \underset{(FADH_2)}{FMNH_2} + NH_4^+$$

Amino acid oxidases also catalyze reoxidation of the reduced coenzyme $FMNH_2$ ($FADH_2$) by oxygen. Recall that any oxidase catalyzes the direct interaction of a substrate with molecular oxygen:

$$\underset{(FADH_2)}{FMNH_2} + O_2 \rightarrow \underset{(FAD)}{FMN} + H_2O_2$$

The toxic hydrogen peroxide (H_2O_2) formed in this reaction is decomposed by action of the enzyme *catalase:*

$$2H_2O_2 \longrightarrow 2H_2O + O_2$$

Table 14.1. Glucogenic and Ketogenic Amino Acids

Amino acid	Glucogenic	Ketogenic
Alanine	×	
Arginine	×	
Asparagine	×	
Aspartate	×	
Cysteine	×	
Glutamate	×	
Glutamine	×	
Glycine	×	
Histidine	×	
Isoleucine	×	×
Leucine		×
Lysine		×
Methionine	×	
Phenylalanine	×	×
Proline	×	
Serine	×	
Threonine	×	×
Tryptophan	×	×
Tyrosine	×	×
Valine	×	

14.2.3. Metabolic Fates of Amino Acids

Following oxidative deamination, carbon skeletons of all amino acids undergo degradation to one of seven common intermediates. Based on these end products of their catabolism, we can place amino acids in two large groups (Table 14.1). **Glucogenic (glycogenic) amino acids** are catabolized to *precursors of carbohydrates.* Catabolism of these amino acids yields pyruvate, oxaloacetate, α-ketoglutarate, succinyl CoA, and fumarate—all compounds that lead to synthesis of glucose and glycogen. **Ketogenic amino acids** are catabolized to *precursors of lipids.* Catabolism of these amino acids yields acetyl CoA and acetoacetyl CoA—compounds that lead to synthesis of ketone bodies. Most of the amino acids (13) are strictly glucogenic; leucine and lysine represent the only amino acids that are strictly ketogenic. Five amino acids turn out to be both glucogenic and ketogenic; their degradation yields two products, one of which is glucogenic, and the other ketogenic.

Both amino acid catabolism and anabolism are intimately linked to operation of the citric acid cycle (Figure

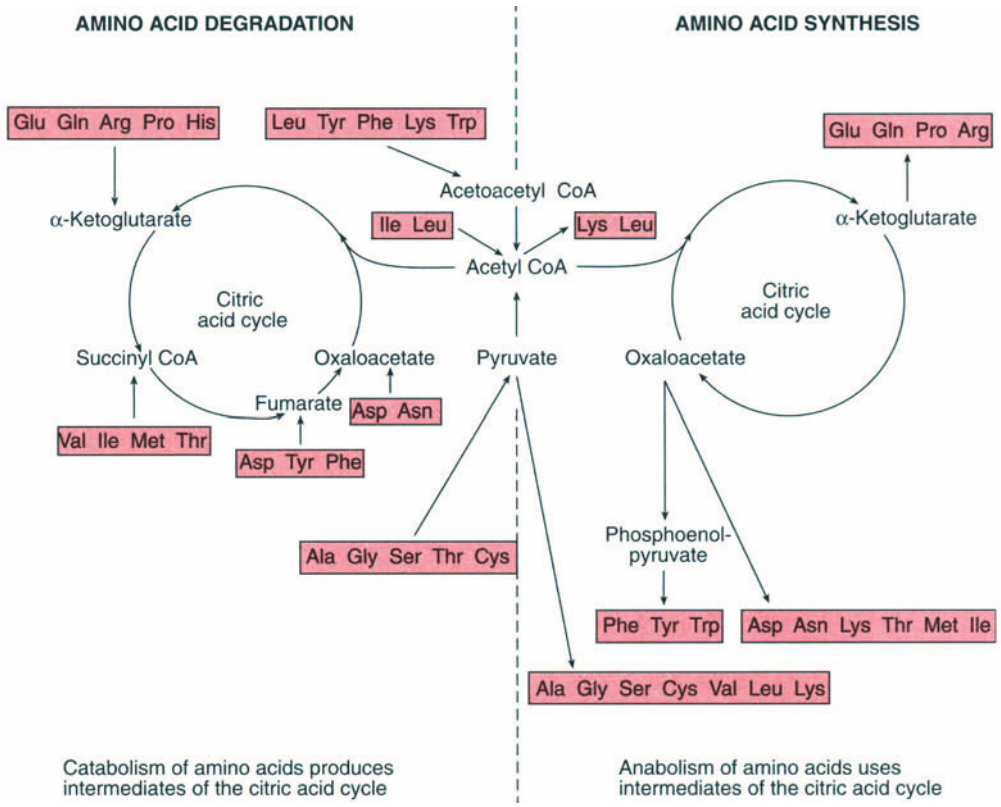


Figure 14.8. The relationships between amino acid metabolism and the citric acid cycle.

14.8). When degraded, amino acids yield intermediates of the cycle or precursors of such intermediates. When synthesized, amino acids form from intermediates of the cycle or from precursors of such intermediates.

14.2.4. Catabolism of Phenylalanine and Tyrosine

As an illustration of amino acid degradation, consider the specific catabolic pathways of phenylalanine and tyrosine (Figure 14.9). The major catabolic pathway of these two aromatic amino acids proceeds from phenylalanine through tyrosine to fumarate and acetoacetate. Because fumarate is glucogenic whereas acetoacetate is ketogenic, we classify phenylalanine and tyrosine as being both glucogenic and ketogenic. As Figure 14.9 indicates, tyrosine is normally synthesized by hydroxylation of phenylalanine, a reaction catalyzed by *phenylalanine hydroxylase*. Because of this reaction, we consider tyrosine to be a nonessential amino acid. The first step in tyrosine catabolism, as in the catabolism of most amino acids, is an oxidative deamination.

Phenylalanine hydroxylase catalyzes the conversion of phenylalanine to tyrosine in a single step. The enzyme requires molecular oxygen and the cofactor *tetrahydrobiopterin.* Phenylalanine hydroxylase is a *monooxygenase,* or *mixed-function oxygenase,* since one atom of O_2 appears in the product (tyrosine) and the other is converted to water (Figure 14.10). The quinonoid form of dihydrobiopterin produced in the hydroxylation of phenylalanine is reduced back to tetrahydrobiopterin by NADH in a reaction catalyzed by *dihydropteridine reductase.*

In addition to initiating the major catabolic route, phenylalanine can undergo a second set of reactions. This pathway functions much as the "overflow pathway" of ketone body formation. Normally of minor significance, the pathway becomes accentuated when phenylalanine accumulates because of a block in the major pathway. In some individuals such a block occurs at the phenylalanine hydroxylase step.

We refer to the genetic disease associated with this metabolic defect as **phenylketonuria,** or **PKU.** It is caused by an autosomal recessive gene. In phenylketonuria, phenylalanine cannot be hydroxylated to tyrosine and is, instead, converted to phenylpyruvate and phenyllactate (Figure 14.9). The first step of this catabolic pathway also involves an oxidative deamination. The PKU pathway leads to accumulation of large quantities of phenylpyruvate and phenyllactate in the blood, followed

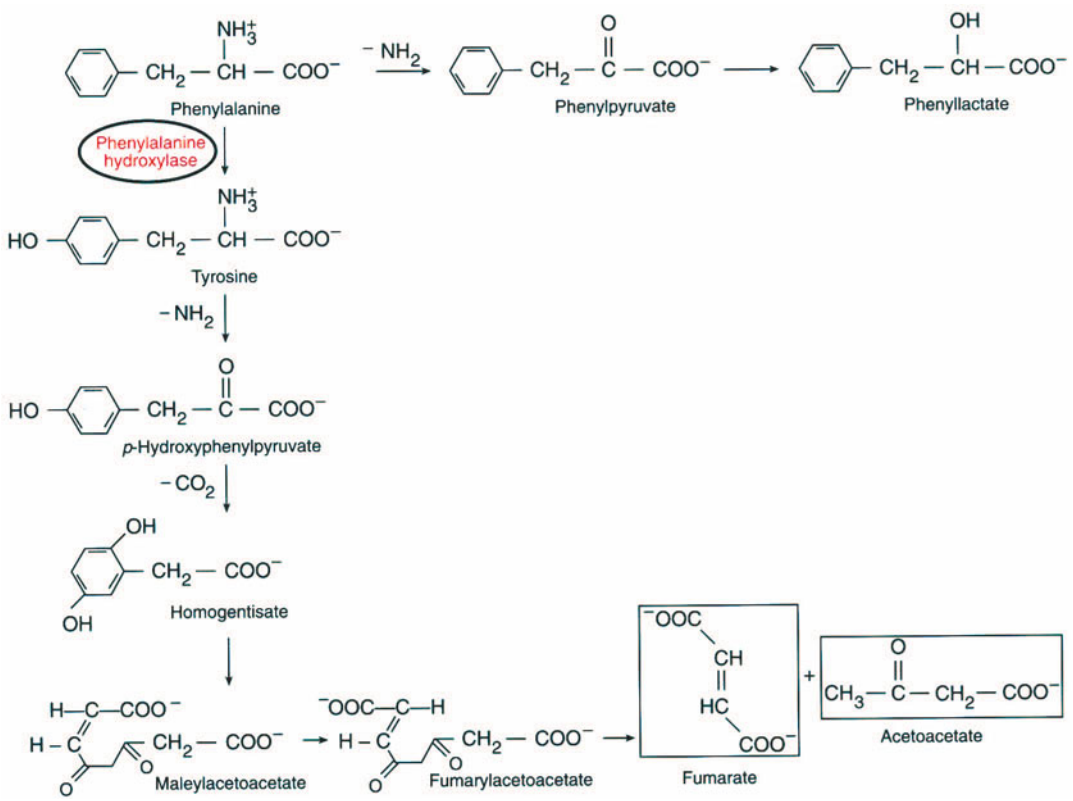


Figure 14.9. Catabolism of phenylalanine and tyrosine. The major pathway leads to fumarate and acetoacetate.

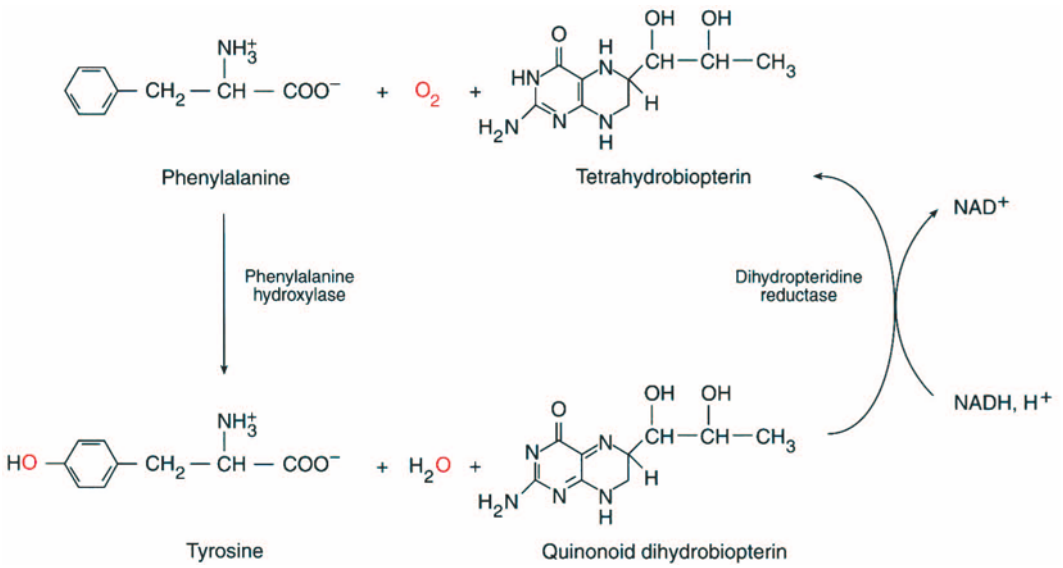


Figure 14.10. Tyrosine conversion to phenylalanine, catalyzed by phenylalanine hydroxylase. The reaction involves an oxidation of tetrahydrobiopterin, which is then reduced by means of NADH in a reaction catalyzed by dihydropteridine reductase.

by their excretion in the urine. Physicians consider a concentration of over 20 mg of phenylalanine per 100 ml of blood to be a positive indicator for PKU.

Someone suffering from PKU in early childhood becomes mentally retarded. Researchers generally attribute mental retardation to toxic effects of phenylalanine, possibly on the transport of other aromatic amino acids in the brain. Because of the severity of the disease, hospitals today routinely check children at birth for PKU.

When laboratory analysis positively identifies the condition, the child is placed on a synthetic diet low in phenylalanine. This regimen is continued at least through early childhood (up to the age of 5). Following that, protein is restricted in the diet for three to five more years. At that age, the deleterious effects of PKU seem to disappear since, by the time a child reaches age 8, a large portion of the brain has been formed. However, whether and when phenylalanine can be added to the diet at normal levels is still a controversial subject.

14.2.5. Amino Acids as Biosynthetic Precursors

In addition to their major function as building blocks of proteins, amino acids serve as precursors for a variety of biomolecules. Glycine, for example, functions as a building block of heme. The entire tetrapyrrole ring system is assembled from glycine and acetate (Figure 14.11). Heme degradation entails an opening of the ring system and its conversion to linear tetrapyrroles termed *bile pigments.* Tyrosine serves as a precursor of *melanins,* dark-colored substances responsible for skin pigmentation.

A number of physiologically active amines, including epinephrine (Figure 10.3), histamine (Figure 14.12), and serotonin (Figure 14.12), are derivatives of amino acids. Epinephrine is a hormone derived from tyrosine that functions in regulating glycogen metabolism and fatty acid mobilization. Histamine derives from histidine and has multiple functions. It acts as a strong vasodilator and is released during inflammation and allergic reactions. It also promotes increased capillary permeability and contraction of smooth muscle and enhances secretion of hydrochloric acid in the stomach.

Serotonin derives from tryptophan. It is a neurotransmitter and acts as a potent vasoconstrictor. It helps to control blood pressure and regulate peristalsis in the small intestine. A reaction common to the synthesis of epinephrine, histamine, and serotonin consists of a pyridoxal phosphate-dependent decarboxylation of the amino acid from which each derives.

Glutathione is a tripeptide composed of glutamic acid, cysteine, and glycine (see Figure 2.12) that functions in oxidation–reduction reactions of sulfhydryl groups. Glutamic acid residues also occur as structural components of folate coenzymes.

14.2.6. Biosynthesis of Amino Acids

Biosynthesis of amino acids proceeds from simple metabolic precursors (see Figure 14.8). Based on the relationships among their biosynthetic pathways, we group amino

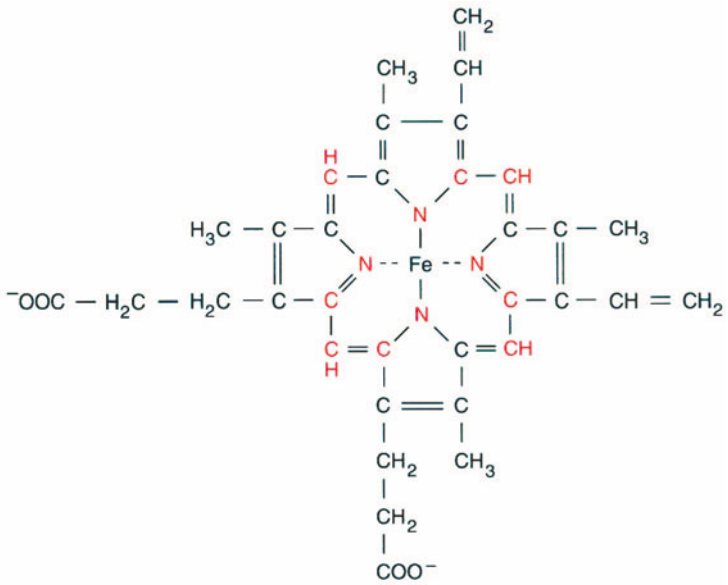
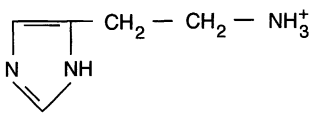


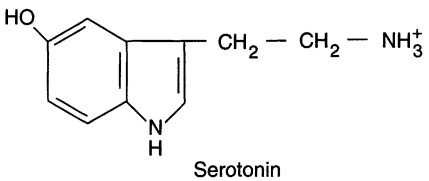
Figure 14.11. Assembly of the heme nucleus. Except for the iron, all atoms are derived from glycine (red) and acetate (black).

acids into six families (Figure 14.13). Some nonessential amino acids are produced directly from intermediates of carbohydrate and lipid metabolism, and some are formed by way of other amino acids. Hydroxylation of phenylalanine yields tyrosine. Serine serves as a precursor of glycine and cysteine, and glutamate serves as a precursor of proline and arginine. Aspartate and glutamate are converted to asparagine and glutamine, respectively, by *asparagine synthase* and *glutamine synthase*. Glutamine synthase, an allosteric enzyme, occupies a key position in regulating nitrogen metabolism. Glutamine represents a storage form of ammonia and functions as an amino group donor in many biosynthetic reactions.

Pathways for the biosynthesis of essential amino acids occur only in plants and microorganisms and usually require a larger number of steps than pathways for the synthesis of nonessential amino acids. Histidine forms from ribose 5-phosphate by way of phosphoribosyl pyrophosphate.



Histamine



Serotonin

Figure 14.12. Some physiologically active amines derived from amino acids.

14.3. UREA CYCLE

The **urea cycle** constitutes a general pathway of amino acid catabolism whereby amino groups of amino acids are converted to urea that can be excreted. The cycle was elucidated by Hans Krebs and Kurt Henseleit in 1932, several years before Krebs discovered the citric acid cycle. The urea cycle was the first known metabolic cycle.

14.3.1. Operation of the Cycle

14.3.1A. Individual Reactions. The urea cycle consists of five enzymatic reactions, three of which occur in the cytosol and two of which take place in the mitochondrial matrix (Figure 14.14). Specific carrier mechanisms, located in the inner mitochondrial membrane, provide for transport of ornithine and citrulline into and out of the matrix. The reactions of the cycle are:

1. $NH_4^+ + HCO_3^- + 2ATP^{4-} + H_2O \rightarrow$ carbamoyl phosphate^{2-} + $2ADP^{3-} + P_i^{2-} + 2H^+$
2. Carbamoyl phosphate^{2-} + ornithine$^+ \rightarrow$ citrulline + $P_i^{2-} + H^+$
3. Citrulline + aspartate$^-$ + $ATP^{4-} \rightarrow$ $AMP^{2-} + PP_i^{3-}$ + argininosuccinate$^-$ + H^+

4. $Argininosuccinate^- + H_2O \rightarrow arginine^+ + fumarate^{2-}$
5. $Arginine^+ + H_2O \rightarrow ornithine^+ + urea$

Reaction 1 represents one of the ammonia fixation reactions in which ammonia from oxidative deamination of amino acids gives rise to **carbamoyl phosphate.** The reaction requires expenditure of energy in the form of ATP, which is cleaved to ADP and P_i. Carbamoyl phosphate formation constitutes the *committed step* of the urea cycle and its major control point. The reaction is catalyzed by *carbamoyl phosphate synthase,* an allosteric enzyme. Recall that in eukaryotes, this reaction is catalyzed by *carbamoyl phosphate synthase I* (see Section 14.1).

Carbamoyl phosphate reacts with **ornithine** to yield **citrulline** (reaction 2). Both ornithine and citrulline are α-amino acids, but do not occur as structural components of proteins. In reaction 3, *aspartate* reacts with citrulline, and ATP is cleaved to AMP and PP_i. Because the reaction product can be considered to be an addition compound of arginine and succinate, we call it **argininosuccinate.**

Reaction 4 involves a cleavage of argininosuccinate, yielding *arginine* and *fumarate.* **Arginase** catalyzes the last step of the cycle (reaction 5), which results in a cleavage of arginine to **urea** ($H_2N-CO-NH_2$) and ornithine. In mammals, urea synthesis occurs exclusively in the liver. Urea is then transported by the blood from the liver to the kidney, where it is excreted in the urine.

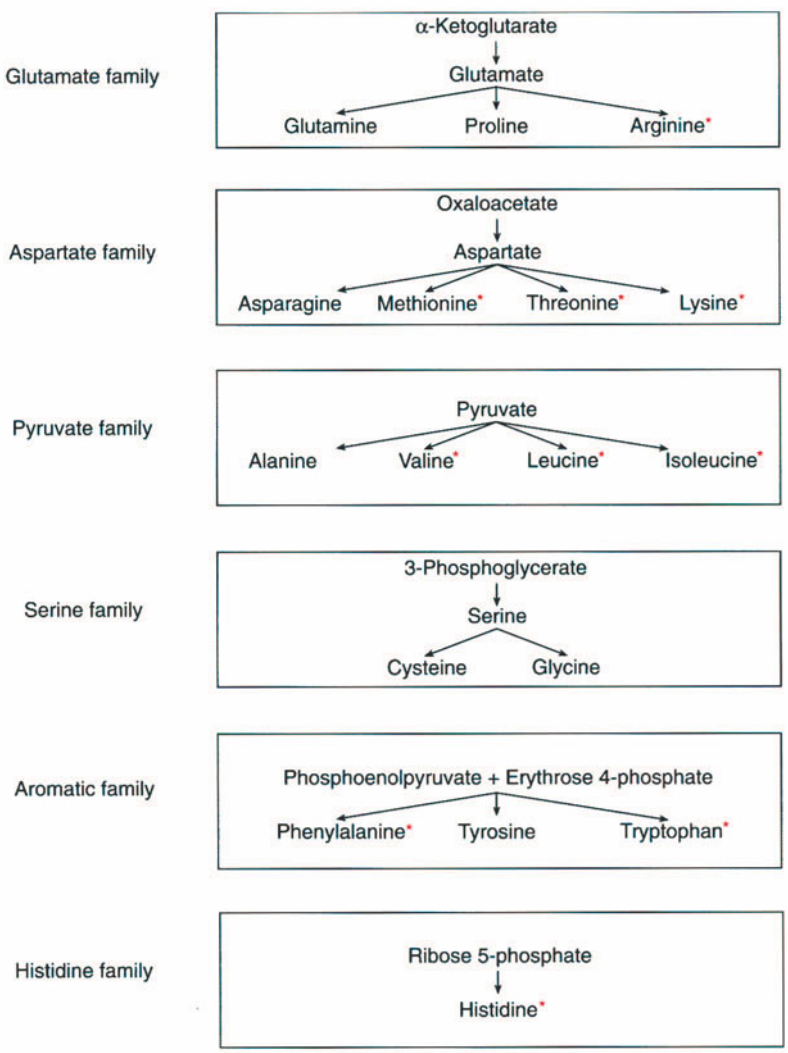


Figure 14.13. Families of amino acids based on their biosynthetic pathways (* = essential amino acid).

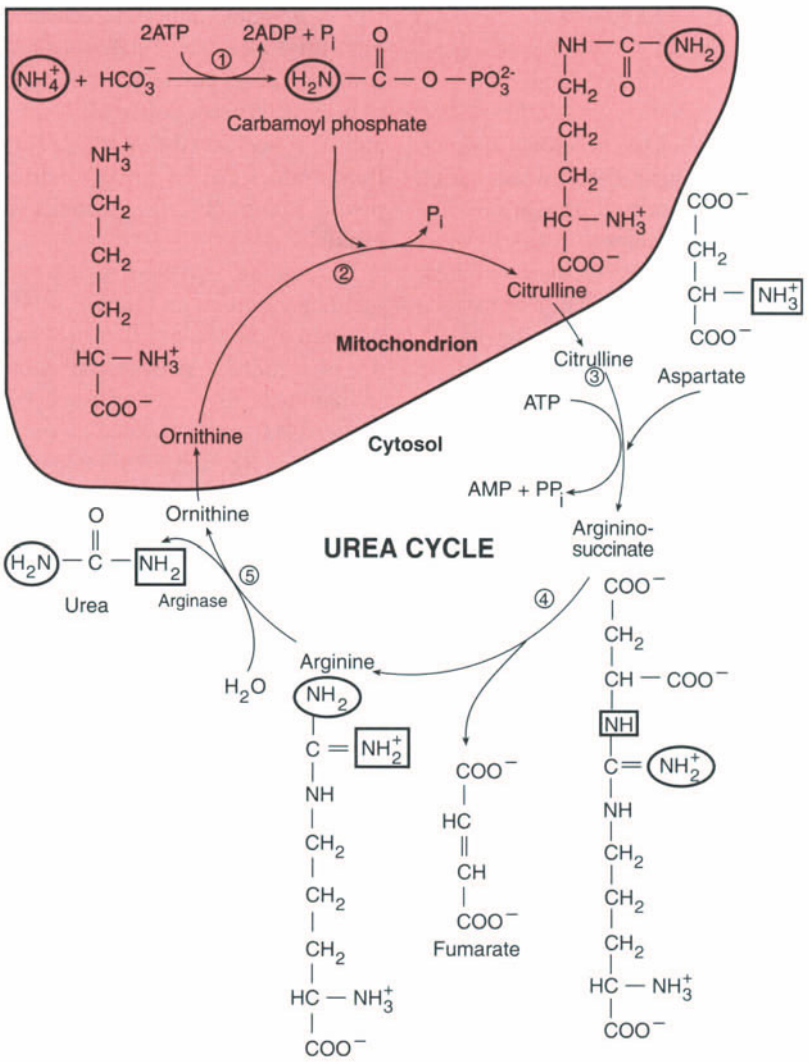


Figure 14.14. The urea cycle. Numbers designate enzymes: ① Carbamoyl phosphate synthase I; ② ornithine transcarbamoylase; ③ argininosuccinate synthetase; ④ argininosuccinase; ⑤ arginase.

14.3.1B. Overall Reaction. By adding the above five reactions, you obtain the overall reaction of the urea cycle:

$$NH_4^+ + HCO_3^- + 3ATP^{4-} + 2H_2O + \text{aspartate}^-$$
$$\downarrow$$
$$\text{Urea} + \text{fumarate}^{2-} + 2ADP^{3-} + AMP^{2-} + PP_i^{3-} + 2P_i^{2-} + 4H^+$$

One turn of the cycle achieves the synthesis of one molecule of urea and requires the input of *one carbon atom and two amino groups*. The carbon derives from CO_2 in reaction 1, and the two amino groups derive from two amino acids. One amino group comes from aspartate, which is thereby converted to fumarate (reactions 3 and 4). The second amino group (reaction 1) comes from deamination of some other amino acid, primarily glutamate via the glutamate dehydrogenase reaction.

14.3.1C. Energetics. Although the urea cycle constitutes a catabolic pathway, it does not yield usable energy. On the contrary, this pathway is endergonic and requires the input of energy. Between them, reactions 1 and 3 require three molecules of ATP. Additionally, as is common in metabolism, when pyrophosphate forms, it undergoes hydrolysis to inorganic phosphate, an exergonic reaction catalyzed by *pyrophosphatase*. Such hydrolysis always helps to drive the overall process that includes pyrophosphate formation. Allowing for pyrophosphate hydrolysis, the energy requirement for urea

synthesis becomes equivalent to the hydrolysis of four energy-rich bonds:

Reaction	Energy-rich bonds
$2ATP \rightarrow 2ADP + 2P_i$	2
$ATP \rightarrow AMP + PP_i$	1
$PP_i \rightarrow 2P_i$	1
Total	4

14.3.1D. Metabolic Interrelationships.

Operation of the urea cycle illustrates the interconnections of metabolic pathways (Figure 14.15). The substrates for reaction 1 derive from deamination (NH_3) and from oxidations via the citric acid cycle and the electron transport system (CO_2, H_2O, ATP). Reaction 1 "fixes" ammonia into the organic compound carbamoyl phosphate. The aspartate in reaction 3 comes from transamination of oxaloacetate, a citric acid cycle intermediate. Fumarate, the product of reaction 4, enters the citric acid cycle and is converted back to oxaloacetate. Conversion of fumarate to oxaloacetate generates NADH, which can be oxidized via the electron transport system to produce ATP and thereby furnish energy for operation of the urea cycle. You can see that the cycle has many links to major aspects of metabolism.

14.3.2. Comparative Biochemistry of Nitrogen Excretion

Because ammonia is highly toxic for living cells, its concentration is generally maintained at low levels. We do not fully understand the reasons for ammonia's great toxicity, but we know of two important contributing factors—the high pK'_a of ammonia (9.3 at 25°C) and the free movement of NH_3 across cell and mitochondrial membranes.

Because of its high pK'_a only a small fraction of ammonia exists as NH_3 at physiological pH (7.0); most occurs in the form of the ammonium ion, NH_4^+. However, because of its great permeation capacity, NH_3 passes readily across the membranes of brain cells and the membranes of their mitochondria. Once inside the mitochondria, ammonia reacts with α-ketoglutarate via the glutamate dehydrogenase reaction. Conversion of α-ketoglutarate to glutamate lowers the α-ketoglutarate concentration in the citric acid cycle pool, thereby lowering the cycle's activity. Decreased citric acid cycle activity translates into decreased oxidation of glucose, the principal fuel of the brain.

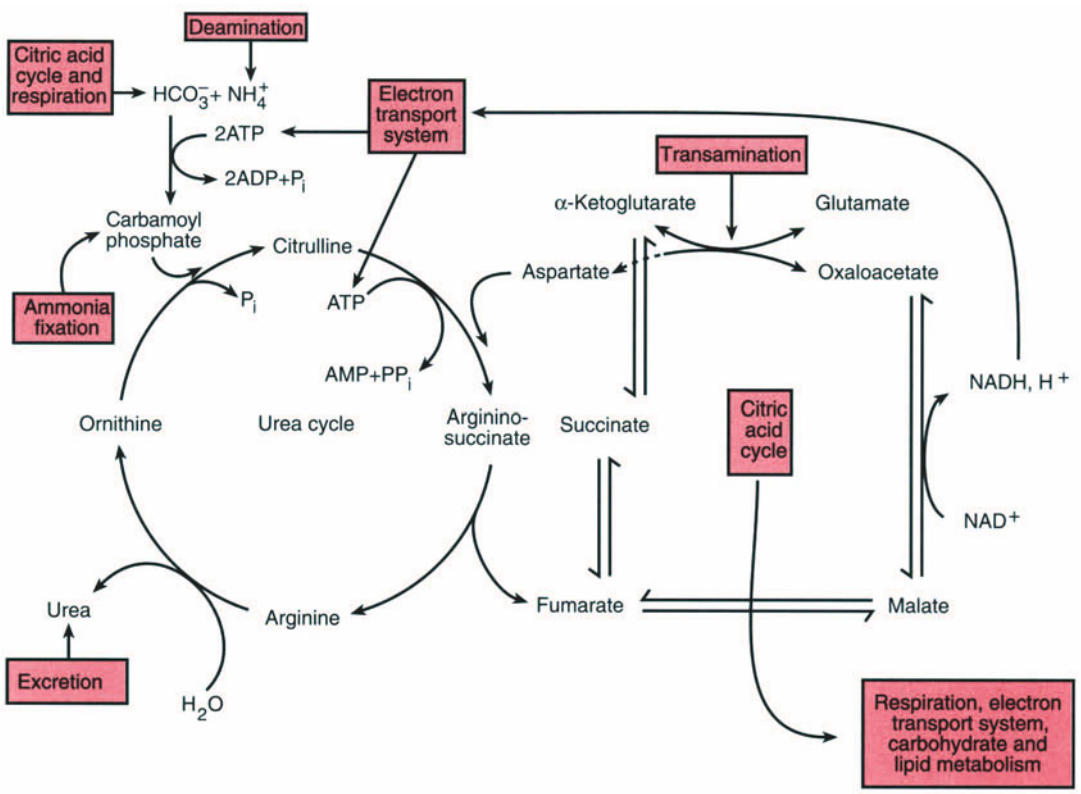


Figure 14.15. Linkage of the urea cycle with other metabolic systems.

Excess ammonia produced by amino acid catabolism must be excreted. The form in which amino acid nitrogen is excreted by living organisms varies. Some organisms excrete ammonia directly; others excrete waste nitrogen in the form of urea or uric acid. We may classify living organisms according to the form in which they excrete their nitrogen. **Ureotelic, uricotelic,** or **ammonotelic organisms** excrete waste nitrogen primarily in the form of *urea, uric acid,* or *ammonia,* respectively. The diversity of nitrogen excretion represents an interesting case of comparative biochemistry. It relates to water intake and availability and to solubility and toxicity of the three excretory nitrogen-containing compounds.

Ammonia represents the most toxic but also the most soluble of the three compounds. Aquatic animals and teleost fish have an unlimited supply of water. These organisms excrete ammonia despite its great toxicity. Because of its great solubility, ammonia becomes effectively diluted by the environment as soon as it is excreted and is thereby rendered harmless.

Birds and land-dwelling reptiles, whose water intake is limited and whose excretion consists largely of semisolid material, excrete nitrogen as uric acid. Uric acid is less toxic than ammonia and is the least soluble of the three compounds. Its low toxicity results in part from its low solubility in water.

Mammals and most terrestrial vertebrates fall somewhere in between as regards their water supply. These organisms excrete nitrogen in the form of urea, a nontoxic compound. Urea is intermediate in its solubility, being less soluble than ammonia but more soluble than uric acid.

Researchers have suggested that the "choice" between excretion of urea and uric acid depends on the conditions under which an embryo develops. In the case of mammals, the early embryonic stage takes place in an *internal environment* and in close contact with the circulatory system of the mother. Urea, a nontoxic compound of good water solubility, can safely be formed by the embryo for ready removal by the maternal circulation.

The early embryonic stage of birds and reptiles, however, takes place in an *external environment* and in an egg surrounded by a hard shell. The egg contains sufficient water to allow hatching but not enough to permit excretion of large quantities of toxic substances. Excretion of ammonia by these organisms would be fatal, and they have evolved to eliminate nitrogen as uric acid, which precipitates out in a fluid-filled sac (amnion) located on the interior surface of the shell.

Scientists believe that these metabolic systems, necessary for development of the embryo, are then carried over to the adult organism. Support for this theory comes from a number of living systems. Amphibia fall in between ureotelic and ammonotelic organisms. The tadpole lives in an aqueous environment and excretes nitrogen as ammonia. However, during metamorphosis, its liver acquires the necessary enzymes for urea synthesis, and the organism begins to excrete urea. By the time that metamorphosis has been completed and the adult frog has formed, nitrogen excretion occurs predominantly in the form of urea.

Nitrogen excretion of the lungfish also varies. So long as the lungfish can exist in plenty of water, its nitrogen excretion consists primarily of ammonia. When the riverbeds and lakes become dry, the fish settles down in the mud and begins to accumulate nitrogen as urea. When the rainy season returns to fill the rivers and lakes, the lungfish first eliminates a large amount of accumulated urea and then reverts to excreting nitrogen as ammonia.

Lastly, the order Chelonia (tortoises and turtles) includes some species that are strictly aquatic, some that are strictly terrestrial, and some that are semiterrestrial. The aquatic species excrete a mixture of urea and ammonia, the terrestrial ones excrete predominantly uric acid, and the semiterrestrial organisms excrete urea.

14.4. PURINE AND PYRIMIDINE METABOLISM

Purines and pyrimidines occur widely in cells in the form of nucleotides and nucleic acids. We will discuss nucleotide metabolism in this chapter and that of nucleic acids in later chapters. In addition to being *building blocks* of nucleic acids, nucleotides function in many other capacities in living cells. Nucleoside triphosphates, such as ATP and GTP, represent *energy-rich compounds* that drive metabolic reactions. Nucleotides form parts of *coenzymes,* like NAD^+, $NADP^+$, FAD, FMN, and coenzyme A. These coenzymes participate in many metabolic systems and pathways, such as the electron transport system, the citric acid cycle, and β-oxidation. Nucleotides constitute versatile *"carriers"* of metabolites. AMP serves as a carrier of methionine (*S*-adenosylmethionine), UDP as a carrier of glucose (UDP-glucose), and CDP as a carrier of choline (CDP-choline). Lastly, nucleotides play an important role in the *regulation* of metabolism. The levels of ATP, ADP, and AMP, the levels of NAD^+ and NADH, and the mediation of hormonal signals by means of cAMP all function to control the network of metabolic reactions.

14.4.1. Purine Biosynthesis

14.4.1A. Synthesis of IMP. Atoms of the purine ring come from five sources: formate, glycine, aspartate (amino group), glutamine (amide group), and CO_2 (Figure 14.16). **Inosine 5′-monophosphate (IMP)** serves

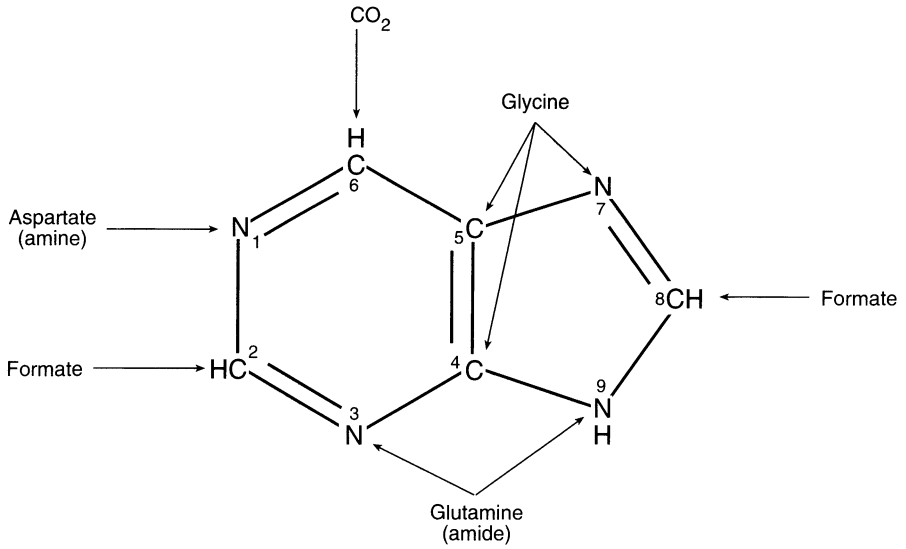


Figure 14.16. Assembly of the purine nucleus.

as the common precursor of all purine nucleotides. It is synthesized *de novo* ("anew," from simple precursors) from ribose 5-phosphate by a series of 11 reactions (Figure 14.17). Assembly of the purine ring takes place on the sugar phosphate, which serves as an anchor.

Synthesis of IMP begins with a conversion of ribose 5-phosphate to 5-phospho-α-D-ribosyl-1-pyrophosphate, or **phosphoribosyl pyrophosphate** (**PRPP**). The reaction is catalyzed by **PRPP synthase,** an unusual kinase that catalyzes the transfer of a pyrophosphate group, rather than a phosphate group, from ATP. PRPP synthase occupies a key position in purine metabolism and is subject to control via a number of activators and inhibitors. PRPP also serves as a precursor in the biosynthesis of pyrimidines.

The second step of purine biosynthesis represents the committed step of the pathway. The reaction, catalyzed by **glutamine-PRPP amidotransferase,** involves displacement of the pyrophosphate group of PRPP by the amide nitrogen of glutamine. The proper anomeric (β) configuration of nucleotides becomes established at this step. Pyrophosphatase-catalyzed hydrolysis of the PP_i produced drives the reaction to completion. Glutamine-PRPP amidotransferase is subject to feedback inhibition by purine nucleotides.

Among the subsequent steps of IMP synthesis occur two that require **tetrahydrofolate** (**THF**), the coenzyme form of *folic acid* (Figure 14.18). THF functions in the metabolism of *one-carbon fragments.* This area of metabolism includes decarboxylations, methylations, and transfers of other groups containing only one carbon atom. Biotin (see Section 13.5) serves as coenzyme for most carboxylation reactions, and *S*-adenosylmethionine (see Section 13.1) serves as coenzyme for most methylation reactions. THF serves as coenzyme for other one-carbon fragment reactions. THF is a versatile coenzyme that can carry one-carbon fragments in various oxidation states, except the most oxidized state (CO_2). As such, THF can serve as a carrier for methyl ($-CH_3$), methylene ($-CH_2-$), formyl ($-CH=O$), formimino ($-CH=NH$), and methenyl ($-CH=$) groups. Figure 14.18 illustrates how THF carries some of these groups.

14.4.1B. Interconversions of Purine Nucleotides. Inosine monophosphate does not accumulate in cells because it is rapidly converted to AMP or GMP. Each conversion involves a two-step mechanism. Synthesis of AMP proceeds via adenylosuccinate, and synthesis of GMP proceeds via xanthosine monophosphate. All other purine nucleotides are derived from AMP and GMP (Figure 14.19).

Regulation of purine biosynthesis occurs via multiple feedback inhibitions (Figure 14.20). Both AMP and GMP serve as negative allosteric effectors for the enzymes at steps 1 and 2 of the purine biosynthetic pathway. In addition, AMP inhibits the first step of its synthesis from IMP, and GMP inhibits the first step of its synthesis from IMP.

Both purine and pyrimidine nucleotides are also synthesized by means of **salvage pathways.** A salvage pathway uses catabolic compounds for biosynthetic purposes even though they do not constitute true intermediates of the corresponding normal biosynthetic pathway. Purine and pyrimidine salvage pathways use bases obtained from nucleotide and nucleoside catabolism for biosynthesis of nucleosides and nucleotides. Salvage pathways constitute "recycling" systems in that they allow for reuse of bases

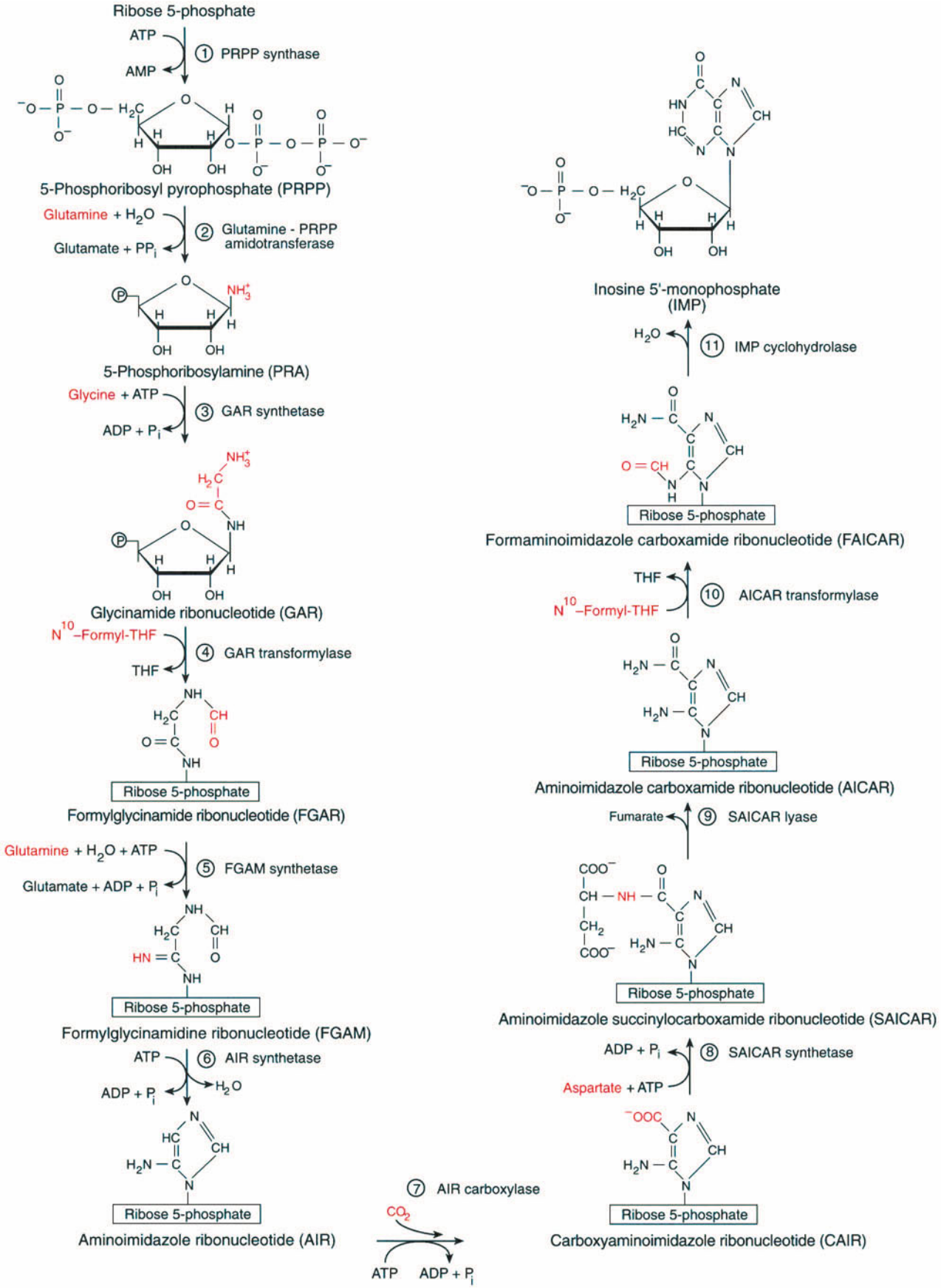


Figure 14.17. Purine biosynthesis. IMP is synthesized *de novo* using ribose 5-phosphate as anchor.

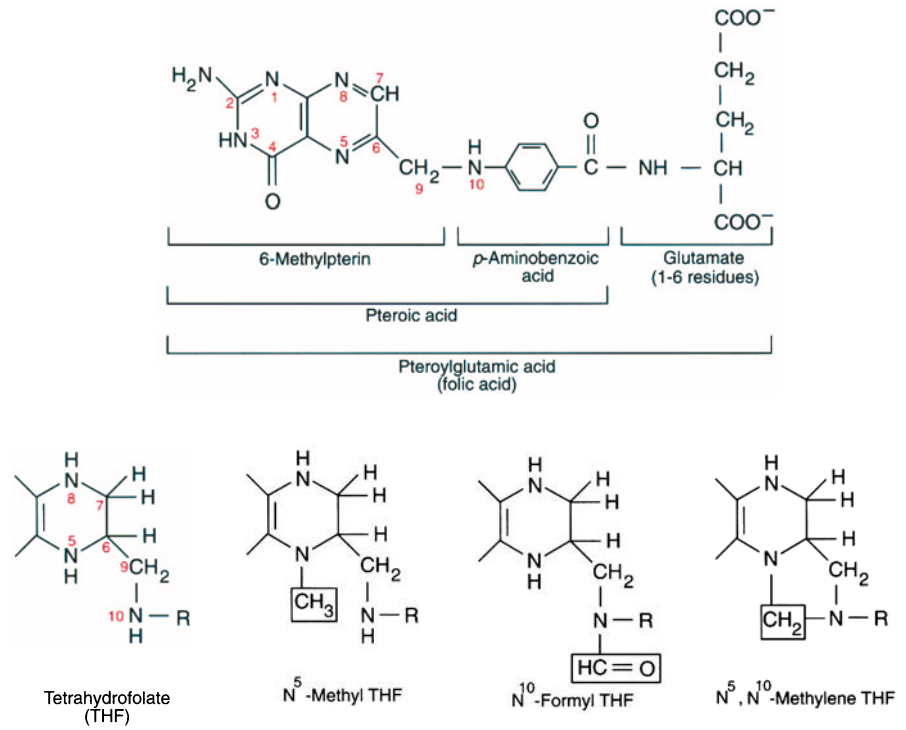


Figure 14.18. Folic acid and its coenzyme forms.

formed in the more complicated and energy-requiring *de novo* biosynthetic pathways. PRPP plays an important role in salvage of purines and pyrimidines because these pathways involve *phosphoribosyl transferases,* enzymes that catalyze reactions of the type

$$\text{Base} + \text{PRPP} \rightleftarrows \text{ribonucleoside 5′-phosphate} + PP_i$$

This readily reversible reaction is pulled to completion by pyrophosphatase-catalyzed hydrolysis of PP_i to inorganic phosphate.

Purine deoxyribonucleotides are synthesized from corresponding ribonucleotides by reduction at C(2′), a reaction catalyzed by **ribonucleotide reductase.** Several forms of the enzyme occur in different species. The most widely distributed form in both prokaryotes and eukaryotes contains a nonheme iron. The enzyme has a reactive disulfide bond at its active site.

In some organisms, ribonucleotide reductase reduces nucleoside triphosphates (NTP) directly to deoxynucleoside triphosphates (dNTP). In most organisms, however, reduction occurs at the nucleoside diphosphate level. Reduction of nucleoside diphosphates (NDP) produces deoxynucleoside diphosphates (dNDP), which are then converted to deoxynucleoside triphosphates (dNTP) by phosphorylation, catalyzed by *nucleoside diphosphate kinase:*

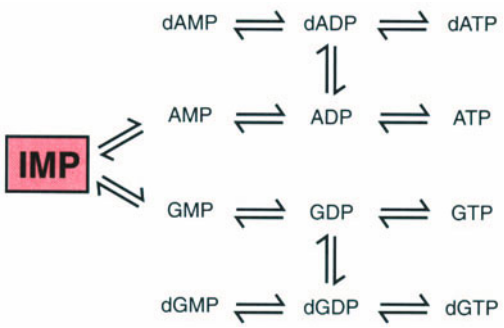
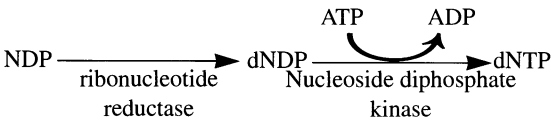


Figure 14.19. Interconversions of purine nucleotides.



NADPH provides the reducing power for reduction of nucleotides to deoxynucleotides by ribonucleotide reductase. Electrons are transferred from NADPH to the reductase by means of a mini electron transport system (Figure 14.21) centered about **thioredoxin,** a small electron carrier (MW = 12,000) that contains a reactive disulfide bond.

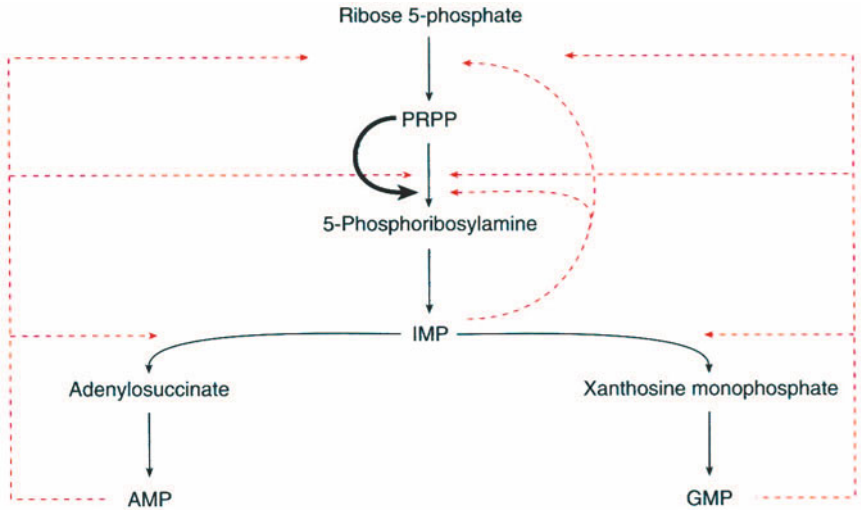
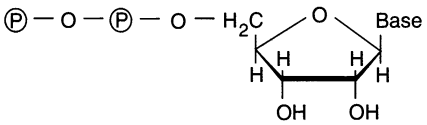


Figure 14.20. Regulation of purine biosynthesis. Feedback inhibiton is shown in red, and activation is shown in black.

As electrons flow from NADPH, they first reduce *thioredoxin reductase,* an FAD-containing flavoprotein with a reactive disulfide bond. FAD is reduced to $FADH_2$ and then regenerated by coupling of the oxidation of $FADH_2$ to reduction of the enzyme's reactive disulfide bond. From thioredoxin reductase, electrons flow to the reactive disulfide bond of thioredoxin. Reduced thioredoxin then reduces the reactive disulfide bond of ribonucleotide reductase to two sulfhydryl groups. Formation of the two SH groups at the enzyme's active site allows



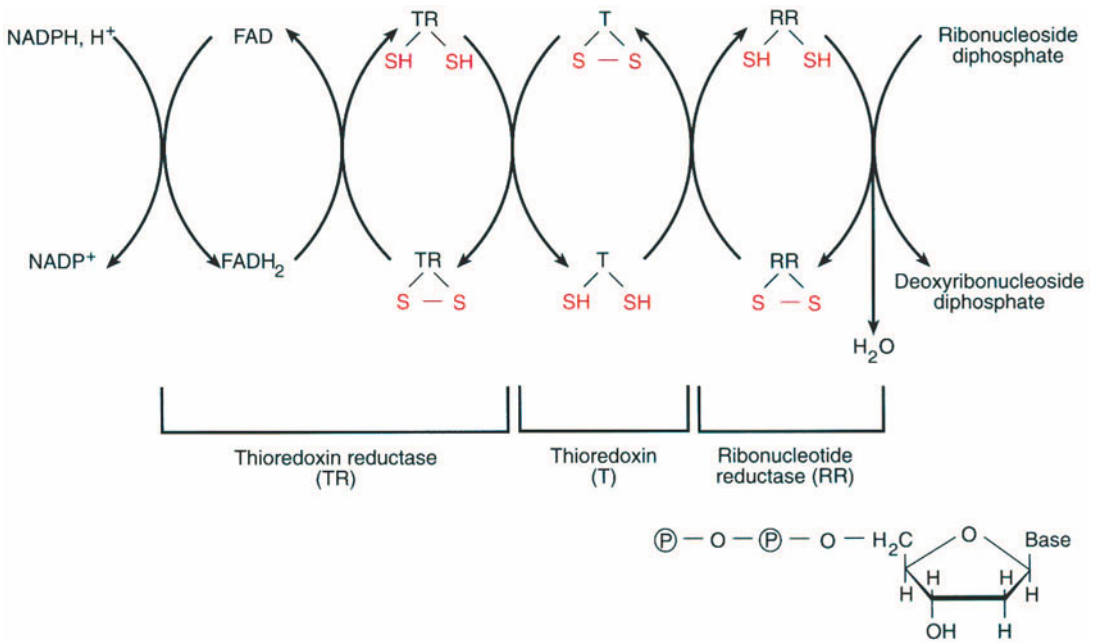


Figure 14.21. An electron transport system of ribonucleotide reductase, centered about thioredoxin.

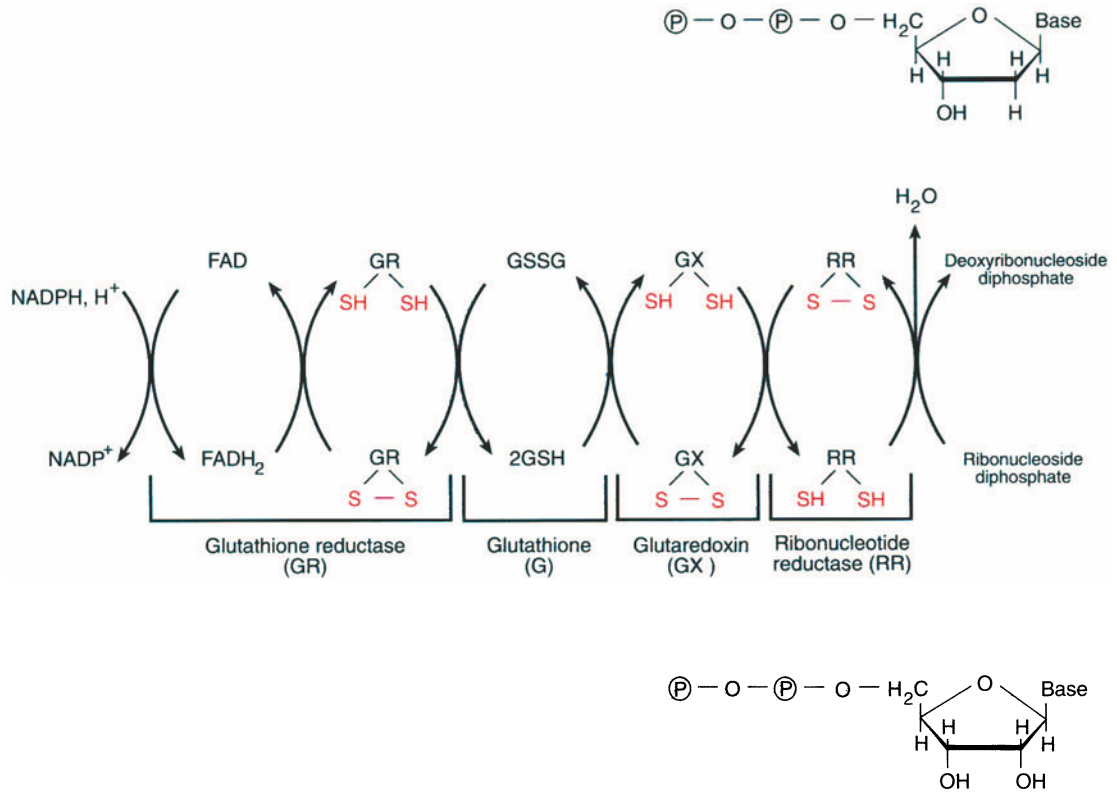


Figure 14.22. An electron transport system of ribonucleotide reductase, centered about glutaredoxin.

reduction of ribose to deoxyribose by a complex free-radical mechanism.

In the absence of thioredoxin (as in *E. coli* mutants lacking this protein), electron flow from NADPH to ribonucleotide reductase proceeds via a second electron transport system, centered about *glutathione* and also containing glutathione reductase and **glutaredoxin** (Figure 14.22). *Glutathione reductase,* like thioredoxin reductase, is an FAD-containing flavoprotein with a reactive disulfide bond. Glutaredoxin resembles thioredoxin in its structure, particularly in the segment containing the two reactive SH groups.

14.4.2. Pyrimidine Biosynthesis

14.4.2A. Synthesis of UMP. Atoms of the pyrimidine ring derive from aspartate, glutamine (amide group), and CO_2 (Figure 14.23). Pyrimidine synthesis differs from purine synthesis in that assembly of the ring system proceeds without a sugar anchor. Only after the ring has been completed does it become coupled to ribose 5-phosphate. **Uridine 5′-monophosphate (UMP)** serves as the common precursor of all the pyrimidines. UMP is synthesized *de novo* in six steps that proceed through *orotate,* a pyrimidine that does not normally occur in nucleic acids (Figure 14.24). The first step of UMP synthesis, catalyzed by carbamoyl phosphate synthase, leads to formation of carbamoyl phosphate. Recall that in eukaryotes the enzyme that catalyzes carbamoyl phosphate synthesis in this pathway is a cytosolic enzyme, *carbamoyl phosphate synthase II.* This enzyme uses glutamine and differs from the mitochondrial carbamoyl phosphate synthase I, which uses ammonia and functions in arginine synthesis in the urea cycle (Section 14.3). Prokaryotes have a single carbamoyl phosphate synthase that functions in both pyrim-

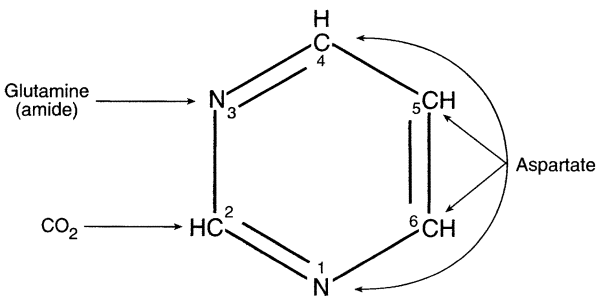


Figure 14.23. Assembly of the pyrimidine nucleus.

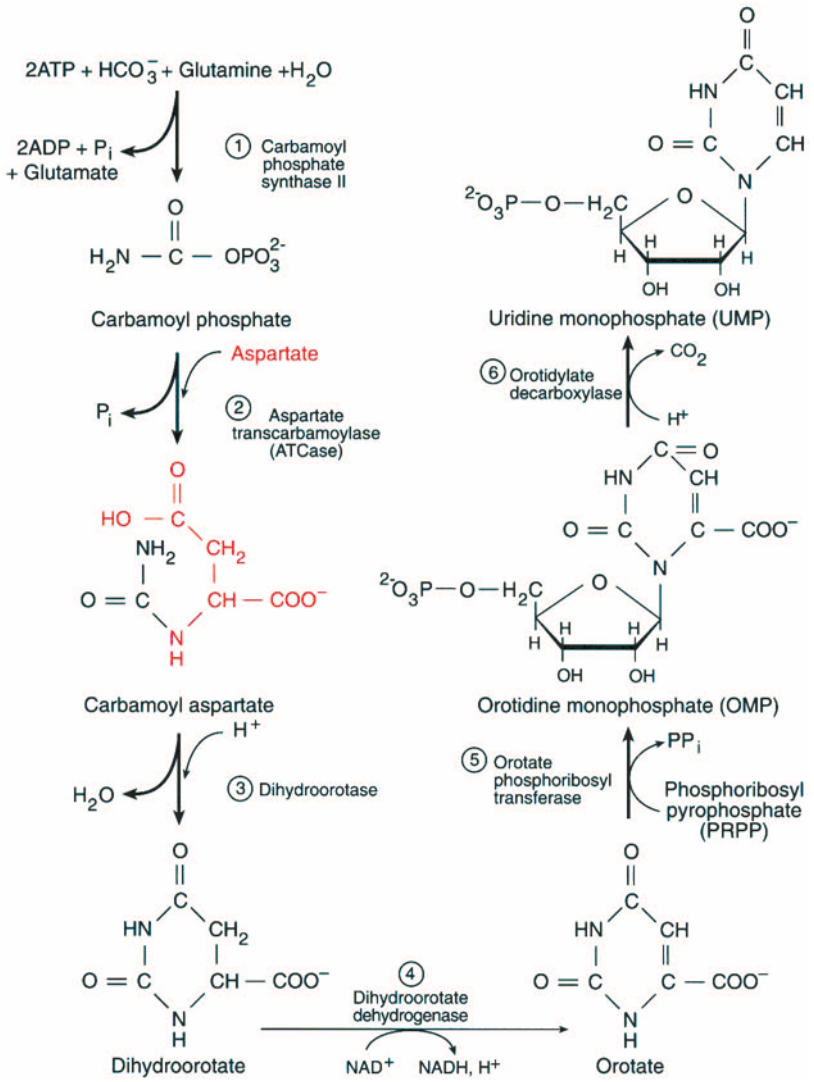


Figure 14.24. Pyrimidine biosynthesis. The pyrimidine ring is assembled first (as orotate) and then linked to ribose 5-phosphate.

idine and arginine biosynthesis and uses glutamine for both processes.

The second reaction of the UMP biosynthetic pathway involves catalysis by **aspartate transcarbamoylase (ATCase).** Bacterial ATCase is an allosteric enzyme that constitutes a major control point for purine biosynthesis. *E. coli* ATCase is one of the most thoroughly studied allosteric enzymes. Note that phosphoribosyl pyrophosphate (PRPP), which serves as a precursor in purine biosynthesis, also serves as a precursor in pyrimidine biosynthesis.

14.4.2B. Interconversions of Pyrimidine Nucleotides.

All of the pyrimidine nucleotides are derived from UMP (Figure 14.25). Thymine nucleotides are derived from dUMP. Methylation of dUMP yields dTMP in a reaction catalyzed by *thymidylate synthase* (Figure 14.26). The enzyme uses N^5,N^{10}-methyl-

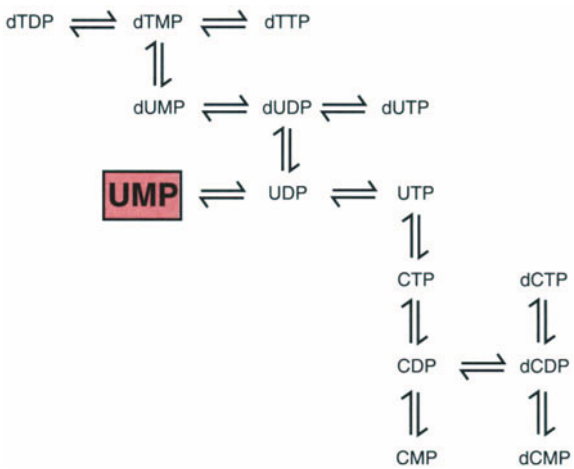


Figure 14.25. Interconversions of pyrimidine nucleotides.

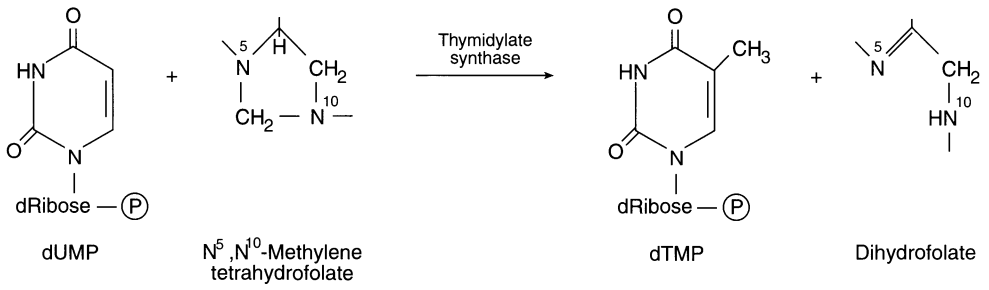


Figure 14.26. Synthesis of dTMP from dUMP by thymidylate synthase.

enetetrahydrofolate as methyl-group donor. The folate coenzyme has a dual function in this reaction, serving as both a carrier of a one-carbon fragment and as a reducing agent. Tetrahydrofolate carries and donates a methylene group ($-CH_2-$) during the reaction. However, since the methylene group becomes converted to a methyl group ($-CH_3-$) in thymidine monophosphate, the coenzyme also serves as a reducing agent. In the process, it is oxidized from N^5,N^{10}-methylenetetrahydrofolate to 7,8-dihydrofolate. The active tetrahydrofolate carrier must be regenerated, since only it serves as a methylene group carrier for this reaction. Reduction of dihydrofolate to tetrahydrofolate is catalyzed by *dihydrofolate reductase,* an enzyme that requires NADPH as coenzyme. Because of this reduction, the overall conversion of dUMP to dTMP requires reducing power in the form of NADPH.

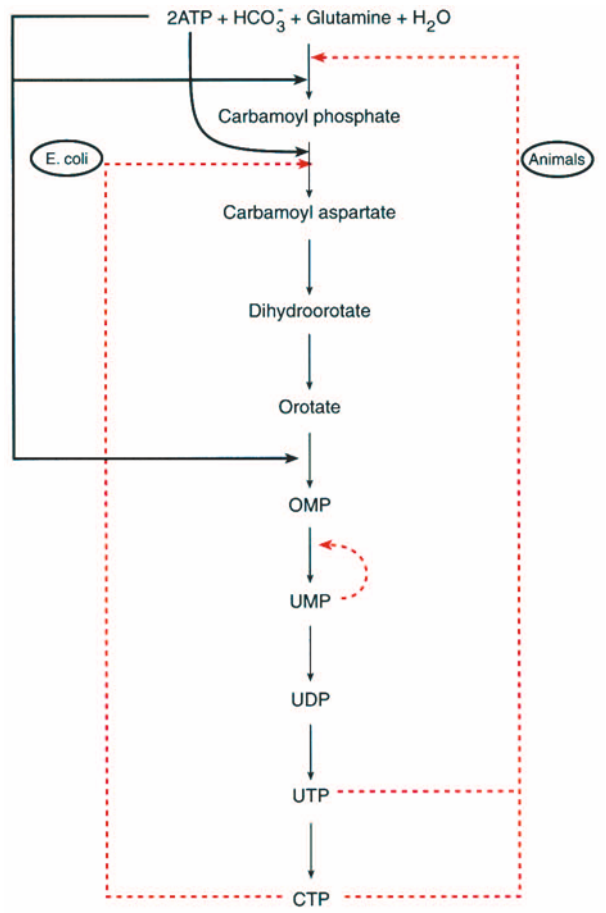
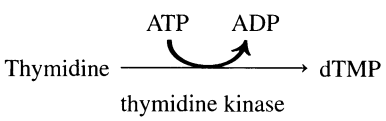


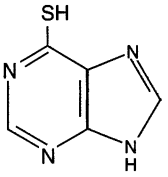
Figure 14.27. Regulation of pyrimidine biosynthesis in *E.coli* and animals. Feedback inhibiton is shown in red, and activation is shown in black.

Major control points in the biosynthesis of pyrimidines occur at the first two steps of the UMP biosynthetic pathway (Figure 14.27). *E. coli* ATCase is activated by ATP and is subject to feedback inhibition by CTP. In many other bacterial systems, UTP is the major inhibitor of ATCase. The ATCase of animals is not an allosteric enzyme. In animals, control of pyrimidine biosynthesis occurs via feedback inhibition of UTP and CTP on carbamoyl phosphate synthase II.

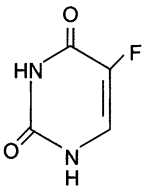
Synthesis of pyrimidine nucleotides and deoxynucleotides by means of salvage pathways proceeds via the same reactions described for purines. Deoxynucleotides are synthesized from the corresponding nucleoside diphosphates by ribonucleotide reductases. As in the case of purines, pyrimidine salvage pathways involve catalysis by phosphoribosyl transferases. A special salvage pathway exists for dTMP, which can be synthesized from thymidine (deoxythmidine) in a reaction catalyzed by an ATP-dependent *thymidine kinase:*



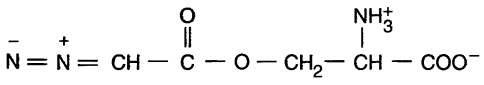
This pathway has a practical application for experiments requiring the labeling of DNA by means of radioactive thymidine. Thymidine enters cells readily and is largely metabolized to dTMP by the above reaction. Subsequent conversion of dTMP to dTTP produces a substrate for DNA polymerase and results in incorporation of labeled thymidine into newly synthesized DNA.



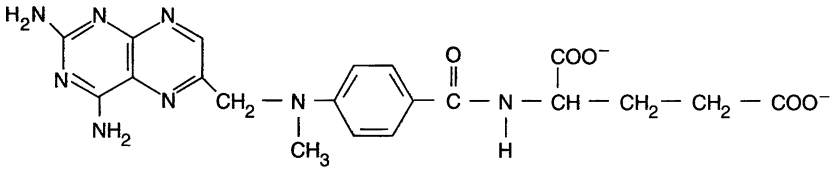
6-Mercaptopurine



5-Fluorouracil



Azaserine



Methotrexate

Figure 14.28. Some drugs used in cancer chemotherapy.

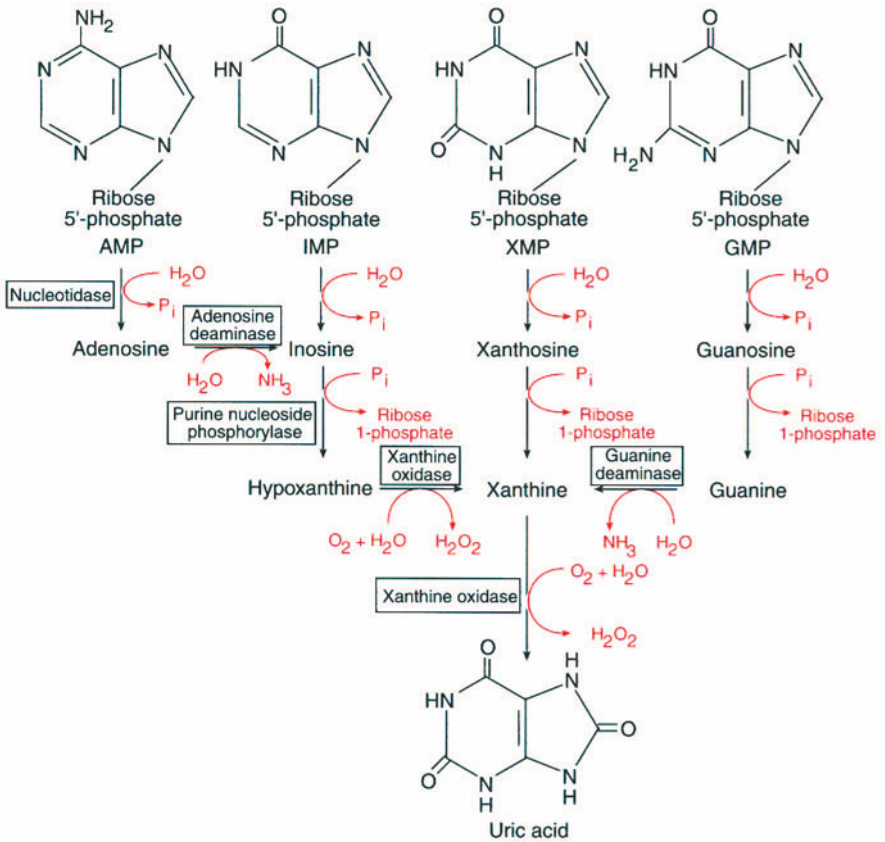


Figure 14.29. Catabolism of purines in animals. XMP, Xanthosine 5′-monophosphate. Xanthine oxidase catalyzes both the conversion of hypoxanthine to xanthine and that of xanthine to uric acid. Deoxyribonucleotides are degraded likewise.

14.4.2C. Cancer Chemotherapy. Our understanding of the pathways for *de novo* synthesis of purines and pyrimidines has been of great benefit in the design of drugs for chemotherapy of cancer. Knowledge of the biosynthetic pathways allows for synthesis of drugs that inhibit purine and pyrimidine biosynthesis at specific steps. Such inhibitors are very toxic to all cells, since all cells require an adequate supply of nucleotides to synthesize DNA and RNA for chromosome duplication and protein synthesis, respectively. However, inhibitors have particularly great toxicity for cells that grow rapidly and require an above-average supply of nucleotides. Because many tumors contain cells that grow and proliferate rapidly, they are affected by the anticancer drugs to a greater extent than normal cells. Consequently, destruction of cancer cells exceeds that of normal cells.

6-Mercaptopurine and **5-fluorouracil** (Figure 14.28) are two useful drugs in cancer chemotherapy. Neither compound by itself possesses cytotoxicity, but their conversion in humans leads to formation of potent nucleotide metabolism inhibitors. 6-Mercaptopurine undergoes conversion to the corresponding purine nucleotide, which acts as a powerful inhibitor of glutamine-PRPP amidotransferase in purine biosynthesis (step 2). 5-Fluorouracil undergoes conversion via a salvage pathway to its deoxyribonucleotide, 5-fluorodeoxyuridylate, which serves as an irreversible inhibitor of thymidylate synthase. 5-Fluorodeoxyuridylate constitutes a *suicide substrate* of the enzyme (see Section 11.3) because it forms an inactive ternary complex composed of thymidylate synthase, 5-fluorodeoxyuridylate, and tetrahydrofolate.

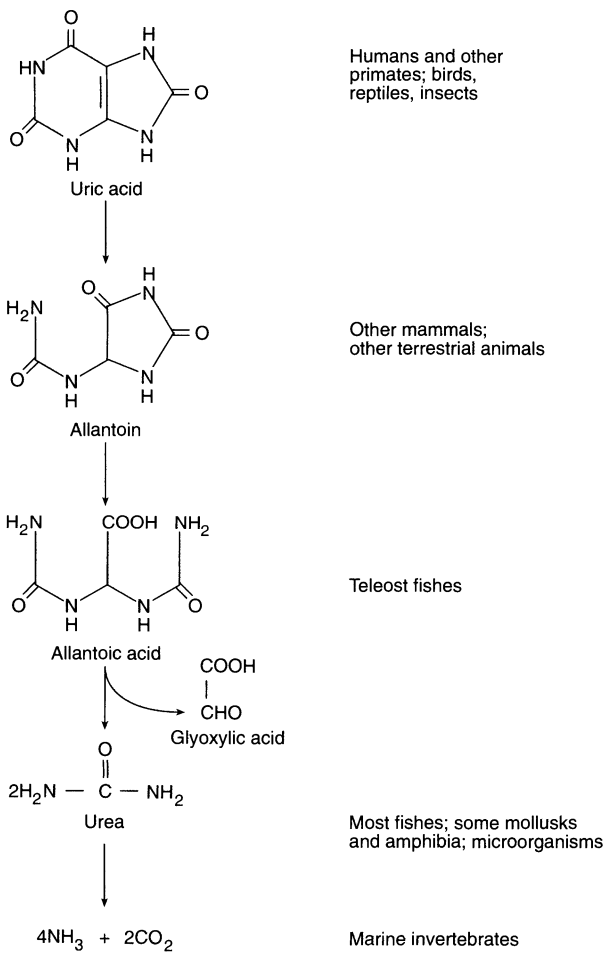


Figure 14.30. End products of purine catabolism. The nitrogenous compounds shown are excreted by the animals listed on the right.

Azaserine and **methotrexate (amethopterin)** are two other drugs that have proven effective in treating some cancers (Figure 14.28). Azaserine is an analog of glutamine, and methotrexate is an analog of folic acid. These drugs belong to two groups of compounds called *glutamine antagonists* and *antifolates,* respectively. Azaserine inhibits glutamine amidotransferases and therefore inhibits purine biosynthesis. Methotrexate is a potent inhibitor of dihydrofolate reductase, which functions in the thymidylate synthase reaction. Methotrexate binds at least 1000 times more tightly to the enzyme than the normal substrate. Administration of methotrexate results in decreased levels of N^5,N^{10}-methylenetetrahydrofolate and greatly decreased dTMP synthesis by thymidylate synthase. The effect of methotrexate on thymidylate synthase is so pronounced because the enzyme is particularly sensitive to depletion of tetrahydrofolate compared to other tetrahydrofolate-requiring enzymes. The high sensitivity of thymidylate synthase results from the dual role that tetrahydrofolate plays in the reaction.

Compounds that interfere with purine and pyrimidine metabolism may also be useful used as antiviral drugs, since viruses also constitute rapidly replicating systems critically dependent on a supply of nucleotides. Antiviral drugs, like cancer drugs, must be such that the destruction of the rapidly replicating viruses exceeds the destruction of the more slowly replicating host cells. An example of an antiviral drug is *AZT* (Figure 7.8), used for treatment of AIDS, caused by a virus termed HIV (human immunodeficiency virus). AZT inhibits DNA polymerase. The viral enzyme (an RNA-dependent DNA polymerase) is at least 100 times more sensitive to the drug than the host cell enzyme (a DNA-dependent DNA polymerase).

14.4.3. Purine Catabolism

Purine catabolism begins with hydrolysis of nucleotides to nucleosides, a reaction catalyzed by *nucleotidases.* Cleavage of nucleosides by *phosphorolysis* (see Section 10.5) yields ribose 1-phosphate (or deoxyribose 1-phos-

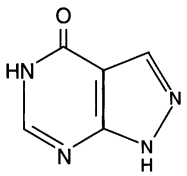


Figure 14.31. Allopurinol, an inhibitor of xanthine oxidase.

phate) and free purines. Phosphorolysis is catalyzed by *purine nucleoside phosphorylase,* an enzyme widely distributed in both mammalian tissues and microbial cells. All of the purines ultimately undergo conversion to **xanthine,** a purine not normally found in nucleic acids. Xanthine is converted to **uric acid** by means of **xanthine oxidase** (Figure 14.29).

The subsequent fate of uric acid varies with the organism (Figure 14.30). In some organisms, uric acid constitutes the ultimate breakdown product of purines and is excreted as such. In others, uric acid undergoes further degradation to **allantoin, allantoic acid,** urea, or ammonia.

Gout is a disease caused by abnormal uric acid metabolism. The disease is fairly common, afflicting about 3 out of every 1000 people, mostly males. It is associated with an elevated plasma level of uric acid. The high concentration of uric acid causes formation of painful deposits of sodium urate in the cartilage of joints, especially that of the big toe,

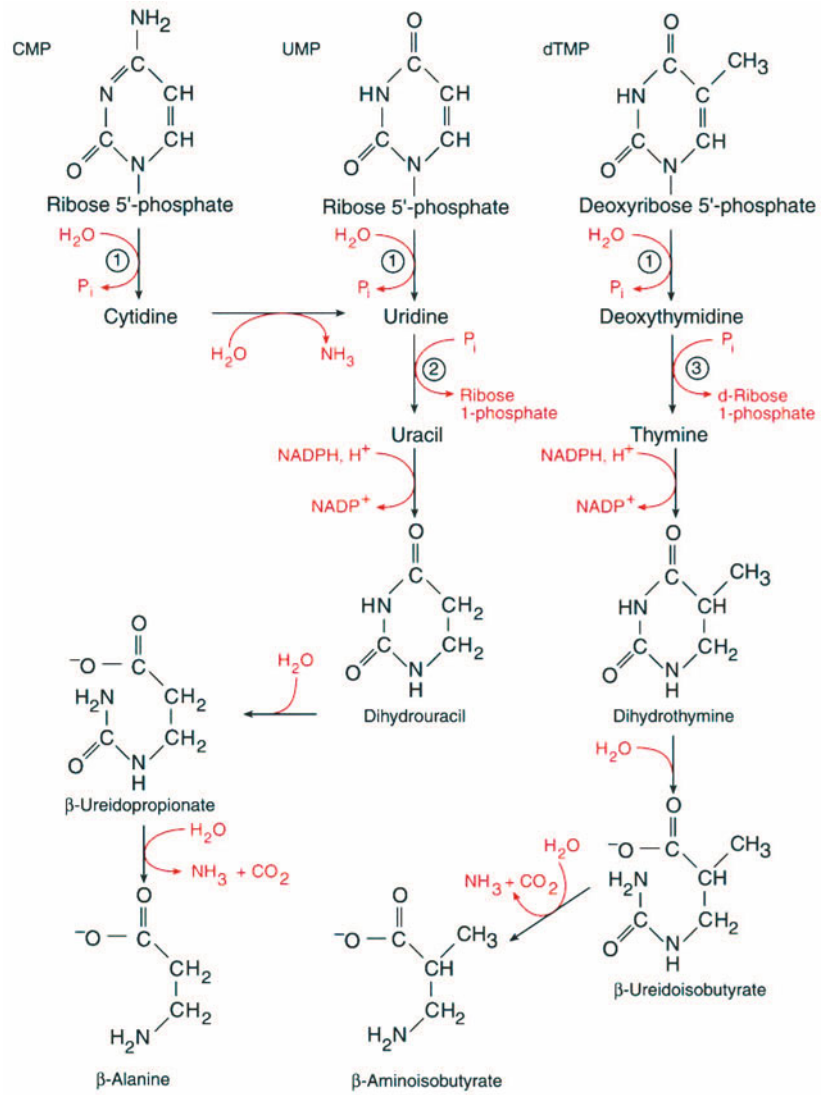


Figure 14.32. Catabolism of pyrimidines in animals. β-Alanine and β-aminoisobutyrate are subsequently metabolized to malonyl CoA and methylmalonyl CoA, respectively. Deoxyribonucleotides are degraded likewise. Numbers designate enzymes: ① Nucleotidase; ② uridine phosphorylase; ③ thymidine phosphorylase.

and can also form kidney stones. Many cases of gout can be treated successfully with the antimetabolite **allopurinol** (Figure 14.31), an inhibitor of xanthine oxidase.

14.4.4. Pyrimidine Catabolism

Pyrimidine catabolism also begins with nucleotidase-catalyzed hydrolysis of nucleotides to nucleosides, followed by phosphorolysis of nucleosides. The latter reaction, catalyzed by *pyrimidine nucleoside phosphorylase,* yields ribose 1-phosphate (or deoxyribose 1-phosphate) and free pyrimidines. Cytosine and uracil are subsequently degraded to β-alanine; thymine is degraded to β-aminoisobutyrate (Figure 14.32).

SUMMARY

Some bacteria can reduce atmospheric nitrogen to ammonia (nitrogen fixation) by means of nitrogenase. Other bacteria, plants, and fungi can reduce nitrate and nitrite to ammonia (nitrate assimilation) using nitrate reductase. All living organisms can use ammonia and convert it to nitrogen-containing organic compounds. Three key reactions lead to conversion of ammonia to carbamoyl phosphate, glutamate, or glutamine. These nitrogenous compounds react with intermediates of carbohydrate and lipid metabolism to form other nitrogen-containing compounds. One such reaction, catalyzed by transaminases, results in conversion of an amino acid to a keto acid while a different keto acid undergoes conversion to an amino acid.

Essential amino acids are those that an organism cannot synthesize or cannot synthesize in sufficient quantity; they must be obtained through the diet. Essential amino acids in humans were determined by nitrogen balance studies in which the difference between the amount of nitrogen ingested and that excreted was measured. Glucogenic and ketogenic amino acids are catabolized to precursors of carbohydrates and lipids, respectively.

The first step in amino acid catabolism usually consists of an oxidative deamination. The ammonia removed enters the urea cycle, where it is converted to urea that can be excreted in the urine. The urea cycle constitutes a catabolic pathway that requires input of energy and has close links to many other metabolic pathways. Mammals and terrestrial vertebrates excrete waste nitrogen as urea. Other organisms excrete waste nitrogen as uric acid or ammonia.

Individuals suffering from the genetic disease phenylketonuria have a deficiency of phenylalanine hydroxylase. This enzyme catalyzes the hydroxylation of phenylalanine to tyrosine in the major pathway of phenylalanine catabolism. In afflicted individuals, a minor pathway becomes accentuated and leads to conversion of phenylalanine to phenylpyruvate and phenyllactate.

Purine and pyrimidine ring systems are synthesized from small metabolic precursors. The purine ring is assembled on a sugar phosphate molecule. The pyrimidine ring is assembled by itself and subsequently linked to a sugar phosphate. A number of compounds function as anticancer drugs by inhibiting specific steps in purine or pyrimidine biosynthesis. Hydrolysis of purine and pyrimidine nucleotides to nucleosides is catalyzed by nucleotidases. The nucleosides are then cleaved by phosphorolysis to yield free bases.

SELECTED READINGS

Bender, D., *Amino Acid Metabolim,* 2nd ed., Wiley, Chichester, England (1985).

Carreras, C. W., and Santi, D. V., The catalytic mechanism and structure of thymidylate synthase, *Annu. Rev. Biochem.* 64:721–762 (1995).

Cavenee, W. K., and White, R. L., The genetic basis of cancer, *Sci. Am.* 272:72–79 (1995).

Dou, Q. P., and Pardee, A. B., Transcriptional activation of thymidine kinase, a marker for cell cycle control, *Prog. Nucleic Acid Res. Mol. Biol.* 53:197–217 (1996).

Fontecave, M., Nordlund, P., Eklund, H., and Reichard, P. The redox centers of ribonucleotide reductase of *E. coli, Adv. Enzymol.* 65:147–183 (1992).

Greenberg, G. R., and Hilfinger, J. M., Regulation of ribonucleotide reductase and relationship to DNA replication in various systems, *Prog. Nucleic Acid Res. Mol. Biol.* 53:345–395 (1996).

Howard, J. B., and Rees, D. C., Nitrogenase: A nucleotide-dependent molecular switch, *Annu. Rev. Biochem.* 63:235–264 (1994).

Kaufman, S., The phenylalanine hydroxylating system, *Adv. Enzymol.* 67:77–264 (1993).

Lipscomb, W. N., Aspartate transcarbamylase from *E. coli:* Activity and regeneration, *Adv. Enzymol.* 68:67–151 (1994).

Peters, J. W., Fisher, K., and Dean, D. R., Nitrogenase structure and function: A biochemical-genetic perspective, *Annu. Rev. Microbiol.* 49:335–366 (1995).

Smith, J. L, Zaluzec, E. J., Wery, J. P., Niu, L., Switzer, R. L., Zalkin, H., and Satow, Y., Structure of the allosteric regulatory enzyme of purine biosynthesis, *Science* 264:1427–1433 (1994).

Traut, T. W., and Jones, M. E., Uracil metabolism—UMP synthesis from orotic acid or uridine and conversion of uracil to β-alanine: Enzymes and cDNAs, *Prog. Nucleic Acid Res. Mol. Biol.* 53:1–78 (1996).

REVIEW QUESTIONS

A. Define each of the following terms:

Transamination	Nitrogen balance
Oxidative deamination	Phenylketonuria
Salvage pathway	Phosphoribosyl pyrophosphate
Nitrate assimilation	Nitrogenase
Ribonucleotide reductase	Tetrahydrofolate

B. Differentiate between the two terms in each of the following pairs:

UMP/IMP	SGOT/SGPT
Ureotelic organisms/ uricotelic organisms	Nitrogen fixation/ ammonia fixation
Nitrate reductase/ nitrite reductase	Glucogenic amino acids/ ketogenic amino acids
Glutamine synthase/ glutamate dehydrogenase	Aspartate transcarbamoylase/carbamoyl phosphate synthase

C. (1) Outline the steps whereby living organisms convert atmospheric nitrogen to nitrogen-containing organic compounds.
(2) Write two balanced equations showing the action of adenosine phosphorylase and uridine phosphorylase.
(3) Complete the following transamination reactions, using structural formulas for reactants and products:

Phenylalanine + pyruvate ⇄
Glycine + oxaloacetate ⇄

(4) Outline the operation of the urea cycle. Write balanced equations for the individual steps. What are the energy requirements of the cycle and how is the cycle linked to other metabolic pathways?
(5) What are some of the similarities and differences between purine and pyrimidine biosynthesis and between purine and pyrimidine catabolism?

PROBLEMS

14.1. Which of the following conditions is likely to lead to a positive nitrogen balance in humans?

(a). Inability to digest dietary protein
(b). Chronic deterioration of muscle tissue
(c). Inhibition of amino acid deamination
(d). Inhibition of protein biosynthesis
(e). Deficiency of amino acid absorption from the intestine

14.2. We classify arginine as an essential amino acid for young organisms but *not* for adults. How do you explain this, considering that arginine is synthesized, even in young organisms, via the urea cycle?

14.3. In terms of energy-rich bonds, how many molecules of urea could be synthesized from the energy released by the complete aerobic oxidation to CO_2 and H_2O of one molecule of: (a) glucose; (b) palmitic acid?

14.4. Name the labeled compound formed by transamination of α-ketoglutarate labeled with ^{14}C in its

keto group. Where is the label located in the compound formed?

14.5. You place a rat on a diet including [^{15}N]alanine. Will urea excreted by the rat become labeled with ^{15}N and, if so, will it become labeled in one or both of its amino groups?

14.6. Write the overall, balanced equation describing the oxidative deamination of D-phenylalanine with D-amino acid oxidase and catalase.

14.7. Name the enzyme that catalyzes each of the following reactions:

(a) $HN{=}NH + 2H^+ + 2e^- \rightleftarrows H_2N\text{-}NH_2$
(b) $NO_3^- + 2H^+ + 2e^- \rightleftarrows NO_2^- + H_2O$
(c) $NO^- + 3H^+ + 2e^- \rightleftarrows NH_2OH$
(d) $NH_4^+ + HCO_3^- + 2ATP^{4-} + H_2O \rightarrow$ carbamoyl phosphate^{2-} + $2ADP^{3-}$ + P_i^{2-} + $2H^+$

14.8.* We define the *oxidation number* of an atom in a polyatomic ion or in a compound as the actual or assigned charge that the atom would have if all the electrons of each bond were assigned exclusively to the more electronegative of the bonded atoms. The sum of the oxidation numbers in a polyatomic ion equals the charge of the ion; that in a compound equals zero. On that basis, determine the oxidation number of nitrogen in each of the following: (a) N_2; (b) NO^-; (c) NO_2^-; (d) NO_3^-; (e) $H_2N\text{-}NH_2$; (f) $HN{=}NH$; (g) NH_2OH.

14.9. Write a balanced equation for the synthesis of alanine from glucose using the reactions of glycolysis and transamination.

14.10. What is the energy requirement, in terms of the number of molecules of ATP used, for the biosynthesis of IMP and UMP?

14.11. Consider two human populations—one subsisting on a diet rich in meat, and the other subsisting on a diet rich in rice. Which of the two populations is likely to have a higher incidence of gout? Why?

14.12. Write a balanced equation for the reaction catalyzed by each of the following enzymes:

(a) Glutamate dehydrogenase
(b) Glutamine synthase
(c) Phenylalanine hydroxylase
(d) Arginase
(e) Adenine phosphoribosyl transferase
(f) Cytidine 5′-triphosphate reductase

14.13.* A student isolates ribonucleotide reductase from bacteria. The cell-free extract needs to be stored in the refrigerator for some time. Would you suggest adding some glutathione to the preparation to stabilize the enzyme and minimize the loss of activity upon storage? Why or why not? Would you suggest adding some iodoacetamide when assaying the enzyme *in vitro?* Why or why not?

14.14. When rabbits are grown for a long time on a diet that includes acetyl CoA labeled with ^{14}C at both carbons of the acetyl group, will urea excreted by the rabbits become labeled with ^{14}C as soon as the acetyl CoA enters the citric acid cycle (i.e., during the first turn of the cycle)? Why or why not? Will the urea become labeled at a later time? Why or why not?

14.15. Draw the structures, and name the α-keto acid formed, when the following amino acids undergo transamination with α-ketoglutarate: (a) alanine; (b) glutamate; (c) aspartate; (d) phenylalanine; (e) tyrosine.

14.16. A researcher assays the transamination between alanine and α-ketoglutarate spectrophotometrically by including lactate dehydrogenase and NADH in the incubation mixture. What does the investigator measure? Why can this measurement be used as an assay for the transamination?

14.17. Deficient absorption of vitamin B_{12} from the intestine causes the disease *pernicious anemia.* In amino acid catabolism, isoleucine, methionine, and valine produce propionyl CoA, which undergoes conversion to succinyl CoA. From your knowledge of this latter set of reactions, predict the changes you would expect in isoleucine, methionine, and valine catabolism in pernicious anemia.

14.18.* Calculate the net yield of ATP when alanine serves as a fuel and undergoes complete oxidation to CO_2 and H_2O under aerobic conditions.

14.19.* Cats that have fasted are given a single meal containing all of the amino acids except arginine. Would you expect the catabolism of glucogenic amino acids to be accelerated or slowed down following the meal? Would the blood concentration of ammonia increase or decrease during the same time interval? Explain your answers. (Hint: Arginine is synthesized in insufficient quantities to compensate for its absence from the cat's diet.)

14.20.* Patients suffering from acute leukemia have extensive breakdown of nucleic acids. When treated with anticancer drugs, allopurinol is often added as well. What is the reason for this?

14.21.* The enzyme *hypoxanthine-guanine phosphoribosyl transferase (HGPRT)* catalyzes the following two reactions:

$$\text{Hypoxanthine} + \text{PRPP} \rightarrow \text{IMP} + PP_i$$
$$\text{Guanine} + \text{PRPP} \rightarrow \text{GMP} + PP_i$$

A severe HGPRT deficiency causes an extreme neurological disorder called the *Lesch-Nyhan syndrome.* The syndrome constitutes a sex-linked trait because the gene for HGPRT is located on the sex

chromosome. Afflicted individuals exhibit mental retardation, aggressive behavior, and a compulsion for self-mutilation. Biochemically speaking, the disorder is characterized by an overproduction of uric acid (6 times normal) and excessive *de novo* biosynthesis of purines (200 times normal). How are these metabolic changes brought about?

14.22. A physician advises individuals on a high-protein diet to drink large amounts of water. What is the rationale for this?

14.23. What energy saving, in terms of energy-rich bonds, is achieved by producing IMP via the salvage pathway of Problem 14.21 as opposed to its *de novo* synthesis as shown in Figure 14.17?

14.24. How many α-amino acids participate in the operation of the urea cycle? How many of these are used for the biosynthesis of proteins?

14.25.* A nurse constructs a *phenylalanine tolerance curve* (similar to a glucose tolerance curve) by injecting a large dose of phenylalanine into a person's bloodstream and then determining the phenylalanine concentration in the serum as a function of time. What would such a curve look like for (a) a normal person and (b) a person suffering from PKU disease?

14.26. Assume that a person suffering from phenylketonuria is also diet-conscious and wants to avoid sugar. Would you advise this individual to use aspartame (see Section 2.5) or saccharin (see Figure 5.17) as a food sweetener? Why?

14.27. One of the symptoms of children afflicted with kwashiorkor is a depigmentation of the skin and hair. What is the biochemical basis for this?

14.28.* The daily energy requirement for an average young man is 12,134 kJ (see Table 8.6). You know that proteins contain 16.0% (by weight) of nitrogen. If a man excretes 40.0 g of urea per day, what is his protein intake in grams per day? Given that the energy derivable from proteins is 17.0 kJ/g, calculate the percentage of the man's energy requirement contributed by protein.

14.29. Oxidation of 1.0 g of protein in humans generally yields a smaller number of ATP than oxidation of 1.0 g of carbohydrate or lipid. Why is this so?

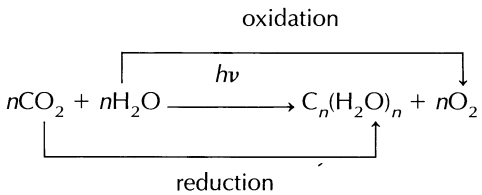
14.30. Why might a potent inhibitor of purine or pyrimidine biosynthesis be useful as an anticancer and/or an antiviral drug?

Photosynthesis

15

Photosynthesis is *the reaction by which algae, bacteria, and plants (photosynthetic organisms), in the presence of certain pigments (especially chlorophyll) and radiant energy from the Sun, synthesize carbohydrates from carbon dioxide and water.* Solar energy is the driving force of photosynthesis and the ultimate source of all the energy living systems require.

Photosynthesis constitutes an oxidation–reduction reaction in which water is oxidized to oxygen and carbon dioxide is reduced to carbohydrate (*CO_2 fixation*):



where $h\nu$ is the energy of a photon of light, and $C_n(H_2O)_n$ is a carbohydrate.

The reactions of photosynthesis lead to production of fructose 6-phosphate. Other monosaccharides (e.g., glucose), oligosaccharides (e.g., sucrose), and polysaccharides (e.g., starch) are then made from fructose 6-phosphate. Because the monomeric unit in these oligomers and polymers is a six-carbon sugar, we usually use a coefficient of $n = 6$ in the above equation, corresponding to the synthesis of glucose:

$$6CO_2 + 6H_2O \xrightarrow[\text{chlorophyll}]{\text{sunlight}} C_6H_{12}O_6 + 6O_2$$

From the air	From the soil	Monosaccharide (glucose)	Returned to the air

This reaction constitutes the reverse of carbohydrate metabolism in plants and animals:

$$\underset{\text{Sugar ingested}}{C_6H_{12}O_6} + \underset{\text{From the air}}{6O_2} \xrightarrow{\text{respiration}} \underset{\text{Expired to the air}}{6CO_2} + \underset{\text{Returned to the soil}}{6H_2O} + \text{energy}$$

Photosynthesis and respiration provide a balance between plant and animal life, between photosynthetic and nonphotosynthetic organisms (Figure 15.1). Photosynthetic organisms convert sunlight energy to chemical energy in the form of ATP and NADPH, both of which then serve to synthesize glucose and other organic compounds from carbon dioxide and water. Simultaneously, photosynthetic organisms release oxygen into the atmosphere.

Both photosynthetic and nonphotosynthetic organisms use the oxygen produced to degrade energy-containing products of photosynthesis to carbon dioxide and water. In the process, energy is released and conserved in the form of ATP. The stored energy of ATP later drives endergonic metabolic reactions.

15.1. THE SCOPE OF PHOTOSYNTHESIS

15.1.1. Brief Historical Perspective

The concept of photosynthesis—plants synthesizing carbohydrates from light and air—took almost two centuries to become firmly established. In 1727, Stephen Hales proposed that plants obtain some of their matter from air. Joseph Priestley, an English clergyman and chemist, investigated the nature of this gaseous component (1770–1780). He showed that a mouse died rapidly if placed in a closed container in which the air had been "depleted" by burning a candle until it went out. However, if the container was connected to a second one containing a green plant exposed to light, the air regained its ability to support a mouse's life.

This classic experiment provided the first indication that plants produce a gas when they are irradiated with light. We now know that this gas consists of oxygen produced by photosynthesis. Priestley later discovered oxygen and called it "dephlogisticated air," but it was Antoine Lavoisier who elucidated oxygen's role in combustion and respiration.

In 1779, Jan Ingen-Housz, a Dutch court physician to the Austrian empress, made another important discovery. He showed that light is essential for photosynthesis

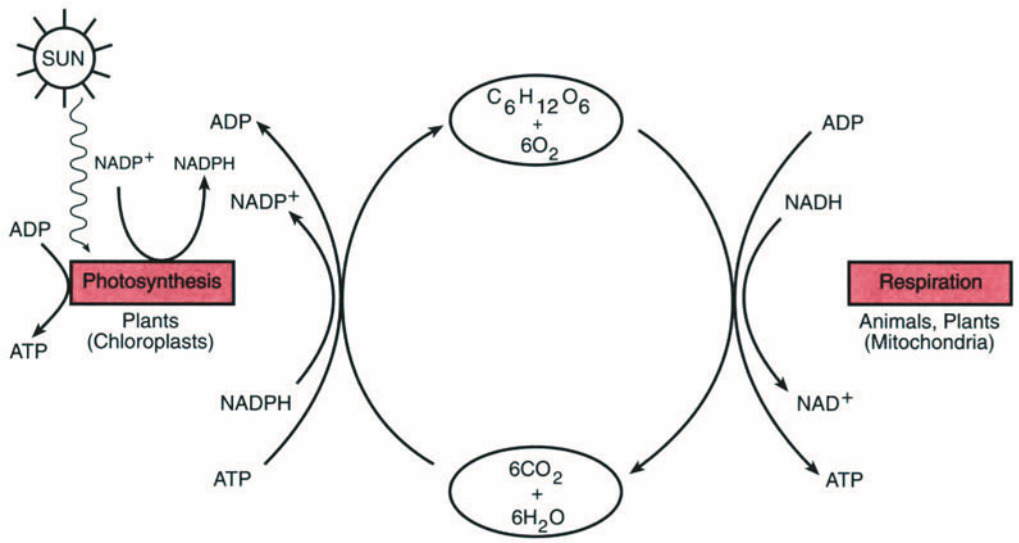


Figure 15.1. The balance between photosynthesis and respiration.

and that only the green parts of a plant produce oxygen. Shortly thereafter, Jean Senebier, a Swiss pastor, showed that CO_2 was taken up during photosynthesis (1782).

By the early 19th century, researchers initiated quantitative measurements of photosynthesis. Some made attempts to determine the amount of CO_2 assimilated, the amount of oxygen evolved, and the amount of plant material produced. Some also considered the energetics of photosynthesis. Robert Mayer, a German surgeon and one of the formulators of the first law of thermodynamics, proposed that plants convert light energy to chemical energy (1842).

At the beginning of the 20th century (1905), F. Blackman proposed that photosynthesis consists of two stages: (1) a *light reaction* that is light dependent but temperature independent, like typical photochemical reactions, and (2) a *dark reaction* that is temperature dependent, like typical enzymatic reactions.

In 1931, Cornelius van Niel showed that photosynthetic sulfur bacteria use H_2S to generate sulfur and proposed a general equation to describe photosynthesis in both plants and sulfur bacteria:

$$CO_2 + 2H_2A \xrightarrow{\text{light}} C(H_2O) + H_2O + 2A$$

where H_2A and 2A are H_2O and O_2 in green plants but H_2S and 2S in photosynthetic sulfur bacteria. Van Niel proposed that the first stage of photosynthesis consists of the oxidation of H_2A to 2A and the formation of a reducing agent,

$$2H_2A \xrightarrow{\text{light}} 2A + \text{reducing agent}$$

followed by a second stage in which the reducing agent converts CO_2 to carbohydrate:

$$\text{Reducing agent} + CO_2 \longrightarrow C(H_2O) + H_2O$$

Based on this reasoning, van Niel predicted that the oxygen evolved in photosynthesis comes from water. His prediction was confirmed when ^{18}O became available. By using ^{18}O-labeled H_2O, Samuel Ruben and Martin Kamen showed that the evolved oxygen was derived from water (1941). The oxygen atoms of carbon dioxide appear in the other two products of the reaction, carbohydrate and water:

$$CO_2^* + 2H_2{}^{18}O \xrightarrow{\text{light}} C(H_2O^*) + H_2O^* + {}^{18}O_2$$

In 1932, Robert Emerson made a crucial discovery while measuring photosynthesis as a function of the wavelength of light. He observed that photosynthesis decreased when he exposed plants to monochromatic light of longer wavelengths (about 700 nm), even though the plants still absorbed light significantly at these longer wavelengths. We refer to this decrease in photosynthetic activity as **red drop** (Figure 15.2). When the same plants were supplemented with light of shorter wavelengths (about 650 nm), photosynthetic activity increased, a phenomenon called the **Emerson enhancement effect.** Emerson's observations suggested that there are *two photosystems* in photosynthesis: one operating at about 700 nm, and one at about 650 nm.

In the late 1930s, the work of Robert Hill led to another milestone in the study of photosynthesis. Hill showed that isolated *chloroplasts* (first described by A. Meyer in 1883) could produce oxygen in response to light and that ferric ions could substitute for CO_2 in the reaction. The photochemical evolution of oxygen was coupled to the reduction of ferric to ferrous ions, resulting in what became known as the **Hill reaction:**

$$4Fe^{3+} + 2H_2O \xrightarrow[\text{chloroplasts}]{\text{light}} 4Fe^{2+} + 4H^+ + O_2$$

Hill's discovery ushered in an era of cell-free studies of photosynthesis. In 1954, D. I. Arnon and his co-workers showed that isolated chloroplasts could carry out complete photosynthesis—that is, produce oxygen *and* fix carbon dioxide. They also discovered *photophosphorylation,* the synthesis of ATP coupled to operation of a pho-

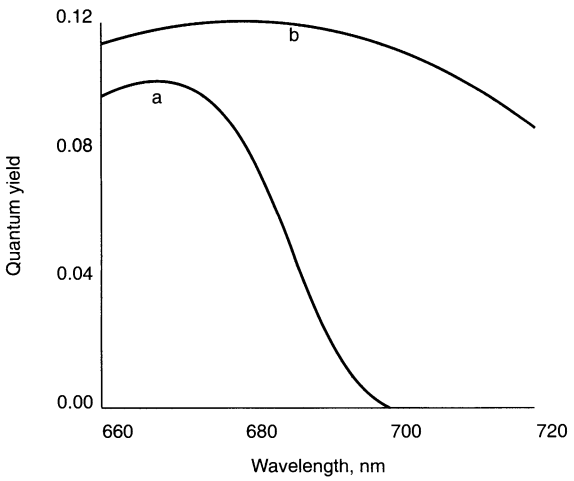


Figure 15.2. The "red drop" of photosynthesis. The quantum yield (defined in Section 15.2) decreases (a) when the illumination consists of monochromatic light of longer wavelengths but increases (b) when the illumination is supplemented with light of shorter wavelengths (<650 nm).

tosynthetic electron transport system. Detailed steps of the dark reaction were not unraveled until the advent of radioactive isotopes. Melvin Calvin and his associates finally elucidated the pathway of carbon dioxide fixation in the 1950s.

15.1.2. Evolution of Photosynthesis

The scenario outlined below is based on the assumption that the Earth's gaseous cover evolved from a reducing to an oxidizing atmosphere. As you read in Section 1.1, recent findings have led many researchers to postulate that the primordial atmosphere was oxidizing. If this view prevails, the scenario must be modified accordingly.

Assuming that photosynthesis arose when the Earth had a reducing atmosphere and an abundance of reducing agents, anaerobic organisms are thought to have been the first to carry out photosynthesis. Modern photosynthetic sulfur bacteria (Table 15.1), which use H_2S to reduce CO_2, are believed to resemble these early photosynthetic organisms.

As anaerobic photosynthetic organisms multiplied and became established, they began sealing their own doom. Their very success at carrying out reduction reactions resulted in continuous depletion of reducing agents in the environment. Some theorists believe that forerunners of modern *cyanobacteria* (Table 15.1) were the first to adapt to these changing environmental conditions and evolved a system capable of extracting electrons from water. This probably occurred some 3×10^9 yr ago and led to accumulation of oxygen, a toxic waste product (see Section 12.5).

The advent of aerobic photosynthesis produced a gradual change in the Earth's environment, converting the reducing atmosphere to an oxidizing one. Conversion of the atmosphere paved the way for development of aerobic metabolism, evolution of animals, and establishment of a balance between photosynthesis and respiration. As the oxidizing atmosphere spread, anaerobic photosynthetic organisms decreased in significance. Today, they constitute only a small fraction of the total number of photosynthetic organisms.

Photosynthesis has evolved to become the single most important chemical reaction in the biosphere. It exceeds any other kind of "manufacturing" reaction in the world in the quantities of reactants and products involved. You can appreciate the magnitude of photosynthesis by considering the following statistics:

- An estimated 4×10^{11} tons of CO_2 (about 10^{11} tons of carbon) are fixed by photosynthesis per year.
- Fixation of 10^{11} tons of carbon represents an energy trapping of about 4×10^{18} kJ/yr. This energy, derived from the Sun, amounts to well over 10 times the total energy generated by all fossil fuels used the world over.
- As immense as the energy trapped is, it constitutes only about 0.1% of the total radiant energy striking the Earth.
- At least one-half of the total photosynthetic activity on Earth takes place in oceans, rivers, and lakes due to algae and microorganisms. The rest is due to terrestrial plants.
- Agricultural crops of the United States alone produce approximately 6×10^6 tons of chlorophyll per year.
- Approximately 3 centuries (for CO_2) and 20 centuries (for O_2) are required to process by photosynthesis an amount of gas equal to that present in the atmosphere.

15.2. LIGHT AND ENERGY

To understand the essence of photosynthesis, we must review some of the properties of light. Recall that light has both wavelike and particulate character. We refer to the particles of light as **photons.** Each photon represents a fixed package of light energy, termed a **quantum.**

Table 15.1. Photosynthesis in Prokaryotes and Eukaryotes

Organisms	Formation of O_2	Presence of chloroplasts	Type of chlorophyll	Number of photosystems
Prokaryotes				
Sulfur bacteria	No	No	Bacteriochlorophyll *a* or *b*	1
Cyanobacteria	Yes	No	Chlorophyll *a*	2
Eukaryotes				
Plants and algae	Yes	Yes	Chlorophyll *a* and *b*	2

15.2.1. Energy of Photons

A photon has no charge and is believed to have no mass. The energy of a photon varies with the wavelength of light as defined by **Planck's law:**

$$E = h\nu \tag{15.1}$$

where E is the energy of a photon, h is Planck's constant (6.626×10^{-34} J s), and ν is the frequency of the light (cycles/s).

The frequency is equal to the velocity of light in a vacuum ($c = 3.00 \times 10^{10}$ cm s^{-1}) divided by the wavelength (λ):

$$\nu = c/\lambda \tag{15.2}$$

Because photon energy varies inversely with photon wavelength, the shorter the wavelength, the greater the energy. A photon of blue light (e.g., $\lambda = 500$ nm) has greater energy than a photon of red light (e.g., $\lambda = 700$ nm). You can calculate the energy of a photon by means of Eq. (15.1). For example, a photon of red light having a wavelength of 700 nm has an energy equal to

$$E_{700} = (6.626 \times 10^{-34} \text{ J s})\,(3.00 \times 10^{10} \text{ cm s}^{-1}/7.00 \times 10^{-5} \text{ cm})$$

$$= 28.4 \times 10^{-20} \text{ J} = 28.4 \times 10^{-23} \text{ kJ}$$

A mole of photons contains Avogadro's number (6.023×10^{23}) of photons. We call a mole of photons an **einstein.** An einstein of 700-nm light represents a quantity of energy equal to

$$(28.4 \times 10^{-20} \text{ J/photon})(6.023 \times 10^{23} \text{ photons}) = 171 \text{ kJ}$$

Table 15.2 lists the energies of several other photons.

15.2.2. Absorption of Light

A molecule can exist in numerous energy states defined by the distribution of electrons in orbitals of different energies. Associated with each of these electronic states are various vibrational and rotational substates. When a molecule absorbs energy in the form of light, a photon strikes the molecule and boosts an electron from an orbital of lower energy to one of higher energy. Both the electron and the entire molecule progress from the **ground state** to an **excited state.** Energy differences between orbitals are **quantized.** It takes a specific amount of energy (*quantum*) to raise an electron from one orbital to another.

To be effective, the striking photon must have a *minimum* amount of energy equal to the energy difference between the lower and higher electronic orbitals. If the photon does not have that energy (has a smaller quantum of energy; is of longer wavelength), it will be ineffective in promoting the electron to a higher energy level; the molecule will not absorb light of that wavelength. If the photon has more energy than is required for boosting the electron (has a greater quantum of energy; is of shorter wavelength), the extra amount of energy may be used to increase the kinetic energy of the boosted electron and/or the vibrational and rotational energies of the molecule.

Thus, a given molecule can only absorb photons of specific wavelengths, and those must provide at least sufficient energy to bring about the particular electronic transitions between orbitals. Once a molecule has been electronically excited, it can dissipate its excitation energy in a number of ways. Four common ones involve internal conversion, fluorescence, resonance energy transfer, and electron transfer (Figure 15.3).

In many cases, an excited molecule returns to its ground state by a radiationless transfer of energy called **internal conversion.** In this process, excitation energy is converted to kinetic energy of surrounding molecules; the energy is converted to heat. Internal conversion occurs very rapidly and is completed in less than 10^{-11} s.

At times, an excited molecule *loses only part* of its excitation energy by emitting radiation in the form of **fluorescence.** In this case, the excited molecule emits a photon that is *less energetic* than the one that excited the molecule; the emitted photon has a *longer wavelength* than the exciting photon. Fluorescence is slower than internal conversion and requires about 10^{-8} s for completion.

When excitiation energy is dissipated by **resonance energy transfer,** the excited molecule directly excites an adjacent molecule. Excitation energy is transferred from one molecule to a neighboring one through interaction of the molecular orbitals of the participating molecules. Resonance energy transfer is a radiationless process, also called *exciton transfer.*

Lastly, excitation energy can be dissipated via **electron transfer.** In this radiationless process, the excited electron itself is transferred from one molecule to a neigh-

Table 15.2. Energy Content of Photons

Color	Wavelength (nm)	Energy/einstein (kJ)
Red	700	171.0
Yellow	600	199.5
Blue	500	239.5
Violet	400	299.3
Ultraviolet	300	399.1

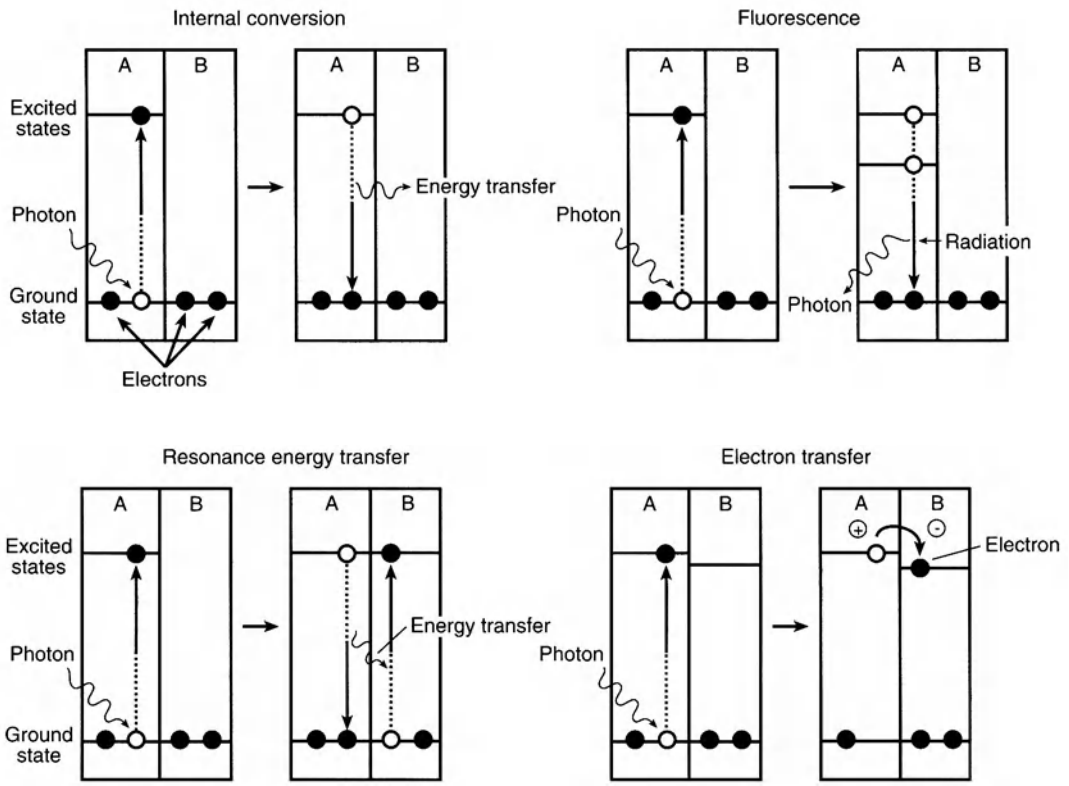


Figure 15.3. Common modes for the dissipation of excitation energy. A and B denote two molecules.

boring one with a slightly lower excited state. The electron donor molecule undergoes oxidation (*photooxidation*) while the acceptor molecule undergoes reduction (*photoreduction*). Electron transfer occurs because the excited electron is bound less tightly to the donor molecule in its excited state than it is to the same molecule in its ground state. Both resonance energy transfer and electron transfer play major roles in the light reactions of photosynthesis.

15.2.3. Quantum Yield

In a photochemical reaction, not all of the incident photons are absorbed. Moreover, as you just saw, excitation energy can be dissipated in different ways, not all of which lead to a chemical reaction. To allow for these factors, we define the efficiency of a photochemical reaction in terms of its quantum yield.

The **quantum yield** of a photochemical reaction *is a ratio equal to the number of molecules (moles) that react or are formed, divided by the number of photons (einsteins) absorbed.* Recall that the biochemical standard free energy change for the complete oxidation of glucose to CO_2 and water is $\Delta G^{\circ\prime} = -2870$ kJ mol^{-1}. In photosynthesis, six moles of CO_2 are reduced. Hence, the amount of energy that has to be expended to reduce *one mole* of CO_2 is $+2870/6 = 478$ kJ.

Photosynthesis uses light that has a wavelength of about 700 nm. Photons of such light have an energy of 171 kJ/einstein. Assuming an efficiency of 100%, three moles of these photons are required to reduce one mole of CO_2, or three photons are required per molecule of CO_2. This represents a quantum yield of 0.33. At an efficiency of 50%, the same reaction would require six photons, resulting in a quantum yield of 0.17. As you will see, photosynthesis actually requires *eight photons* for the reduction of *one molecule of* CO_2. This represents a quantum yield of 0.13 and an efficiency of 37%, which is similar to the efficiency of other metabolic pathways. Remember, however, that using $\Delta G^{\circ\prime}$ provides only an approximate answer and that an accurate evaluation of efficiency requires the use of $\Delta G'$.

15.3. PHOTOSYNTHETIC MACHINERY

15.3.1. Prokaryotes and Eukaryotes

Most photosynthetic organisms use water as an electron donor and produce oxygen. Some organisms use other electron donors, such as H_2S, and do not evolve oxygen

during photosynthesis. These organisms are generally strict anaerobes for which oxygen is toxic.

Both prokaryotes and eukaryotes carry out photosynthesis. Photosynthetic prokaryotes include cyanobacteria, purple and green sulfur bacteria (Table 15.1), and nonsulfur photosynthetic bacteria. Biologists formerly called cyanobacteria blue-green algae and considered them to be plants because of their content of chlorophyll and their evolution of oxygen. Scientists now consider them a class of bacteria. Cyanobacteria use water as an electron donor. Sulfur bacteria use hydrogen sulfide, sulfite, thiosulfate, or other sulfur compounds as electron donors and do not evolve oxygen. Nonsulfur bacteria use a variety of compounds such as hydrogen gas, lactate, succinate, or acetate as reducing agents. In prokaryotes, photosynthesis takes place in **chromatophores,** vesicular structures formed by invaginations of the cell membrane.

Photosynthetic eukaryotes include plants, multicellular algae, dinoflagellates, and diatoms. These organisms all use water as an electron donor, and photosynthesis takes place in specialized subcellular structures called chloroplasts.

15.3.2. Chloroplasts

Chloroplasts are subcellular, membrane-bounded organelles that resemble mitochondria in several ways. Like the mitochondrion, a chloroplast has an *inner* and an *outer membrane,* separated by an *intermembrane space* (Figure 15.4). The inner chloroplast membrane, like the mitochondrial membrane, is nearly impermeable whereas the outer membrane is highly permeable. **Stroma,** the soluble portion of a chloroplast, is analogous to the matrix of mitochondria. Mitochondria and chloroplasts are *semiautonomous;* they have their own DNA and machinery for replicating and expressing this DNA, but functions of both organelles depend on additional products encoded in nuclear DNA. According to the *endosymbiotic theory,* both mitochondria and chloroplasts have evolved from prokaryotic cells (specifically, cyanobacteria) that assumed symbiotic relationships with nonphotosynthetic eukaryotic cells.

Within the chloroplast occur the **grana,** bodies consisting of flattened membranes termed **thylakoid disks.** The membrane of thylakoid disks has an unusual lipid composition. It contains only small amounts of phospholipids but large amounts of glycolipids (about 10% and 90%, respectively). The membrane also contains proteins. Specific multisubunit protein complexes embedded in the thylakoid membrane are termed **reaction centers.** They play a key role in the photosynthetic process and represent sites at which chemical reactions occur. Thylakoid disks form by invaginations of the inner chloroplast membrane and resemble the cristae of mitochondria.

A network of membranes, called **stroma lamellae** or **intergranal lamellae,** connects the grana. Thylakoid disks, stacked into grana, are embedded in the stroma much as the cristae of mitochondria are embedded in the matrix. Typically, each chloroplast contains about 40–80 grana.

Chloroplasts vary in number, size, and shape. There

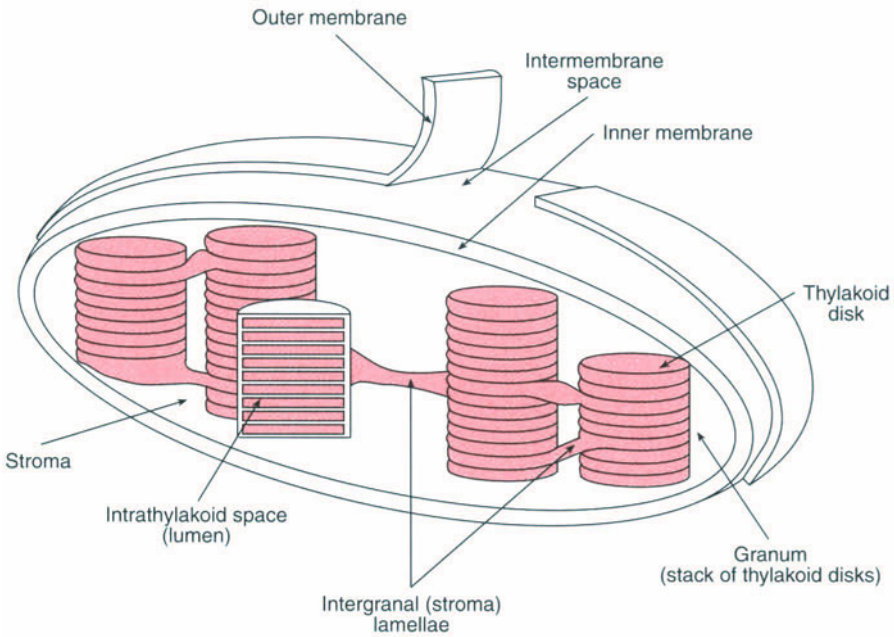


Figure 15.4. Schematic drawing of a chloroplast.

may be from 1 to 1000 per cell. They usually have an elongated and ellipsoidal shape. They have a length of about 5–10 μm and a diameter of about 0.5–2 μm. Chloroplasts are approximately 2–5 times larger than mitochondria. Light trapping and oxygen production (the light reactions) take place in the thylakoid disks. CO_2 fixation (the dark reactions) occurs in the stroma.

15.3.3. Photosynthetic Pigments

15.3.3A. Chlorophylls. The primary light-absorbing pigment is **chlorophyll (Chl).** Major chlorophylls of eukaryotes are **chlorophylls *a*** and ***b,*** both of which occur in plants and algae. Cyanobacteria contain only Chl *a*. Other photosynthetic bacteria have **bacteriochlorophylls *a*** and ***b***. Chlorophylls are planar, metal-containing tetrapyrroles that resemble heme (Figure 15.5). They differ structurally from heme in four respects:

1. Chlorophylls contain magnesium rather than iron; they are magnesium porphyrins rather than iron porphyrins.
2. Chlorophylls contain a cyclopentanone ring (ring V) fused to pyrrole ring III.
3. Chlorophylls contain **phytol,** a long-chain alcohol esterified to the propionyl group of pyrrole ring IV. The nonpolar phytol serves to anchor chlorophyll molecules in the hydrophobic portion of thylakoid membranes.
4. Chlorophylls are more reduced (saturated) than heme. Chlorophylls *a* and *b* contain a partially reduced pyrrole ring IV; bacteriochlorophylls *a* and *b* contain partially reduced pyrrole rings II and IV.

Because of their extensive conjugated double-bond system, chlorophylls absorb light strongly in the visible region. Chlorophylls *a* and *b* have distinct absorption spectra and complement each other in their absorptive properties (Figure 15.6). The composite effect of light absorption by chlorophylls *a* and *b* is responsible for chlorophyll's green color. Bacteriochlorophylls have fewer conjugated double bonds than chlorophylls and have different absorptive properties. They absorb at longer wavelengths, in some cases up to 1100 nm.

15.3.3B. Accessory Pigments. In addition to chlorophylls, several other pigments, termed **accessory pigments,** function in photosynthesis. Two major classes of accessory pigments are the **carotenoids,** which are purple, red, or yellow pigments, and the **phycobilins,**

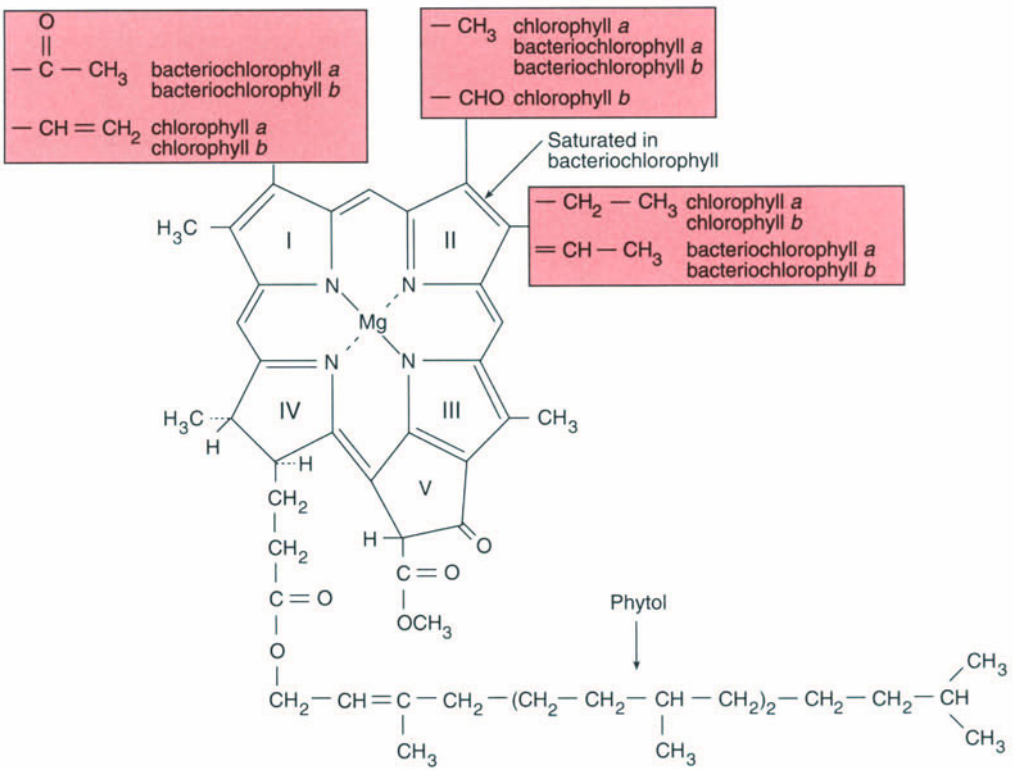


Figure 15.5. Structure of plant and bacterial chlorophylls.

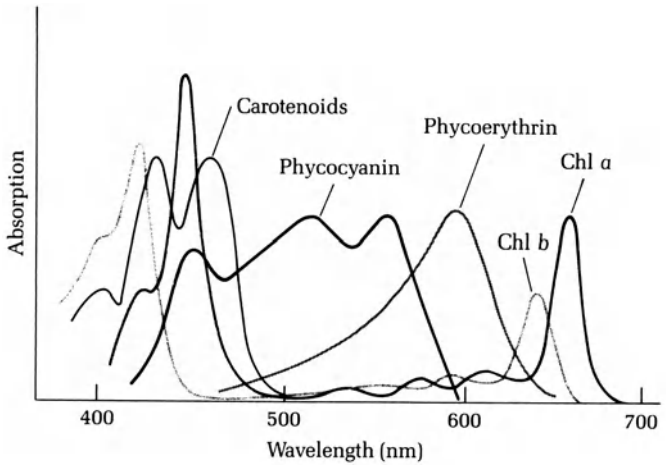


Figure 15.6. Absorption spectra of photosynthetic pigments. Phycocyanins and phycoerythrins are accessory pigments consisting of proteins conjugated to phycobilins. [Reprinted, with permission, from Govindjee and R. Govindjee, *Sci. Am.* 231(6):68–82 (1974). © 1974 by Scientific American, Inc. All rights reserved.]

which are blue or red pigments. In chloroplasts, there occur primarily two types of carotenoids, *carotenes* (precursors of vitamin A; see Figure 6.15) and **xanthophylls** (oxygenated carotenes). Phycobilins are porphyrins in which the ring system has opened up to form a chain of pyrroles. Accessory pigments assist in the transfer of light energy to reaction centers (see below). They absorb light in regions of the spectrum at which chlorophylls *a* and *b* do not absorb well (see Figure 15.6) and extend the spectrum of light that photosynthetic organisms can use. Collectively, chlorophylls and accessory pigments absorb radiant energy across the spectrum of visible light.

15.3.4. Photosystems I and II

Light-absorbing pigments and their associated proteins are assembled into a functional unit called a **photosystem.** Green plants and cyanobacteria possess two such assemblies, **photosystem I (PSI)** and **photosystem II (PSII).** Photosystem I responds to light below 700 nm; photosystem II responds to light below 680 nm. Accordingly, the *reaction centers* of these two photosystems are designated **P700** and **P680,** respectively, where P stands for pigment. Both photosystems function in the light reactions of photosynthesis. Photosystem I leads to production of reducing power in the form of NADPH. Photosystem II leads to evolution of oxygen through *photolysis* (splitting brought about by light) of water.

Absorption of light by these photosystems involves both resonance energy transfer and electron transfer. When a photon is absorbed by a photosystem, it excites a chlorophyll molecule, boosting its electron to a higher energy level. The excited chlorophyll molecule passes its excitation energy by resonance energy transfer to surrounding chlorophyll molecules. The excitation energy passes in stepwise fashion and in a random manner from one chlorophyll molecule to another. We call the molecules that serve as conduits for the excitation energy **antenna chlorophylls.** Much as antennas function to receive incoming radio or TV signals, so antenna chlorophylls serve to gather the incoming energy of light.

Ultimately, the excitation energy becomes channeled by antenna chlorophylls to a reaction center (Figure 15.7), where it brings about an excitation of chlorophyll molecules called **specialized chlorophylls.** Reaction centers containing excited specialized chlorophylls are designat-

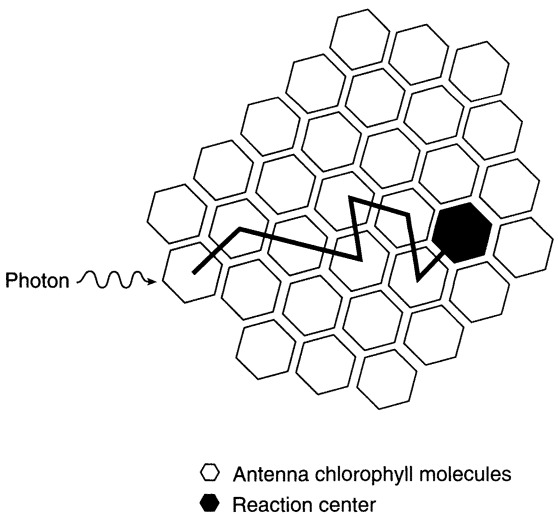


Figure 15.7. Diagram showing the flow of energy through a series of antenna chlorophylls to a photosynthetic reaction center.

ed **P700*** and **P680*,** respectively. Specialized chlorophylls are chemically identical to antenna chlorophylls but have different properties because of their location and immediate environment. They have *lower* excited-state energies than antenna chlorophylls and, therefore, *trap* the excitation energy and prevent it from being passed along by further resonance energy transfer. Instead, the excitation energy is dissipated by electron transfer. A specialized chlorophyll molecule transfers its excited electron to an appropriate acceptor, thereby reducing the acceptor while the specialized chlorophyll is converted to a cation (Chl^+). The chlorophyll ultimately returns to its ground state by replacing its lost electron from another source.

Most chlorophyll molecules are antenna chlorophylls. There are usually several hundred antenna chlorophylls for every one specialized chlorophyll. Antenna chlorophylls transfer excitation energy with great efficiency (over 90%) and convey it rapidly (in less than 10^{-10} s) to the reaction center. Once the excitation energy reaches the specialized chlorophylls in the reaction center, the actual photochemistry of the light reactions begins.

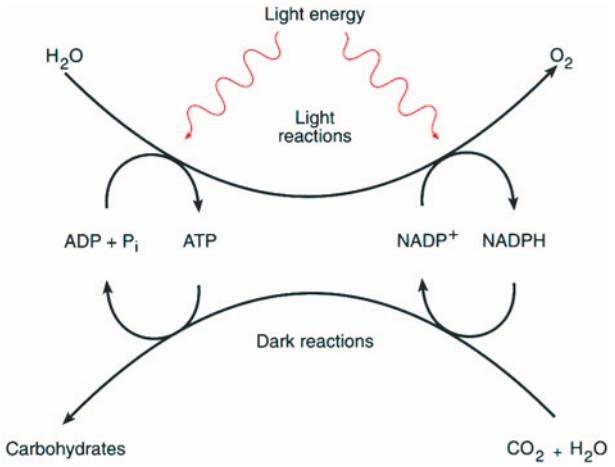


Figure 15.8. The light and dark reactions of photosynthesis. The former include water photolysis, ATP synthesis, and NADPH production. The latter convert CO_2 to carbohydrate, using ATP and NADPH formed by the light reactions.

15.4. LIGHT REACTIONS

The **light reactions** include those parts of the photosynthetic process that are directly dependent on light and do not proceed without it. All components of the light reactions—light-absorbing pigments, electron carriers, and ATP synthase—are located in the thylakoid membrane. The light reactions produce NADPH and ATP, both of which subsequently provide the energy to "fix" CO_2 in the dark reactions (Figure 15.8). The light reactions comprise four individual reactions, which we will discuss in the order listed:

- *Photooxidation of chlorophyll*—photochemical excitation of chlorophyll
- *Photoreduction of* $NADP^+$—production of NADPH
- *Photooxidation of water*—splitting of water (photolysis); evolution of O_2
- *Photosynthetic phosphorylation*—synthesis of ATP

In plants and cyanobacteria, we describe the set of reactions by an energy diagram called the **Z-scheme** (because of its resemblance to the letter Z, looked at sideways; Figure 15.9). In this scheme, photosystems I and II are linked in series.

15.4.1. Photooxidation of Chlorophyll

In a *photochemical excitation* step, photosystem I absorbs a photon of light that boosts an electron to a higher energy level, thereby forming a molecule of excited chlorophyll. The excitation energy passes, by resonance energy transfer and via antenna chlorophylls, to a reaction center (P700), where it leads to formation of an excited specialized chlorophyll (P700*). The P700* loses its excited electron by electron transfer to an appropriate acceptor, believed to be a chlorophyll molecule (A_0). In the process, the acceptor is reduced.

Recall that a half-reaction with a smaller reduction potential represents a stronger reducing agent than one with a larger reduction potential (Section 12.1). Thus, P700 is a weaker reducing agent than A_0 (see Fig. 15.9). Ordinarily, P700 could not reduce A_0; on the contrary, A_0 would reduce P700. Yet here P700 is made to reduce A_0. This constitutes an "uphill" reduction in which a weaker reducing agent (P700) is made to reduce a stronger one (A_0). The "uphill" reduction occurs only because the energy of the absorbed photon excites P700 to P700*, and this *excited form of chlorophyll is a stronger reducing agent* than A_0. The reduction potential of P700* is smaller than that of A_0, and hence P700* can reduce A_0.

15.4.2. Photoreduction of $NADP^+$

From the reduced electron carrier (A_0), the electron passes through a chain of electron carriers constituting a *photosynthetic electron transport system.* In principle, the coupled oxidation–reduction reactions are similar to those of the mitochondrial electron transport system.

Phylloquinone (*vitamin* K_1) is the electron carrier (A_1) that follows A_0 (Figure 15.9). From there, the elec-

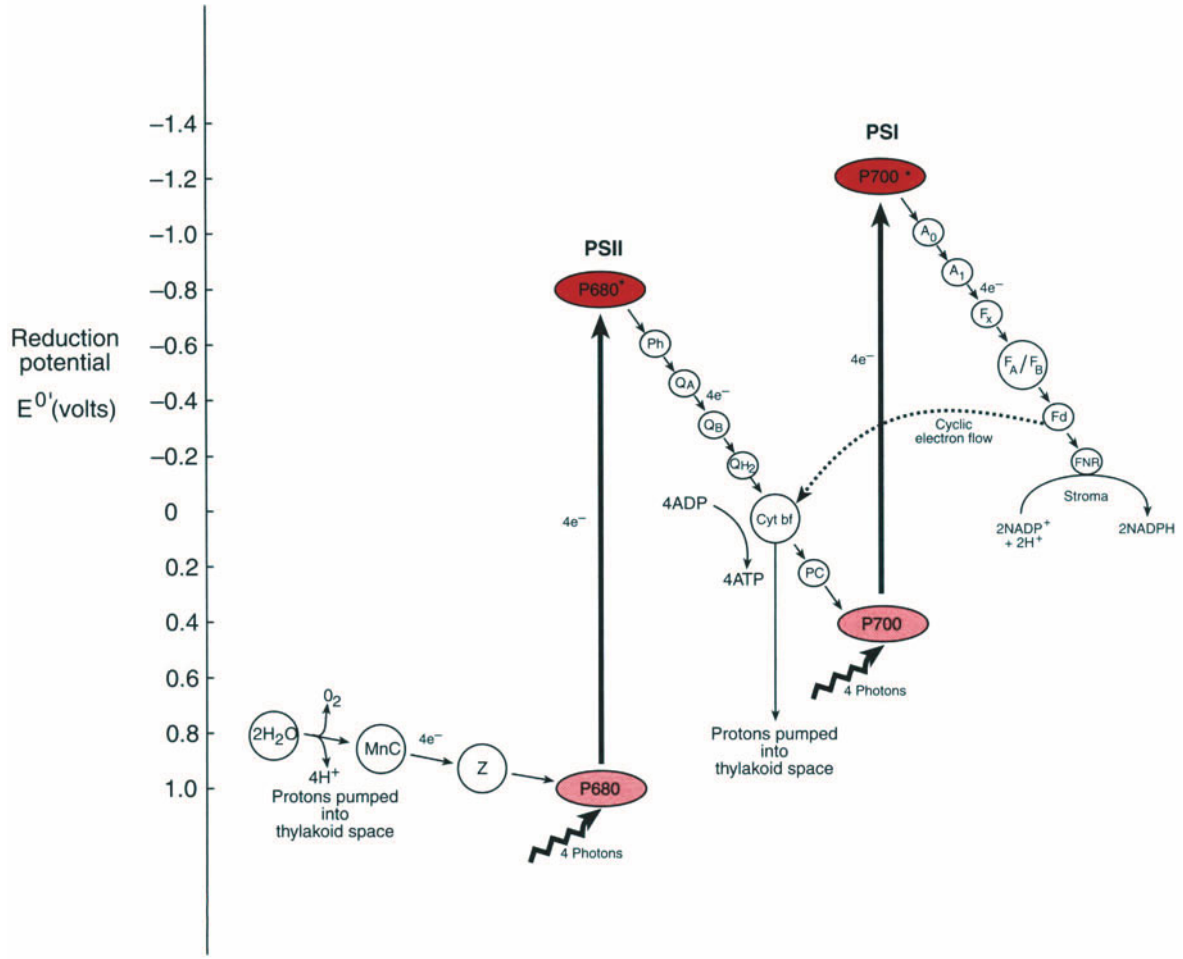


Figure 15.9. The Z-scheme of photosynthesis. Photosystems I and II (PSI and PSII) are linked in series and include two "uphill" and three "downhill" reductions.

tron passes through a series of three *membrane-bound ferredoxins* (F_x, F_A, F_B). Ferredoxins are nonheme, iron–sulfur proteins that have low reduction potentials. From membrane-bound ferredoxins, the electron passes to a *soluble* ferredoxin (Fd) and from there to *Fd:NADP⁺ reductase* (FNR). This enzyme is a flavoprotein, containing FAD as a prosthetic group. The enzyme catalyzes reduction of $NADP^+$ by transfer of a hydride ion (H^-) from $FADH_2$ (see Section 11.1). $NADP^+$ serves as the terminal electron acceptor in this photosynthetic electron transport system and is reduced to NADPH:

$$NADP^+ + H^- \rightarrow NADPH$$

All of the steps from P700* to $NADP^+$ constitute ordinary "downhill" reductions. Each carrier has a smaller reduction potential than the one following it, so each carrier reduces the subsequent one in the series.

15.4.3. Photooxidation of Water

Photooxidation of chlorophyll results in formation of an "electron hole" in the reaction center P700—an electron from the center was used to reduce $NADP^+$. The lost electron must be replaced by obtaining an electron from some other source. That source is water, the splitting of which produces electrons:

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

However, although water constitutes the ultimate source of the electrons, the splitting of water is carried out in an indirect way involving photosystem II. Water splitting is coupled to absorption of a photon (*photolysis*) by photosystem II. The complete photosynthetic system, therefore, requires *absorption of two photons, one by each of the two photosystems.*

Absorption of a photon by photosystem II, as in photosystem I, leads to production of an excited specialized chlorophyll, in this case P680*. P680* loses its electron, via electron transfer, to a *pheophytin (Ph).* Pheophytins are identical to chlorophylls in their structure but have the centrally bound Mg^{2+} replaced by two protons. The reduction of pheophytin by P680 constitutes another "uphill" reduction, again made possible by absorption of a photon that excites P680 to P680*.

From pheophytin, the electron passes through a set of electron carriers constituting a second photosynthetic electron transport system and ultimately fills the electron hole in photosystem I. After pheophytin, the electron passes through a series of *plastoquinones* (Q), compounds structurally similar to coenzyme Q (ubiquinone). A *cytochrome/iron–sulfur complex (cytochrome bf complex)* constitutes the next electron carrier in the sequence. The complex consists of *cytochrome* b_6, *cytochrome f,* and *iron–sulfur clusters.* From the cytochrome/iron–sulfur complex, the electron passes via *plastocyanin (PC)* to P700. Plastocyanin is a copper-containing peripheral membrane protein located on the surface of the thylakoid membrane. All steps from P680* to P700 are ordinary "downhill" reductions. Transfer of the electron from plastocyanin to P700 effectively fills the electron hole in that reaction center.

However, loss of an electron from P680* produces an electron hole in photosystem II. This hole is filled by an electron derived from the splitting of water. Thus, water constitutes the ultimate source of the electrons used to replenish those lost by photochemical excitations resulting from the absorption of two photons. A *manganese-containing protein complex (MnC)* mediates water photolysis. The complex contains four protein-bound manganese ions that cycle through a series of oxidation states comprising various combinations of Mn^{3+} and Mn^{4+}. The sequence of steps allows for the formation of O_2 without generating hazardous partially reduced forms of oxygen. Water photolysis, like the reverse reaction catalyzed by cytochrome oxidase, is a four-electron reaction. The MnC-catalyzed photolytic sequence constitutes an oxidation, involving the loss of four electrons. By contrast, the cytochrome oxidase catalyzed conversion of oxygen to water constitutes a reduction, involving the gain of four electrons.

From MnC, the electron transfers to an electron carrier designated Z and identified recently as a tyrosine residue located on a polypeptide chain of phototosystem II. From Z, the electron passes to P680. Transfer of the electron from water to P680 is an ordinary "downhill" reduction.

Note that both photosystems carry out oxidation–reduction reactions but with different end results. Photosystem I generates a strong reducing agent (P700*) capable of reducing $NADP^+$. Photosystem II generates a strong oxidizing agent (P680) capable of oxidizing water. (In Figure 15.9, four electrons flow through the Z-scheme. We will explain the reason for this below.)

15.4.4. Photosynthetic Phosphorylation

Operation of the electron transport system that links photosystems I and II is *coupled* to synthesis of ATP. We refer to this mode of ATP synthesis as **photosynthetic phosphorylation** or **photophosphorylation.** It represents a third mode of ATP synthesis that differs from *oxidative phosphorylation* and *substrate-level phosphorylation,* discussed earlier. In principle, the mechanism of photophosphorylation is similar to that of oxidative phosphorylation. In both cases, electron transport and ATP synthesis become coupled by means of an electrochemical proton gradient (*chemiosmotic coupling*) across a membrane containing *ATP synthase* (Figure 15.10). However, photophosphorylation differs from oxidative phosphorylation in the way in which the electrochemical gradient is generated, in the type and sequence of electron carriers, and in the dependence on absorption of light.

In photophosphorylation, the transmembrane proton gradient that drives ATP synthesis develops due to two reactions that yield protons. One is the photolysis of water:

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

The other involves the cytochrome *bf* complex that reacts with plastoquinone, which shuttles between an oxidized and a reduced form (Q and QH_2), much like CoQ and $CoQH_2$:

$$QH_2 + 2(\text{Cyt } bf)_{ox} \rightleftarrows Q + 2(\text{Cyt } bf)_{red} + 2H^+$$

These proton-producing reactions pump protons into the *thylakoid space (intrathylakoid space; fluid-filled lumen).* As a result, the pH decreases (high $[H^+]$) in the thylakoid space and increases (low $[H^+]$) in the stroma. The difference that develops amounts to about 2–3 pH units. Due to the gradient, protons move from the thylakoid space into the stroma through the membrane-bound ATP synthase, thereby driving ATP synthesis. Much like the mitochondrial ATP synthase (F_0-F_1 ATPase), the chloroplast enzyme consists of two factors and is also known as **CF_0-CF_1 ATPase,** where C stands for chloroplast.

We can demonstrate the chemiosmotic coupling of

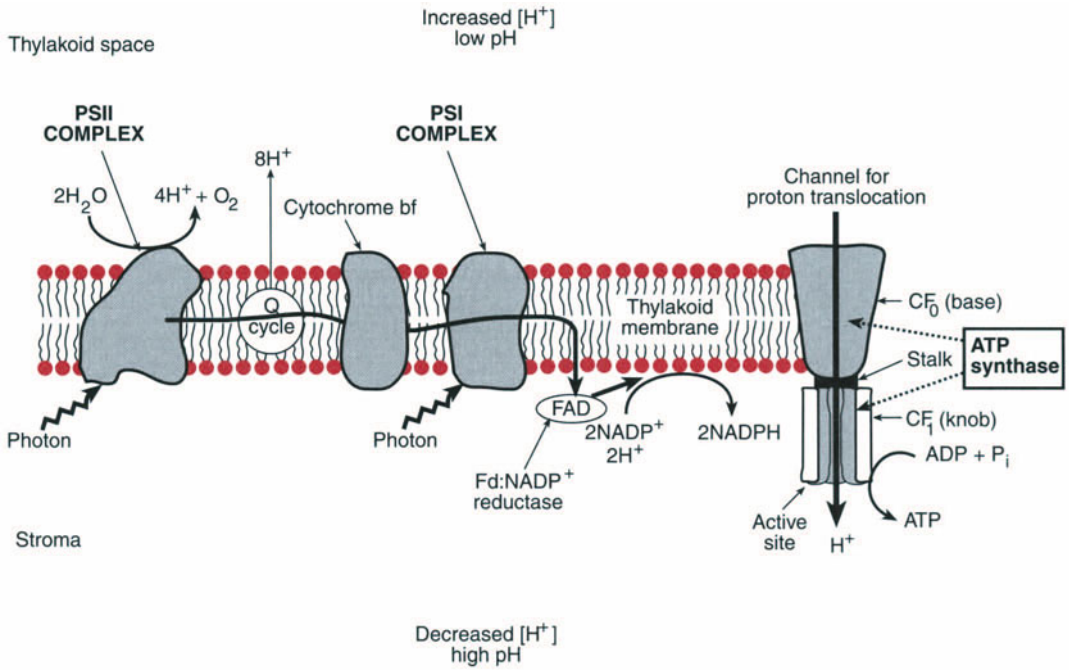


Figure 15.10. Localization of Z-scheme components and ATP synthase in the thylakoid membrane. PSI, PSII, and the cytochrome *bf* complex constitute three large assemblies interconnected by mobile carriers, plastocyanins (PC) and plastoquinones (Q). [Adapted from D. R. Ort and N. E. Good, *Trends Biochem. Sci.* 13:467–469 (1988) with kind permission of Elsevier Science-NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.]

photophosphorylation experimentally. We first soak chloroplasts for several hours in a buffer at pH 4 so that their internal pH becomes stabilized at that value. Then we change the external pH abruptly by transferring the chloroplasts to a buffer at pH 8, containing both ADP and P_i. When we do this, there occurs a sudden burst of ATP synthesis accompanied by disappearance of the pH gradient across the chloroplast membrane. ATP synthesis from ADP and P_i proceeds in the dark; no light is required. This experiment was originally performed by André Jagendorf in 1966. The results provide strong support for the general chemiosmotic coupling hypothesis.

Electron carriers of the photosynthetic electron transport system are largely organized in three complexes that resemble the clustering of mitochondrial electron carriers to form respiratory complexes. The three complexes (photosystem I, photosystem II, and the cytochrome *bf* complex) need not be very close to each other because they are effectively interconnected by highly mobile electron carriers, plastoquinones and plastocyanins. Photosystem I and ATP synthase are located primarily in the stroma lamellae (unstacked regions) while photosystem II is located almost exclusively in the grana (stacked regions). The cytochrome *bf* complex is distributed throughout both the stacked and unstacked regions.

15.4.5. Balance Sheet of the Light Reactions

Reduction of $NADP^+$ to NADPH requires transfer of a hydride ion, equivalent to a proton plus *two electrons:*

$$H^- = H^+ + 2e^-$$

Accordingly, *two* photons must be absorbed by photosystem I. Since reduction of $NADP^+$ results in *two* "electron holes" in photosystem I, *two* electrons must be derived from photosystem II to fill these holes. Thus, photosystem II must also absorb *two* photons, and a total of *four* photons is required to bring about the reduction of one molecule of $NADP^+$.

However, four photons cannot account for the electron balance imposed by water, which constitutes the ultimate electron source for filling the "electron holes" in the two photosystems. Water oxidation must yield molecular oxygen (O_2, not $\frac{1}{2}O_2$) so that two molecules of water must be oxidized by photolysis, releasing four electrons:

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

Therefore, *four* photons must be absorbed by photosystem I, and *four* photons must be absorbed by photosys-

tem II, resulting in the reduction of *two* molecules of $NADP^+$. Consequently, Figure 15.9 shows a flow of *four* electrons through the Z-scheme, so that *eight photons must be absorbed for every two molecules of $NADP^+$ reduced.*

Photosystem I:

$$2NADP^+ + 2H^- + 2H^+ \xrightarrow{4h\nu} 2NADPH + 2H^+$$
$$(4H^+ + 4e^-)$$

Photosystem II:

$$2H_2O \xrightarrow{4h\nu} 4H^+ + 4e^- + O_2$$

Overall reaction for photosystems I and II:

$$2NADP^+ + 2H_2O \xrightarrow{8h\nu} 2NADPH + 2H^+ + O_2$$

The overall reaction is strongly endergonic, having a negative $\Delta E^{\circ\prime}$ and a positive $\Delta G^{\circ\prime}$:

$$2NADP^+ + 4H^+ + 4e^- \rightarrow 2NADPH + 2H^+ \quad (1)$$
$$2H_2O \rightarrow 4H^+ + 4e^- + O_2 \quad (2)$$
$$2NADP^+ + 2H_2O \rightarrow 2NADPH + 2H^+ + O_2$$
$$E_1^{\circ\prime} = -0.32\ V$$
$$E_2^{\circ\prime} = -0.82\ V$$
$$\Delta E^{\circ\prime} = -1.14\ V$$
$$\Delta G^{\circ\prime} = +220\ kJ\ mol^{-1}$$

This reaction cannot go in the absence of light and is made feasible only by absorption of photons, resulting in "uphill" reductions. The light energy absorbed drives the endergonic synthesis of NADPH and the endergonic photolysis of water.

15.4.6. Efficiency of the Light Reactions

15.4.6A. NADPH Formation. We can evaluate the efficiency of the light reactions by considering the two products formed, NADPH and ATP. As you just saw, the energy expenditure for synthesis of a mole of NADPH in the light reactions is 220 kJ. Therefore, the energy required to form two moles of NADPH equals 440 kJ. This energy is provided by absorption of eight moles of photons (eight einsteins). Each einstein has an energy of about 171 kJ so that the total absorbed energy is

$$8\ \text{einsteins} \times 171\ \text{kJ/einstein} = 1368\ \text{kJ}$$

This constitutes a more than sufficient amount of energy to drive NADPH formation. On the basis of these values, reduction of $NADP^+$ to NADPH has an efficiency of

$$\frac{440\ \text{kJ}}{1368\ \text{kJ}} \times 100 \approx 32\%$$

similar to that calculated for CO_2 fixation (see Section 15.2) and other metabolic pathways.

15.4.6B. ATP Production. To evaluate the efficiency of ATP synthesis, we can proceed in two ways. First, let us consider the entire process comprising the light reactions. Experiments indicate that translocation of about 12 H^+ is associated with evolution of one molecule of oxygen (O_2) and synthesis of four molecules of ATP. Thus, photophosphorylation leads to

$$4ADP^{3-} + 4P_i^{2-} + 4H^+ \rightarrow 4ATP^{4-} + 4H_2O$$

When you combine this equation with those given above for photosystems I and II, you obtain an overall equation that sums up all four of the composite light reactions:

$$2NADP^+ + 4ADP^{3-} + 4P_i^{2-} + 2H^+ \xrightarrow{8\ h\nu}$$
$$2NADPH + 4ATP^{4-} + 2H_2O + O_2$$

To calculate the efficiency of ATP synthesis, recall that ATP hydrolysis has a $\Delta G^{\circ\prime}$ of $-30.5\ kJ\ mol^{-1}$. According to the overall equation, four molecules of ATP are synthesized for every eight photons absorbed. In other words, energy equivalent to four *moles* of ATP ($4 \times 30.5 = 122$ kJ) becomes trapped when light energy equivalent to eight *moles* of photons ($8 \times 171 = 1368$ kJ) is expended. This indicates an efficiency of energy conservation of approximately 9%, low in comparison with that of other metabolic pathways.

However, free energy has also been trapped in the form of NADPH synthesis. The free energy *potentially derivable* from oxidation of NADPH via the mitochondrial electron transport system is like that for oxidation of NADH, and oxidation of one molecule of NADPH generates three molecules of ATP. Because the overall reaction yields 2 NADPH, an additional 6 ATP could be produced per molecule of O_2 evolved. Accordingly, 8 moles of photons potentially yield a total of 10 moles of ATP, resulting in an efficiency of

$$\frac{10\ \text{mol} \times 30.5\ \text{kJ/mol}}{1368\ \text{kJ}} \times 100 \approx 22\%$$

For our second estimate of the efficiency of ATP synthesis, we consider just the chain of electron carriers that links photosystems I and II. The potential drop of this system amounts to about 1.2 V per $2e^-$ transferred (see Figure 15.9). Because the drop is traversed twice, by $2e^-$ in each pass, the total potential used equals 2.4 V. Recall that a potential drop of 0.16 V is required for synthesis of ATP when $2e^-$ pass through a chain of electron carriers (Section 12.3). Accordingly, ATP synthesis represents a trapping of $4 \times 0.16 = 0.64$ V or an efficiency of

$$\frac{4 \times 0.16 \text{ V}}{2.4 \text{ V}} \times 100 \approx 27\%$$

On this basis, photophosphorylation appears to be somewhat less efficient than oxidative phosphorylation.

15.4.7. Cyclic Electron Flow

We refer to the flow of electrons through the Z-scheme of photosynthesis as *noncyclic electron flow,* and to ATP synthesis coupled to it as **noncyclic photophosphorylation.** The two constitute the normal mode of operation during photosynthesis. A different mode of operation can, however, occur. It involves a cyclic electron flow in which the excited electron returns to the molecule from which it originated (Figure 15.11).

In *cyclic electron flow,* the excited electron is first transferred from P700* to the initial acceptor (A_0) as before. Following that transfer, the electron passes through the standard noncyclic electron flow carriers up to ferredoxin (Fd). From ferredoxin, the electron is shunted via the cytochrome *bf* complex and plastocyanin to P700 rather than to $NADP^+$. This cyclic set of reactions in-

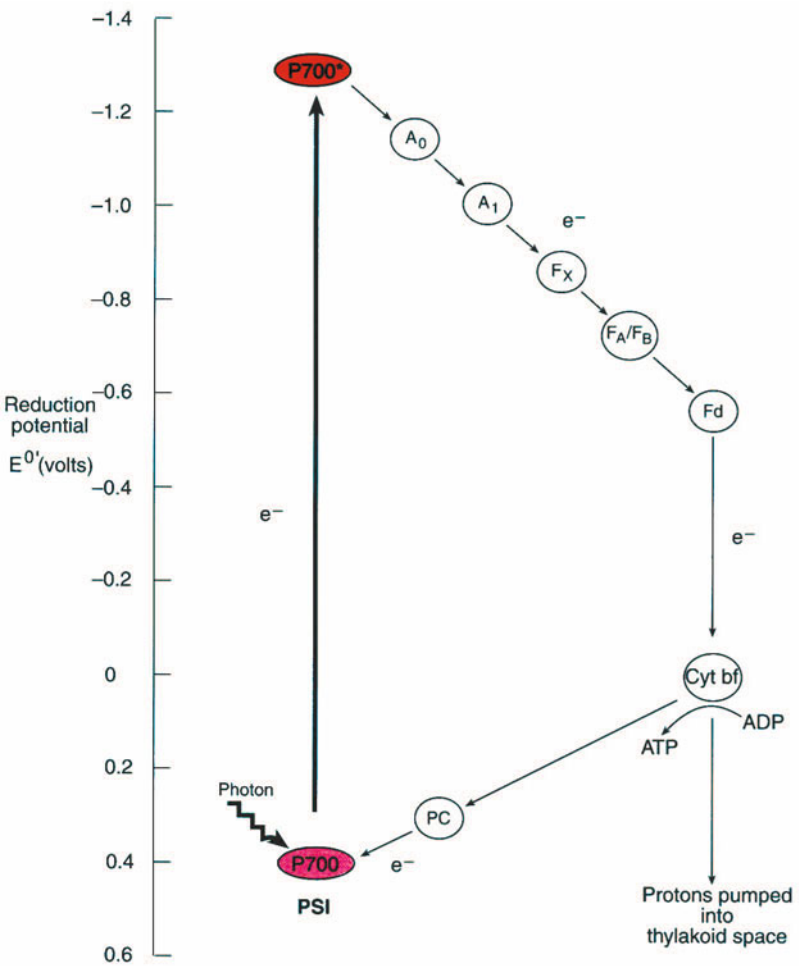


Figure 15.11 Cyclic photophosphorylation. The process involves only photosystem I, features no net oxidation or reduction, and results in ATP synthesis.

volves *only photosystem I* but uses electron carriers from both photosynthetic electron transport systems and leads to production of ATP.

We refer to ATP synthesis coupled to cyclic electron flow as **cyclic photophosphorylation.** Investigators believe it to be somewhat more efficient than noncyclic photophosphorylation and to result in synthesis of two molecules of ATP for every four protons translocated and three photons absorbed:

$$2ADP^{3-} + 2\,P_i^{2-} + 2H^+ \xrightarrow{3\,h\nu} 2\,ATP^{4-} + 2H_2O$$

Cyclic photophosphorylation occurs when cells have a high NADPH/$NADP^+$ ratio but require ATP for various metabolic reactions. Since there exists insufficient $NADP^+$ to accept electrons from reduced ferredoxin, ATP is generated by the cyclic electron flow. Cyclic photophosphorylation may also help to maintain the balance between production of NADPH and ATP. As you will see shortly, the dark reactions require large quantities of ATP. Under some conditions, noncyclic photophosphorylation may not provide enough ATP to drive both the dark reactions and other ATP-requiring reactions of the cell. In that case, cyclic photophosphorylation may supply the additional needed ATP.

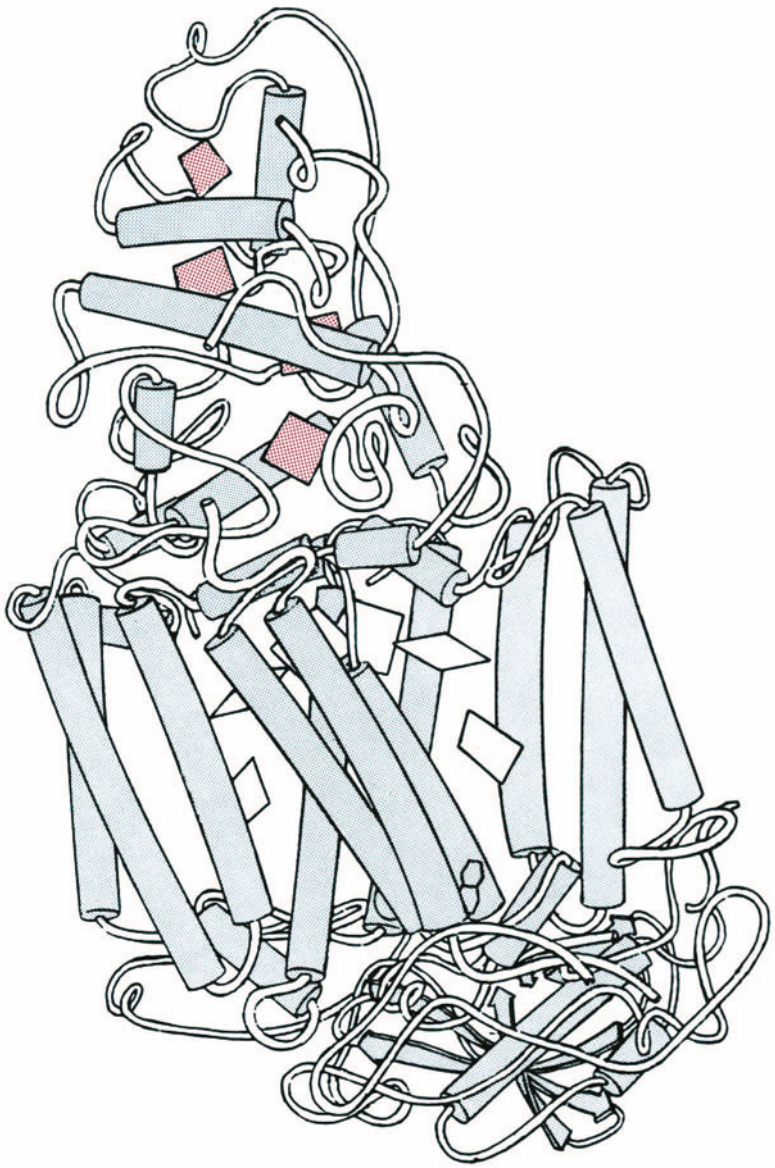


Figure 15.12. Schematic drawing of the photosynthetic reaction center of *Rhodopseudomonas viridis*. Cylinders represent α-helices (11 are transmembrane), and flat arrows represent β-sheets. Dark rectangles designate heme groups, and gray rectangles designate four bacteriochlorophylls and two bacteriopheophytins. (Courtesy of Dr. J. S. Richardson.)

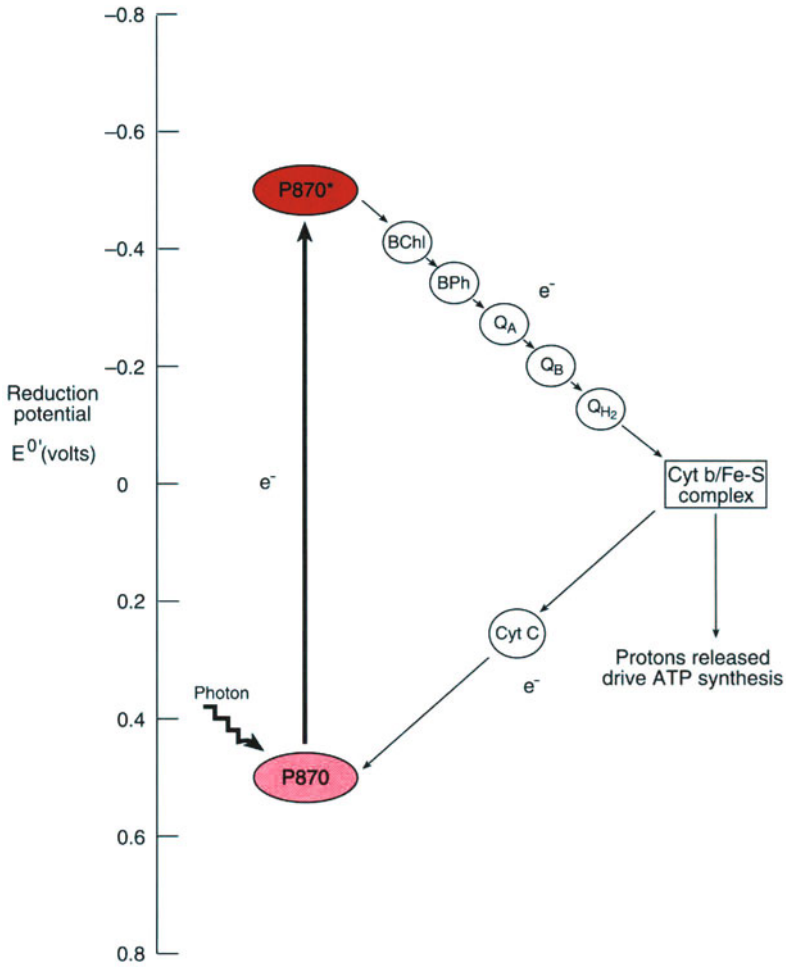


Figure 15.13. Cyclic electron flow in *Rhodopseudomonas viridis*. BChl, Bacteriochlorophyll; BPh, bacteriopheophytin.

Photosynthetic bacteria that do not use oxygen also use a cyclic electron flow. The most extensively studied system is that of the purple sulfur bacterium *Rhodopseudomonas viridis*. The molecular structure of its reaction center was elucidated by Johann Deisenhofer, Hartmut Michel, and Robert Huber in 1984 (Nobel Prize, 1988). The center contains two specialized bacteriochlorophyll molecules ("special pair") embedded in the hydrophobic region of a transmembrane protein and surrounded by other bacteriochlorophylls and accessory pigments (Figure 15.12). A cytochrome, carrying four heme groups, is bound to the reaction center on the external side of the cell membrane (in the periplasmic space).

The special pair becomes excited collectively by light of longer wavelengths (870 nm); hence it is called P870 (Figure 15.13). Loss of an electron by the excited special pair (P870*) to a pheophytin occurs very rapidly, within about four picoseconds (4×10^{-12} s). The extremely rapid transfer prevents the electron from returning to its ground state. It appears that a molecule of bacteriochlorophyll functions in the transfer between P870* and pheophytin.

From pheophytin, the electron passes through a series of plastoquinones and from there to a cytochrome/iron–sulfur complex similar to that linking photosystems I and II in plants. The electron then returns to P870 via cytochrome *c*. ATP synthesis is associated with the cyclic electron flow.

As opposed to plants and cyanobacteria, which obtain their reducing equivalents from photolysis of water and convert them to usable forms by producing NADPH, photosynthetic bacteria must obtain their reducing equivalents from other sources. This they achieve by using inorganic compounds (H_2S, HSO_3^-, $Na_2S_2O_3$) or organic compounds (succinate, lactate, acetate) as reducing agents.

15.5. DARK REACTIONS

The **dark reactions** follow the light reactions. The term "dark reactions" is somewhat misleading. It does not mean that the reactions occur only in the dark; rather, the term emphasizes that these reactions *do not require direct participation of light.* In the dark reactions, or **CO_2 fixation,** CO_2 is converted to carbohydrate. The process consists of a large number of reactions that take place in the stroma and use the ATP and NADPH produced by the light reactions (Figure 15.8). Light, although not directly required, affects the dark reactions since it activates several of the enzymes involved in CO_2 fixation (see Section 15.5.3). In simplified fashion, the following (unbalanced) equation describes the dark reactions:

$$6CO_2 \xrightarrow[\text{enzymes}]{\text{NADPH, ATP}} C_6H_{12}O_6$$

The path that the carbon of CO_2 takes during the dark reactions was elucidated by Melvin Calvin (Nobel Prize, 1961), James Bassham, and Andrew Benson between 1945 and 1953 and is known as the **Calvin cycle** or **reductive pentose phosphate cycle.** These investigators determined the initial steps of the cycle by passing a suspension of unicellular green algae *(Chlorella)* through an illuminated glass tube. They injected radioactive carbon dioxide ($^{14}CO_2$) into the tube and ran the emerging suspension into hot alcohol to inactivate enzymes and stop the reaction. They varied the time elapsed between injecting $^{14}CO_2$ and stopping the reaction to allow the labeled carbon to appear in a larger or smaller number of metabolic intermediates.

After stopping the reaction, the researchers extracted the algae and analyzed the extract by two-dimensional paper chromatography and autoradiography. When they allowed the suspension to carry out photosynthesis for a minute or longer, the extract contained a complex mixture of many ^{14}C-labeled compounds including carbohydrates, amino acids, lipids, and nucleotides. When they reduced the reaction time to 30 s, the extract contained only a limited number of labeled compounds, and when they decreased the time allowed for photosynthesis to 5 s, the extract yielded an autoradiogram that contained only a single spot. Analysis of the material corresponding to this spot showed it to be *3-phosphoglycerate.* The experimenters concluded that 3-phosphoglycerate must be the first *stable* compound formed in the dark reactions of photosynthesis.

15.5.1. Ribulose 1,5-Bisphosphate Carboxylase (Rubisco)

Subsequent work showed that the initial acceptor of carbon dioxide is a five-carbon compound, *ribulose 1,5-bisphosphate,* which is cleaved to two three-carbon compounds (3-phosphoglycerate) so that the flow of carbon is given by

$$C_5 + C\ (CO_2) \longrightarrow 2C_3$$

The reaction whereby CO_2 combines with ribulose 1,5-bisphosphate to produce two molecules of 3-phosphoglycerate is catalyzed by **ribulose 1,5-bisphosphate carboxylase (rubisco)** and proceeds through formation of an enediol intermediate (Figure 15.14).

Ribulose 1,5-bisphosphate carboxylase of plants and most photosynthetic microorganisms is a large oligomer,

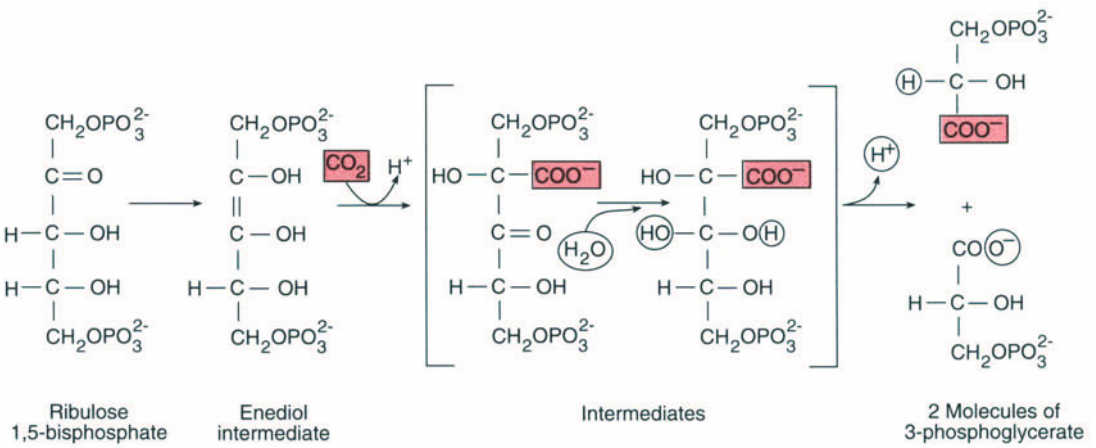


Figure 15.14. The reaction catalyzed by ribulose 1,5-bisphosphate carboxylase (rubisco) when the enzyme functions as a carboxylase in CO_2 fixation.

having a molecular weight of about 560,000. It consists of 16 subunits, eight large ones (MW = 56,000 each) and eight small ones (MW = 14,000 each). The large and small subunits are specified, respectively, by a gene of chloroplast DNA and by a gene of nuclear DNA. Each large subunit contains both a catalytic and a regulatory site. The small subunits enhance the activity of the large subunits.

Because so much photosynthesis occurs all over the globe, scientists consider rubisco to represent the most abundant protein on Earth. The enzyme typically accounts for more than 50% of soluble leaf proteins and constitutes about 15% of all chloroplast proteins. Researchers have estimated that there are about 40×10^6 tons of rubisco in the world and that the enzyme is being synthesized at a rate of about 4×10^{13} g per year! Because the enzyme is so abundant, some scientists have proposed using it as a dietary protein supplement.

15.5.2. Calvin Cycle

We commonly divide the Calvin cycle into two phases (Figure 15.15):

A. *Production phase*
   1. Carboxylation
   2. Phosphorylation
   3. Reduction
   4. Carbohydrate formation

B. *Regeneration phase*

15.5.2A. Production Phase. The **carboxylation stage,** the first part of the production phase, consists of the rubisco-catalyzed reaction whereby ribulose 1,5-bisphosphate is carboxylated to form two molecules of 3-phosphoglycerate.

Carboxylation is followed by a **phosphorylation stage** that requires ATP and in which 3-phosphoglycerate becomes converted to *1,3-bisphosphoglycerate.* A **reduction stage** follows the phosphorylation and leads to reduction of 1,3-bisphosphoglycerate to *glyceraldehyde 3-phosphate* by means of NADPH. Both the ATP and the NADPH required in these reactions are supplied by the light reactions. We can summarize the first three stages of the production phase as follows (Figure 15.16):

$$\text{Ribulose 1, 5-bisphosphate}^{4-} + CO_2 + H_2O \rightarrow 2\,(\text{3-phosphoglycerate}^{3-}) + 2H^+$$

$$\text{3-Phosphoglycerate}^{3-} + ATP^{4-} \rightarrow \text{1, 3-bisphosphoglycerate}^{4-} + ADP^{3-}$$

$$\text{1, 3-Bisphosphoglycerate}^{4-} + NADPH + H^+ \rightarrow \text{glyceraldehyde 3-phosphate}^{2-} + NADP^+ + P_i^{2-}$$

In order to fix enough CO_2 for synthesis of one molecule of hexose, the first reaction must be multiplied by 6 and the remaining two by 12. Thus, 6 ribulose 1,5-bisphosphate, 6 CO_2, and 6 H_2O react to yield 12 glyceraldehyde 3-phosphate, 12 $NADP^+$, and 12 P_i.

After formation of 12 molecules of glyceraldehyde 3-phosphate, the path of carbon is split. Two molecules of

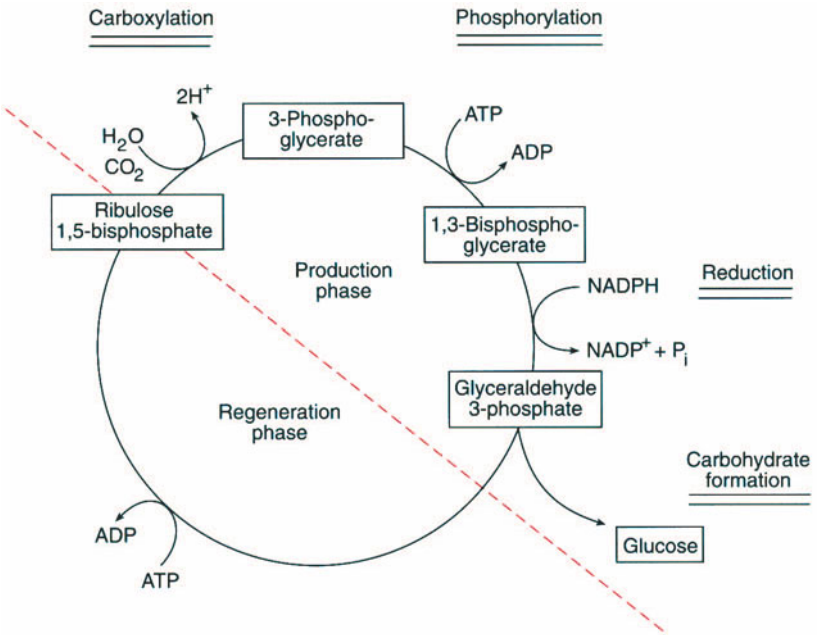


Figure 15.15. The two major parts of the Calvin cycle—a production phase and a regeneration phase.

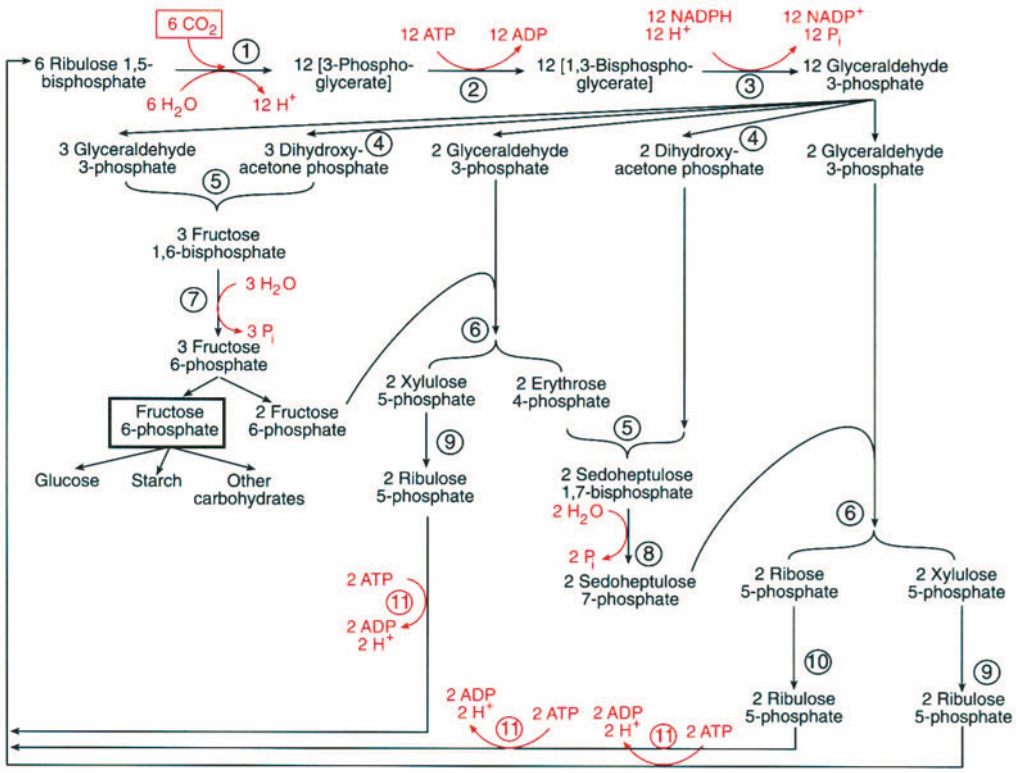


Figure 15.16. Reactions of the Calvin cycle. Six molecules of ribulose 1,5-bisphosphate yield 12 molecules of glyceraldehyde 3-phosphate. Of these, two form one molecule of fructose 6-phosphate, and 10 regenerate six molecules of ribulose 1,5-bisphosphate. Numbers designate enzymes: ① Ribulose 1,5-biphosphate carboxylase; ② phosphoglycerate kinase; ③ glyceraldehyde 3-phosphate dehydrogenase; ④ triose-phosphate isomerase; ⑤ aldolase; ⑥ transketolase; ⑦ fructose 1,6-bisphosphatase; ⑧ sedoheptulose 1,7-bisphosphatase; ⑨ ribulose 5-phosphate epimerase; ⑩ ribulose 5-phosphate isomerase; ⑪ phosphoribulose kinase.

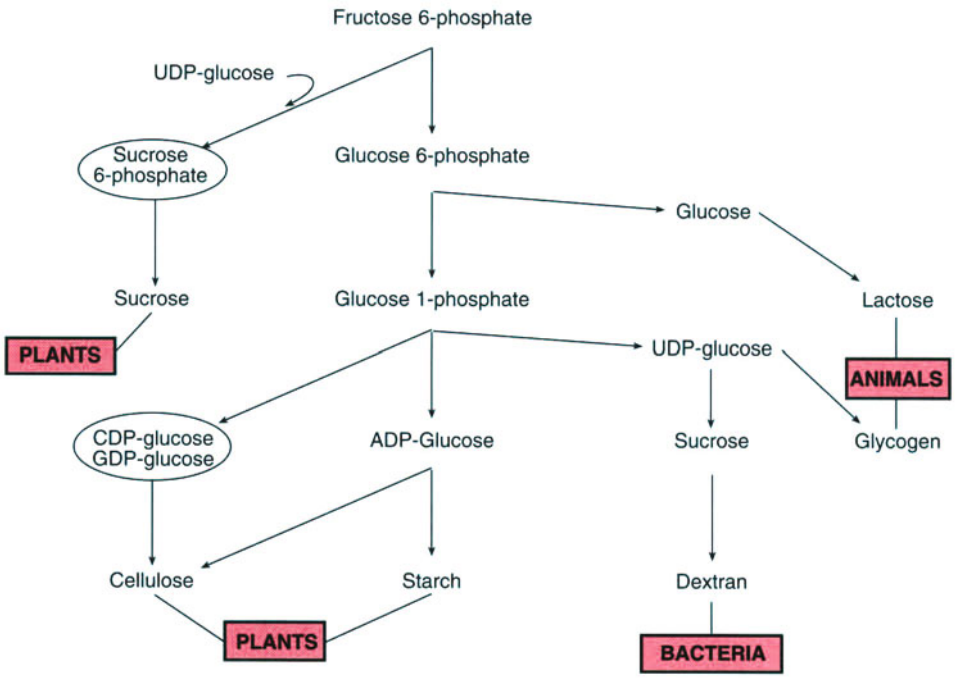


Figure 15.17. Some pathways for carbohydrate synthesis from fructose 6-phosphate.

glyceraldehyde 3-phosphate serve for synthesis of one molecule of hexose as part of the **carbohydrate formation stage** of the production phase. The remaining 10 molecules of glyceraldehyde 3-phosphate are converted to six molecules of ribulose 1,5-bisphosphate in the regeneration phase of the cycle.

Synthesis of fructose 6-phosphate is initiated with six molecules of glyceraldehyde 3-phosphate; three of these are used directly, and three are first converted to the isomeric form dihydroxyacetone phosphate. These precursors lead to formation of three molecules of fructose 1,6-bisphosphate, which undergo dephosphorylation to yield three molecules of fructose 6-phosphate. Two molecules of fructose 6-phosphate (equivalent to four molecules of glyceraldehyde 3-phosphate) become diverted to the regeneration phase. The third fructose 6-phosphate can lead to formation of glucose, starch, or other carbohydrates. These reactions include some of the steps of gluconeogenesis and constitute the carbohydrate formation stage of the production phase (Figure 15.17).

15.5.2B. Regeneration Phase. Ten of the 12 molecules of glyceraldehyde 3-phosphate serve to regenerate six molecules of ribulose 1,5-bisphosphate in the regeneration phase of the cycle. This phase is initiated with six molecules of glyceraldehyde 3-phosphate; four are used directly, and two are first converted to the isomeric form, dihydroxyacetone phosphate. The equivalent of four molecules of glyceraldehyde 3-phosphate become siphoned off from the pathway of hexose synthesis as two molecules of fructose 6-phosphate.

Reactions of the regeneration phase resemble those of the pentose phosphate pathway in that they involve a reshuffling of carbon skeletons (Figure 15.18) and action of *transketolase* (see Section 10.4). Transaldolase does

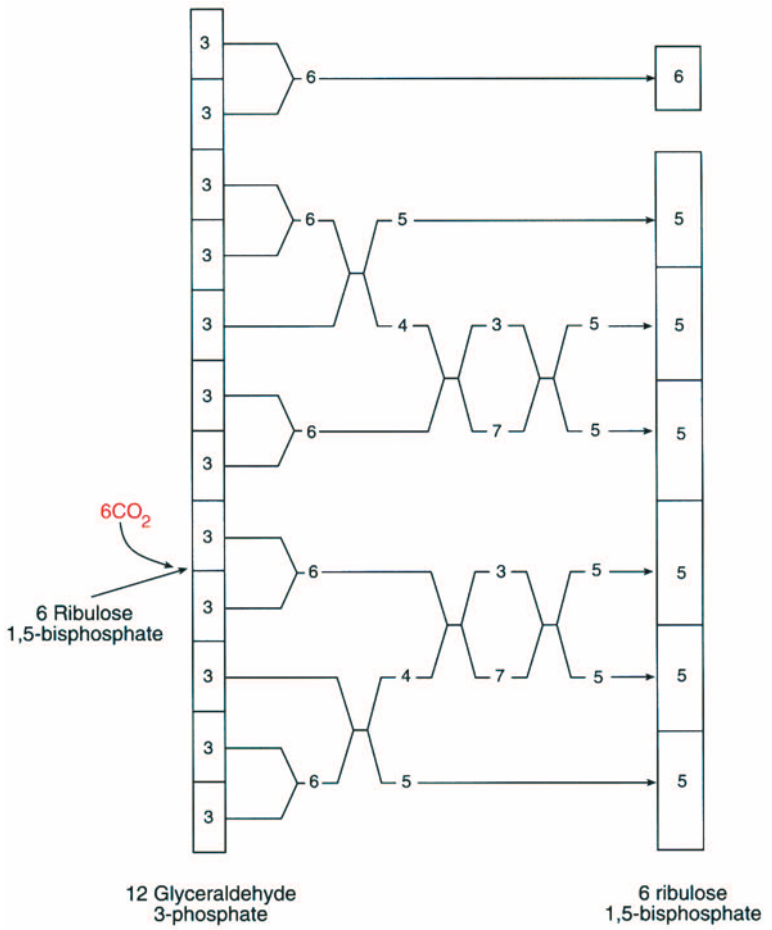


Figure 15.18. Schematic diagram of the rearrangements of carbon skeletons in the Calvin cycle. Twelve molecules of glyceraldehyde 3-phosphate regenerate six molecules of ribulose 1,5-bisphosphate and form one molecule of fructose 6-phosphate. Numbers indicate the number of carbon atoms per molecule.

not participate in the Calvin cycle; instead, two reactions are catalyzed by *aldolase*.

15.5.2C. Overall Reaction. To arrive at the overall reaction of the Calvin cycle, consider the stoichiometric relationships involved:

Reactants

CO_2: 6 CO_2 for rubisco reaction
ATP: 12 ATP for phosphorylation of 12 molecules of 3-phosphoglycerate
6 ATP for phosphorylation of 6 molecules of ribulose 5-phosphate
NADPH: 12 NADPH for reduction of 12 molecules of 1,3-bisphosphoglycerate
H_2O: 6 H_2O for rubisco reaction
5 H_2O for hydrolysis of 3 molecules of fructose 1,6-bisphosphate and 2 molecules of sedoheptulose 1,7-bisphosphate

Products

ADP: 18 ADP from phosphorylation by 18 ATP
P_i: 12 P_i from reduction of 12 molecules of 1,3-bisphosphoglycerate
5 P_i from hydrolysis of 3 molecules of fructose 1,6-bisphosphate and 2 molecules of sedoheptulose 1,7-bisphosphate
$NADP^+$: 12 $NADP^+$ from reduction by 12 NADPH
H^+: 6H^+ from phosphorylation of 6 molecules of ribulose 5-phosphate
Other: 1 molecule of fructose 6-phosphate

Combining these values yields the overall reaction:

$$6CO_2 + 18ATP^{4-} + 12NADPH + 11H_20 \downarrow$$
$$\text{Fructose 6-phosphate}^{2-} + 18ADP^{3-} + 17P_i^{2-} + 12NADP^+ + 6H^+$$

Thus, the Calvin cycle leads to net synthesis of one molecule of hexose as *fructose 6-phosphate*. Ribulose 1,5-bisphosphate does not appear in the overall equation. Although six molecules of ribulose 1,5-bisphosphate enter the cycle, *all six* are ultimately regenerated. Subsequent reactions can convert fructose-6-phosphate to glucose, starch, or other carbohydrates as you saw in Figure 15.17.

15.5.3. Control of the Cycle

The two phases of photosynthesis—the light and dark reactions—are interconnected. Because ATP and NADPH are products of the light reactions and reactants in the dark reactions, the light reactions affect the rate and extent of the dark reactions. As soon as ATP and NADPH start to form, the dark reactions begin. Dark reactions start while light reactions are still in progress and terminate some time after illumination has stopped.

There is a second linkage between the two photosynthetic phases in addition to that of product/reactant. Light is not only a required ingredient in the light reactions but also functions as an activator of several Calvin cycle enzymes. The activity of phosphoribulose kinase, for example, increases 100-fold on illumination. The mechanism of **light activation** involves *ferredoxin* (Section 15.4) and *thioredoxin* (Section 14.4) and is diagrammed in Figure 15.19.

When chloroplasts are subjected to strong illumination, reduced ferredoxin accumulates because of the activity of photosystem I. Under these conditions, ferredoxin donates some of its electrons to thioredoxin rather than to $NADP^+$. Reduction of thioredoxin is catalyzed by *ferredoxin-thioredoxin reductase* and converts a disulfide bond of thioredoxin to two sulfhydryl groups. Thioredoxin's sulfhydryl groups then undergo a *disulfide exchange* with a disulfide bond in the Calvin cycle enzyme being light-

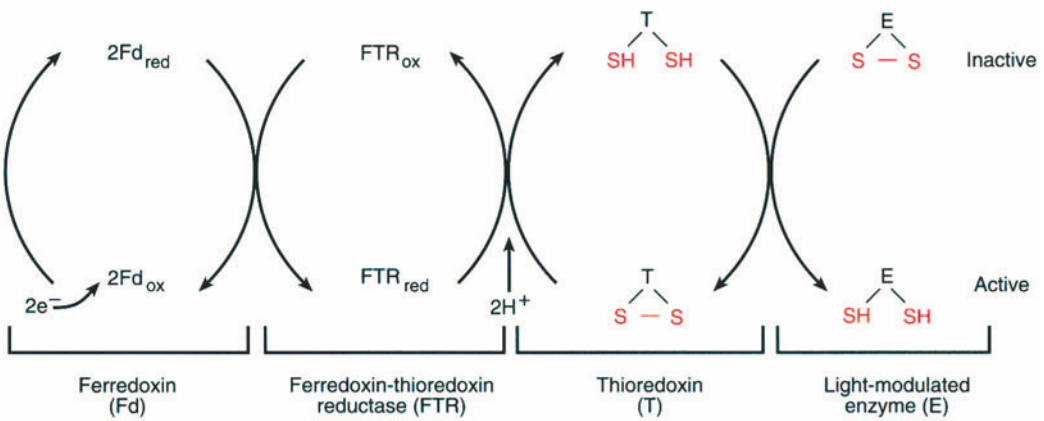


Figure 15.19. Light activation of some Calvin cycle enzymes. Electrons generated by photosystem I reduce ferredoxin. The mini electron transport system leads to reduction of a disulfide bond in a susceptible enzyme.

activated. Reduction of the enzyme's disulfide bond results in activation of the enzyme. Reduced thioredoxin also stimulates ATP synthase (CF_0-CF_1 ATPase) of chloroplasts. This ensures a high rate of ATP synthesis when chloroplast illumination is intense.

Light modulates the activity of several enzymes of carbohydrate catabolism via a similar mechanism except that there the effect is an opposite one, resulting in **light inactivation.** *Phosphofructokinase* and *glucose 6-phosphate dehydrogenase,* for example, are inhibited when their disulfide bonds are reduced by thioredoxin. Recall that these two catalysts constitute key enzymes of glycolysis and the pentose phosphate pathway, respectively.

Consequently, light stimulates carbohydrate synthesis in plants by activating Calvin cycle enzymes but prevents carbohydrate degradation by inhibiting enzymes of glycolysis and the pentose phosphate pathway. In the dark, these effects are reversed, and plants assume some metabolic characteristics of animals. Rather than carry out photosynthesis and produce and store carbohydrates, plants draw on their carbohydrate reserves for growth and respiration and degrade carbohydrates via glycolysis, the citric acid cycle, and the pentose phosphate pathway.

In addition to the regulatory effects of light, the dark reactions are controlled by means of the enzyme rubisco, located in the stroma. Rubisco is stimulated by increases in pH, $[Mg^{2+}]$, and [NADPH], all of which are produced by illumination of chloroplasts.

Light reactions cause pumping of protons from the stroma into the thylakoid space. Hence, the $[H^+]$ in the stroma decreases and the stroma becomes more alkaline, thereby activating rubisco. Movement of the protons is accompanied by movement of magnesium ions in the opposite direction to preserve electrical neutrality. As a result, the $[Mg^{2+}]$ increases in the stroma, which again serves to activate the enzyme. Lastly, as the light reactions progress, NADPH accumulates. NADPH serves as the terminal electron acceptor of the light reactions and as a positive allosteric effector of rubisco. Increasing concentrations of NADPH result in enzyme activation.

15.6. PHOTORESPIRATION

We have known since the 1960s that illuminated plants carry out reactions whose essence is the reverse of photosynthesis: they consume O_2 and evolve CO_2. This process, called **photorespiration,** becomes especially pronounced at low levels of CO_2 and high levels of O_2. Photorespiration differs from oxidative phosphorylation and results from rubisco's unusual ability to function as *either* a carboxylase or an oxygenase. Oxygen competes with carbon dioxide as a substrate for the enzyme. Because of this dual activity, we also refer to rubisco as *ribulose 1,5-bisphosphate carboxylase-oxygenase.*

When rubisco functions as an oxygenase, it binds O_2 at its active site instead of CO_2 (the two molecules are similar in size). Binding of oxygen initiates photorespiration. In this process, ribulose 1,5-bisphosphate undergoes conversion to 3-phosphoglycerate and *phosphoglycolate* (Figure 15.20). 3-Phosphoglycerate can enter the Calvin cycle, but phosphoglycolate is dephosphorylated to *glycolate,* followed by a series of reactions that convert it back to ribulose 1,5-bisphosphate. These reactions lead to evolution of CO_2, require reducing power in the form of NADH, and require energy input in the form of ATP. The entire pathway involves three types of organelles—*chloroplasts, peroxisomes,* and *mitochondria* (Figure 15.21). Photorespiration competes with and decreases the efficiency of photosynthesis owing to:

- loss of both CO_2 and ribulose 1,5-bisphosphate that could have fed into the Calvin cycle;
- evolution of CO_2 and consumption of O_2, which reverses the effects of CO_2 fixation;
- diversion of reducing power from CO_2 fixation, since NADH, expended in photorespiration, could have been converted to NADPH and used in the Calvin cycle; and
- loss of carbon, since not all of the carbon ultimately returns to the chloroplast

Photorespiration has no known metabolic function, so we consider it a wasteful process. In addition to decreasing photosynthetic efficiency, its operation requires both ATP and NADH. Consumption of NADH effectively depletes the supply of NADPH because the two reducing compounds are interconvertible by action of the enzyme *transhydrogenase:*

$$NADPH + NAD^+ \rightleftarrows NADP^+ + NADH$$

Thus, photorespiration uses up both ATP and NADPH produced by the light reactions, thereby undermining the reactions' usefulness. Additionally, as opposed to ordinary cellular respiration, photorespiration is not coupled to an oxidative phosphorylation; reduction of O_2 is *not* accompanied by synthesis of ATP. Therefore, photorespiration results in a threefold wastefulness: O_2 is needlessly reduced, and both NADPH and ATP are needlessly expended.

Photorespiration leads to a tremendous loss of photosynthetic products. In some plants, it results in as much as a 50% reduction of these products. If photorespiration could be avoided, the food supply to the growing world

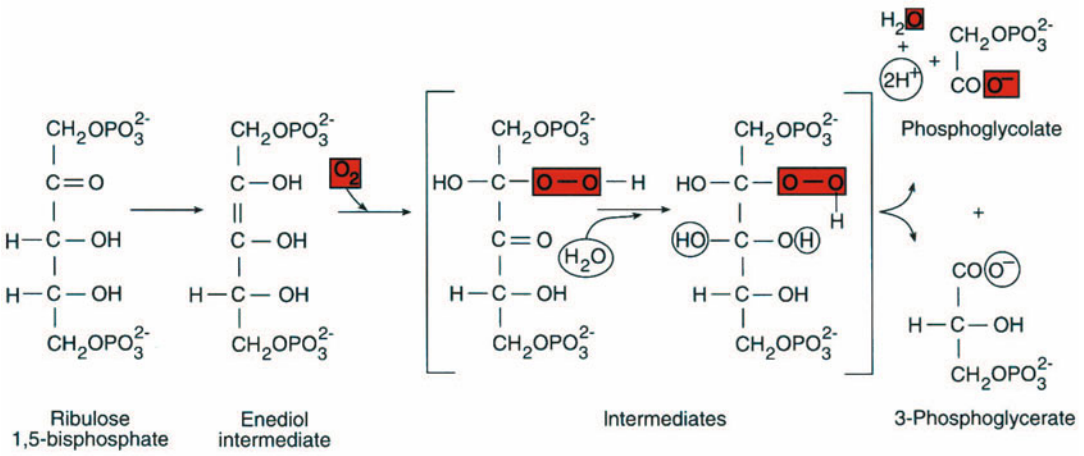


Figure 15.20. The reaction catalyzed by ribulose 1,5-bisphosphate carboxylase when the enzyme functions as an oxygenase in photorespiration.

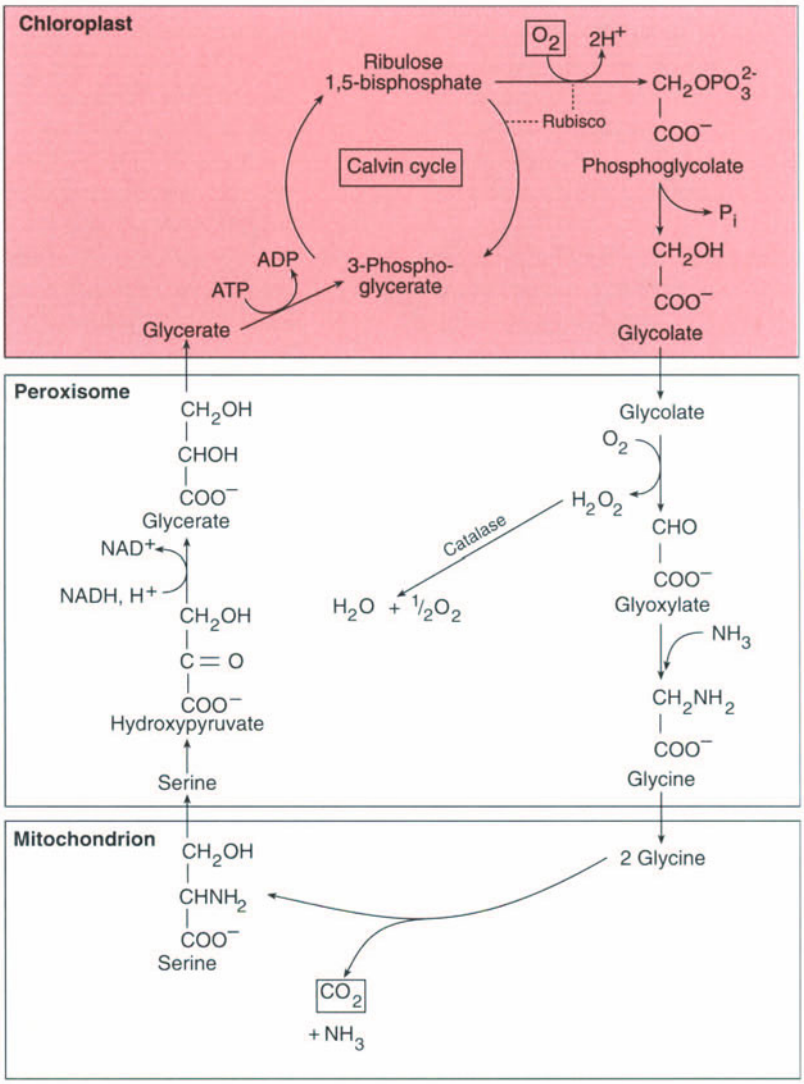


Figure 15.21. The pathway of photorespiration. Peroxisomes are organelles that produce and use hydrogen peroxide.

population would greatly increase. Because of this, efforts are under way to use genetic engineering to either modify the specificity of rubisco or to introduce the C_4 cycle (see below) into plants that do not have it.

15.7. THE C_4 CYCLE

Fixation of CO_2 via the Calvin cycle occurs in all photosynthetic plants, but the mechanism of fixation varies. Plants that grow in *temperate* climatic zones convert CO_2 directly to stable 3-phosphoglycerate, a *three-carbon compound.* These plants are called **C_3 plants.**

Other plants, however, fix CO_2 by a different mechanism. In these plants, a number of reactions *precede* operation of the Calvin cycle. Because these reactions include an initial fixation of CO_2 to stable oxaloacetate, a *four-carbon compound,* these plants are called **C_4 plants.** Following its initial fixation, CO_2 is released and fixed a second time via the Calvin cycle. We term the complete mechanism of CO_2 fixation in these plants the **C_4 cycle** or the **Hatch–Slack pathway** in honor of its discoverers, Marshall Hatch and Roger Slack. These investigators worked on the elucidation of the C_4 cycle between 1966 and 1970.

C_4 plants include not only some crop species, such as sugarcane, sorghum, and corn, but also desert plants, crabgrass, and Bermuda grass. These plants all have two properties in common: they thrive in hot and sunny environments, including tropical and desert zones, and they have evolved a unique leaf anatomy coupled with specific metabolic properties.

Plant leaves contain two types of cells, **mesophyll cells** and **bundle-sheath cells** (Figure 15.22). In C_3 plants, mesophyll cells contain chloroplasts and carry out the reactions of the Calvin cycle; bundle-sheath cells are devoid of chloroplasts.

In C_4 plants, both types of cells contain chloroplasts, but all of the rubisco is concentrated in bundle-sheath cells so that the reactions of the Calvin cycle are carried out only by these cells. Mesophyll cells have become specialized and contain all of the enzymes required for operation of the C_4 cycle.

Mesophyll cells, in C_4 plants, surround bundle-sheath cells in a close concentric packing, separating them effectively from the epidermis and the stoma. This minimizes the leaf's water loss by transpiration and allows bundle-sheath cells to retain their water and function in photosynthesis even in hot, dry environments.

In C_4 plants, the initial acceptor of CO_2 is *phosphoenolpyruvate (PEP).* CO_2 fixation, catalyzed by *PEP carboxylase,* occurs in mesophyll cells and yields oxaloacetate (Figure 15.23). Oxaloacetate subsequently undergoes conversion to malate, which diffuses into bundle-sheath cells, where it is decarboxylated and converted to pyruvate. The CO_2 produced is fixed via the Calvin cycle, and pyruvate returns to the mesophyll cells, where it is converted to phosphoenolpyruvate.

Carbon dioxide is thus *fixed twice;* once in meso-

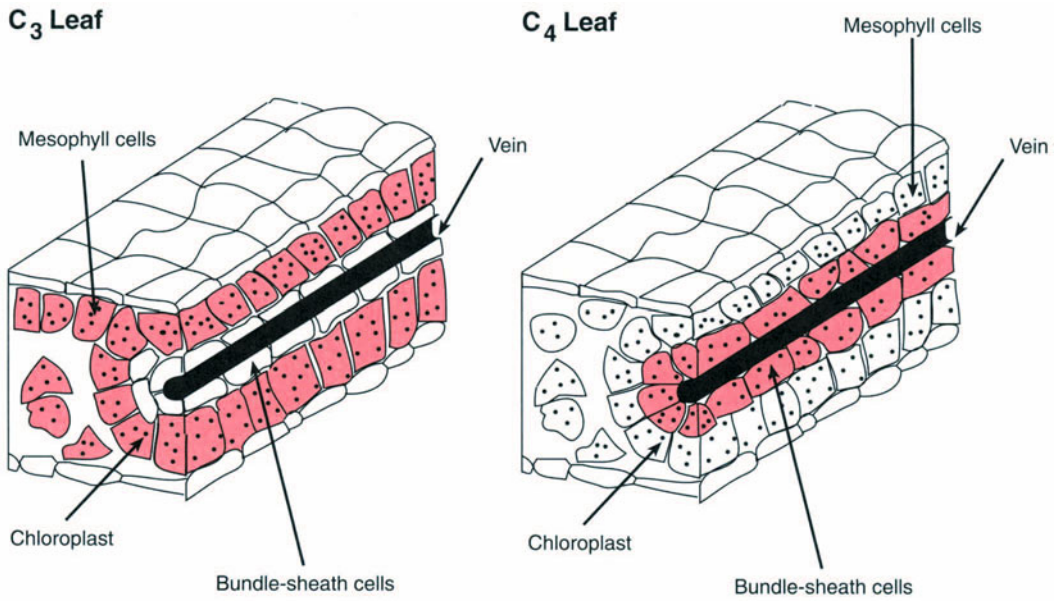


Figure 15.22. Leaf structure of C_3 and C_4 plants. A vein comprises a bundle of tubes transporting water, inorganic ions, and organic compounds through the leaf. Stoma are minute openings in the leaf through which gas exchange takes place. Colored cells have the capacity to carry out normal CO_2 fixation via the Calvin cycle.

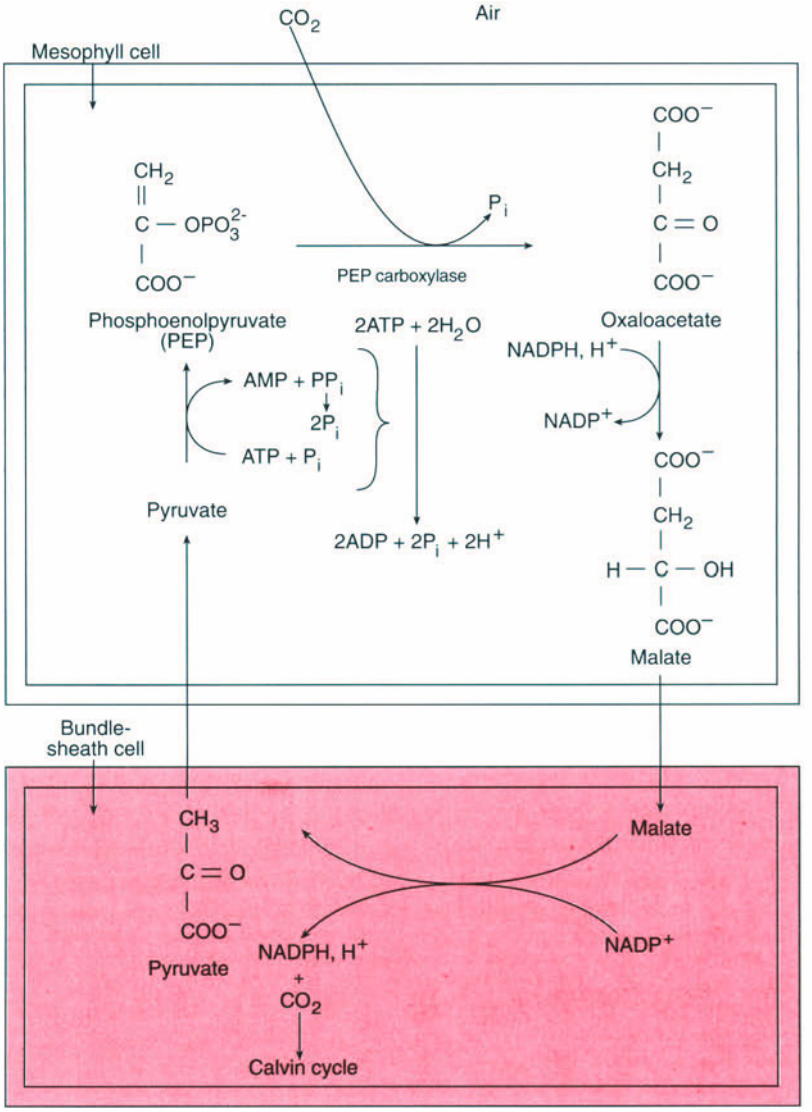


Figure 15.23. The C_4 cycle (Hatch–Slack pathway).

phyll cells to oxaloacetate (C_4 cycle) and once in bundle sheath cells to 3-phosphoglycerate (Calvin cycle). Double fixation increases the efficiency of the Calvin cycle. The initial CO_2-fixing enzyme (PEP carboxylase) has much greater affinity for CO_2 than rubisco. PEP carboxylase serves as an effective scavenger of CO_2, so that much more CO_2 is ultimately "fixed" by rubisco than would otherwise be the case. High [CO_2] also minimizes the possibility that rubisco will use O_2, rather than CO_2, as a substrate and initiate photorespiration.

Double CO_2 fixation in C_4 plants costs more in terms of ATP expended than the single fixation that occurs in C_3 plants. Conversion of pyruvate to PEP uses *pyruvate phosphate dikinase,* an unusual enzyme that catalyzes phosphorylation of *both* pyruvate and P_i by ATP:

$$\text{Pyruvate}^- + \text{ATP}^{4-} + \text{P}_i^{2-} \rightarrow \text{PEP}^{3-} + \text{PP}_i^{3-} + \text{AMP}^{2-} + \text{H}^+$$

As in other reactions producing PP_i, pyrophosphatase catalyzes the hydrolysis of PP_i to $2P_i$. Consequently, the energy expenditure becomes *equivalent* to hydrolysis of two energy-rich bonds (ATP → AMP + PP_i; PP_i → $2P_i$). The energy requirement is large because synthesis of PEP by a simple transfer of a phosphate group from ATP is thermodynamically unfavorable. The overall reaction of the combined C_4 and Calvin cycles is

$$6\text{CO}_2 + 30\text{ATP}^{4-} + 12\text{NADPH} + 23\text{H}_2\text{O} \downarrow \text{Fructose 6-phosphate}^{2-} + 30\text{ADP}^{3-} + 29\text{P}_i^{2-} + 12\text{NADP}^+ + 18\text{H}^+$$

This equation differs from that of the Calvin cycle by increased numbers of ATP, ADP, P_i, H_2O, and H^+. The increases result from the added energy expenditure in the pyruvate phosphate dikinase reaction of the C_4 cycle. That reaction requires an *effective* expenditure of 2 ATP per molecule of CO_2 fixed:

$$2ATP^{4-} + 2H_2O \rightarrow 2ADP^{3-} + 2P_i^{2-} + 2H^+$$

Multiplying this equation by 6 (for fixation of 6 CO_2) and adding it to the Calvin cycle equation yields the overall stoichiometry for the combined operation of the two cycles.

Photosynthetic efficiency of C_4 plants is much greater than that of C_3 plants. C_4 plants synthesize carbohydrate faster and grow faster than C_3 plants. Three reasons explain the increased efficiency:

1. $[CO_2]$ for the rubisco reaction is increased by the prior fixing of CO_2 via PEP carboxylase, which has a much greater affinity for CO_2 than rubisco.
2. Water loss is minimized due to the leaf's unique anatomical structure. Hence the plants can function even in dry environments and can fix CO_2 even when the water supply is low.
3. Wasteful photorespiration is greatly minimized, or eliminated entirely, by the increased $[CO_2]$, which prevents rubisco from using O_2 as a substrate. Prevention of photorespiration exacts a price from C_4 plants; they must expend extra energy to form PEP. However, the high energy price is well worth the photosynthetic efficiency gained.

SUMMARY

Photosynthesis—synthesis of carbohydrate from CO_2 and H_2O in the presence of light and pigments—constitutes essentially the reverse of respiration and is carried out by plants, algae, and bacteria. Most photosynthetic organisms use water as electron donor and produce oxygen. Some organisms use other compounds as electron donors and do not evolve oxygen. Plants and cyanobacteria carry out photosynthesis in chloroplasts by means of two photosystems. Photosystem I (PSI) and photosystem II (PSII) respond, respectively, to light below 700 nm and below 680 nm. The two photosystems function in series, an arrangement called the Z-scheme.

Absorption of a photon excites a molecule of chlorophyll in a photosystem by boosting an electron to a higher energy level. The excitation energy passes via resonance energy transfer through antenna chlorophylls to a reaction center containing specialized chlorophylls. Excited specialized chlorophylls transfer an electron to a suitable acceptor, thereby reducing the acceptor. In PSI, the electron passes through a chain of electron carriers to $NADP^+$, which is reduced to NADPH. In PSII, the electron passes through another set of electron carriers to PSI, where it fills the "electron hole" produced when PSI lost an electron. The electron hole produced in PSII is filled with an electron derived from the splitting of water (photolysis). ATP synthesis is coupled to operation of the chain of electron carriers linking PSI and PSII.

Light reactions comprise four individual reactions, require absorption of eight photons for reduction of two molecules of $NADP^+$, and lead to synthesis of four molecules of ATP by photophosphorylation. Photosynthetic bacteria use only one photosystem, and ATP synthesis is coupled to a cyclic electron flow.

Dark reactions (CO_2 fixation) convert carbon dioxide to carbohydrate via the Calvin cycle. The cycle uses ATP and NADPH produced in the light reactions and results in net synthesis of hexose. Ribulose 1,5-bisphosphate carboxylase (rubisco) initiates the reactions of the Calvin cycle by catalyzing the combination of CO_2 with ribulose 1,5-bisphosphate to yield two molecules of 3-phosphoglycerate. Rubisco has both carboxylase and oxygenase activity. When the the enzyme functions as an oxygenase, it leads to photorespiration, a wasteful process of unknown metabolic function that decreases photosynthetic efficiency.

Certain plants (C_4 plants) use a double fixation of CO_2, once into oxaloacetate and once into 3-phosphoglycerate. The C_4 pathway is more efficient than the single fixation pathway in ordinary plants (C_3 plants).

SELECTED READINGS

Bassham, J. A., and Calvin, M., *The Path of Carbon in Photosynthesis,* Prentice-Hall, Englewood Cliffs, New Jersey (1957).

Deisenhofer, J., and Michel, H., The photosynthetic reaction center from the purple bacterium *Rhodopseudomonas viridis, Science* 245:1463–1473 (1989).

Golbeck, J. H., Structure and function of photosystem I, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 43:293–324 (1992).

Govindjee, and Coleman, W. J., How plants make oxygen, *Sci. Am.* 262(2):50–58 (1990).

Greenbaum, E., Lee, J. W., Tevault, C. V., Blankinship, S. L., and Mets, L. J., CO_2 fixation and photoevolution of H_2 and O_2 in a mutant of *Chlamydomonas* lacking photosystem I, *Nature (London)* 376:438–441 (1995).

Hall, D. D., and Rao, K. K., *Photosynthesis,* Cambridge University Press, Cambridge (1994).

Hartman, F. C., and Harpel, M. R., Structure, function, regulation, and assembly of D-ribulose 1,5-bisphosphate carboxylase/oxygenase, *Annu. Rev. Biochem.* 63:197–234 (1994).

Hatch, M. D., C_4 photosynthesis: A unique blend of modified biochemistry, anatomy, and ultrastructure, *Biochim. Biophys. Acta* 895:81–106 (1987).

Krauss, N., *et al.,* Three-dimensional structure of system I of photosynthesis at 6Å resolution, *Nature (London)* 361:326–331 (1993).

Kühlbrandt, W., Wang, D. N., and Fujiyoshi, Y., Atomic model of plant light-harvesting complex by electron crystallography, *Nature (London)* 367:614–621 (1994).

Ogren, W. L., Photorespiration: Pathways, regulation, and modification, *Annu. Rev. Plant. Physiol.* 35:415–442 (1984).

Vermaas, W., Molecular biological approaches to analyze photosystem II structure and function, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 44:457–481 (1993).

Yachandra, V. K., De Rose, V. J., Latimer, M. J., Mukerji, I., Sauer, K., and Klein, M. P., Where plants make oxygen: A structural model for the photosynthetic oxygen-evolving manganese cluster, *Science* 260:675–679 (1993).

REVIEW QUESTIONS

A. Define each of the following terms:

Phytol	Photon
Internal conversion	Fluorescence
Resonance energy transfer	Quantum yield
Electron transfer	Red drop
Thylakoid disk	Hill reaction
Reaction center	Photophosphorylation
Emerson enhancement effect	Rubisco

B. Differentiate between the two terms in each of the following pairs:

Photosynthesis/ photorespiration	Photosystem I/ photosystem II
Light reactions/dark reactions	Ground state/excited state
Chlorophyll/chloroplast	Stroma/grana
Antenna chlorophylls/ specialized chlorophylls	Light activation/ light inactivation
Cyclic photophosphorylation/noncyclic photophosphorylation	Carboxylation stage/ phosphorylation stage
Calvin cycle/ Hatch-Slack Pathway	Mesophyll cells/ bundle-sheath cells

C. (1) Discuss each of the four individual light reactions that occur in plant photosynthesis and explain the operation of the Z-scheme.
(2) What is the mechanism of photorespiration? Why does photorespiration decrease the efficiency of photosynthesis and why does it represent a wasteful process?
(3) What are the two phases of the Calvin cycle and what are their subdivisions? Outline the path of carbon for each phase. What are the energetics of the Calvin cycle?
(4) How are the dark reactions affected by light?
(5) Compare the photosynthetic process in C_3 and C_4 plants.

PROBLEMS

15.1. A student calculates the free energy changes for various parts of the Z-scheme from the potentials shown in Figure 15.9. Do these calculated values correctly describe the free energy changes in the chloroplast under physiological conditions? Why or why not?

15.2. Write out the equations for the reactions in plant photosynthesis that are responsible for establishing a pH gradient across the thylakoid membrane.

15.3.* A student sets up a number of experiments along the line of Priestley's classic experiment. The student uses (a) a C_3 plant that does not carry out photorespiration; (b) a C_3 plant with extensive photorespiration; and (c) a C_4 plant that does not carry out photorespiration. Everything else being equal, predict the relative lengths of time that the mouse will survive in the apparatus under these conditions.

15.4. Based on the estimated amount of carbon fixation given in the text, calculate the annual amount of glucose that can be produced via photosynthesis.

15.5. Destruction of the ozone layer by fluorocarbons used in spray cans increases the extent of ultraviolet irradiation of the Earth. Would you expect this to affect the rate of photosynthesis? Why or why not?

15.6. What is the quantum yield of ATP synthesis in cyclic photophosphorylation if the absorption of three photons leads to the synthesis of two molecules of ATP?

15.7. Based on the data in Figure 15.6, what are the likely colors of carotenoids, phycocyanins, and phycoerythrins?

15.8.* Assume that photolysis of water in photosynthesis leads to ozone (O_3) formation:

$$3H_2O \rightarrow O_3 + 6H^+ + 6e^-$$

On this basis, write the equations for the overall reactions of photosystem I, photosystem II, and their combined operation. How many photons must strike each photosystem and how many electrons must flow through the Z-scheme?

15.9.* What is the quantum yield for ATP synthesis of the system in the previous problem if we know that a total of two molecules of ATP are synthesized for every two electrons passing from P680* to P700 (Figure 15.9)?

15.10.* What is the efficiency of energy conservation for the electron transport system from P680* to P700 based on Problems 15.8 and 15.9?

15.11. Theoretically, what is the maximum number of molecules of ATP that could be obtained from the flow of one electron in bacterial photosynthesis (Figure 15.13) for the chain of carriers (a) from P870* to the Cyt *b*/Fe-S complex and (b) from the Cyt *b*/Fe-S complex to P870?

15.12. Would you be able to demonstrate the Emerson enhancement effect by using photosynthesizing cells of the bacterium *Rhodopseudomonas viridis* and supplementing light at 870 nm with that at 700 nm? Why or why not?

15.13.* Calculate $\Delta G^{\circ\prime}$ for the excitation of P870 to P870*. What is the minimum number of photons of 870-nm light that will bring about the excitation of one molecule of P870?

15.14.* When photosynthesis is carried out in the presence of $^{14}CO_2$, what fraction of the carbons in glyceraldehyde 3-phosphate will be labeled after 12 molecules of glyceraldehyde 3-phosphate have been formed? Depending upon which two molecules of glyceraldehyde 3-phosphate react to form fructose 6-phosphate, the latter will have different degrees of labeling. What are the various fractions of ^{14}C in fructose 6-phosphate that could be produced from this pool of glyceraldehyde 3-phosphate?

15.16. Leaves of deciduous trees are green in the summer and often have brilliant red and yellow colors in the fall. Given that these leaves are known to lose their chlorophyll in the fall, what must cause the nongreen colors?

15.17.* Cyclic photophosphorylation, as illustrated in Figure 15.11, leads to synthesis of two molecules of ATP for every three photons (700 nm) absorbed. On this basis, what is the efficiency of solar energy conservation?

15.17.* Calculate $\Delta E^{\circ\prime}$ and $\Delta G^{\circ\prime}$ for the reduction of $NADP^+$ by ferredoxin based on the data in Table 12.1.

15.18. You place a potted C_3 plant and a potted C_4 plant in a sealed glass container containing an adequate amount of moisture. You then illuminate the container. As time goes on, the C_4 plant thrives, but the C_3 plant shrivels up and dies. How do you explain these results?

15.19. The herbicide dichlorophenyldimethylurea (DCMU) blocks electron transfer between Q_A and QH_2 in photosystem II (see Figure 15.9). Do you expect DCMU to inhibit cyclic photophosphorylation? Why or why not?

15.20. DCMU-inhibited chloroplasts do not normally evolve oxygen (O_2), but when ferricyanide is added to the chloroplasts, they do. How do you explain this result?

15.21.* A researcher determines the kinetics of CO_2 incorporation by rubisco once in a nitrogen atmosphere and once in an oxygen atmosphere. He plots the data as $1/v$ versus $1/[CO_2]$ and obtains two straight

lines (v = velocity of the reaction). Draw the lines expected, and explain how the Michaelis constants (K_m), calculated from the lines, are related.

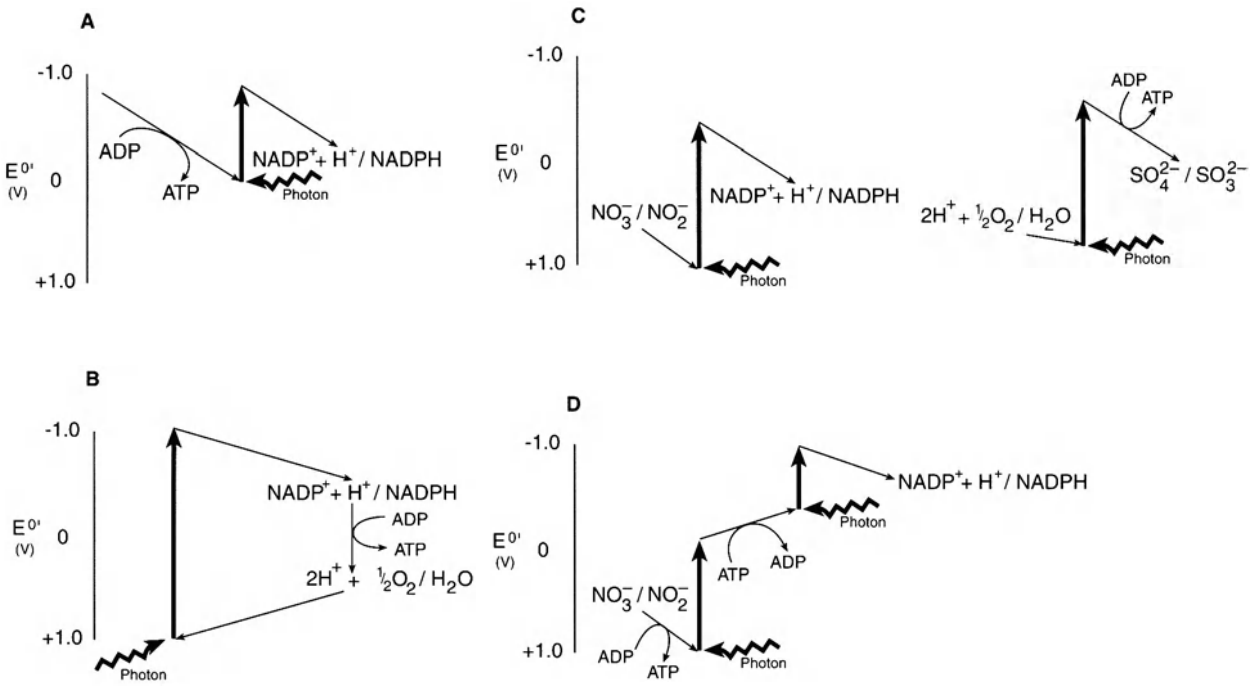
15.22. The steady-state concentrations of ATP, ADP, and P_i in actively photosynthesizing spinach chloroplasts are 120μM, 6.00μM, and 700μM, respectively. Under these conditions, what is the minimum amount of energy required for the synthesis of one mole of ATP? Hint: Use ([ADP][P_i])/[ATP] for the [products]/[reactants] term. $\Delta G^{\circ\prime}$ for ATP hydrolysis is -30.5 kJ mol^{-1}.

15.23.* The chloroplasts of the previous problem are known to lead to the synthesis of four molecules of ATP from ADP for every eight photons absorbed. Assuming that all the photons are of the same wavelength, calculate the maximum wavelength of light that can result in photosynthesis in this system, based on an efficiency of 9% for the entire process comprising the light reactions.

15.24. What is the minimal potential drop per $2e^-$ required for synthesis of ATP in actively photosynthesizing spinach chloroplasts, based on your calculations in Problem 15.22?

15.25. What is the maximum increase in reduction potential per e^- that can be brought about by a photon of blue light (500 nm)?

15.26.* Consider the following schematic representations of photosynthetic systems in *hypothetical* organisms. Each half-reaction ties into the scheme via its oxidant or reductant as shown. For each representation, indicate whether the scheme, as outlined, is theoretically *possible* or *impossible*. If the latter applies, explain why this is so.



Transfer of Genetic Information

IV

The last part of this book deals with the biosynthesis of nucleic acids and proteins. These processes differ radically from the biosynthetic pathways discussed in Part III because of their obligatory link to DNA—the genetic material of the cell. Synthesis of DNA, RNA, and protein involves special mechanisms not required for the synthesis of low-molecular-weight biomolecules or high-molecular-weight polysaccharides. Since newly made DNA, RNA, and protein structures relate to the original DNA, we refer to their synthesis as transfer of genetic information.

Introduction to Molecular Biology

16

The term *molecular biology* came into use around the middle of the 20th century. In its broadest meaning, **molecular biology** comprises the study of a variety of biological processes at the molecular level. The physicochemical properties of cellular components and their functions in nerve-impulse conduction, vision, membrane transport, and molecular genetics constitute some of the topics of molecular biology.

In a more restricted sense, and the one most commonly used today, molecular biology refers to **molecular genetics,** which deals with DNA, RNA, and protein synthesis. The essence of molecular genetics is the storage, transmission, and expression of genetic information. Major concerns include *mapping genes* on chromosomes; mechanisms of *gene expression;* duplication of chromosomes (*DNA replication*); restoration of damaged DNA (*DNA repair*); the two stages of *protein biosynthesis, DNA transcription* and *RNA translation;* and *recombinant DNA technology.*

The phenomenal growth and incredible advances in molecular genetics over the last 50 years have a number of causes. Clear demonstration that DNA is the genetic material of the cell served to focus attention on the properties and metabolism of this biopolymer. These studies received a tremendous boost when Watson and Crick proposed their double-helical model for the structure of DNA in 1953. The proposal constituted a major breakthrough for both biochemistry and genetics, because it provided a physicochemical concept of heredity. The DNA double helix immediately suggested possible mechanisms for chromosome duplication and repair of mutant DNA. The *gene,* the unit of heredity, became identified as a section of DNA, a double-stranded polynucleotide segment having a unique sequence of base pairs. Some genes code for polypeptide chains, others code for RNA molecules, and yet others have regulatory functions.

Confirmation of the double-helical structure of DNA led to a burst of research activity, frequently accompanied by development of new experimental techniques. Both our understanding of the transfer of genetic information and our ability to manipulate the genetic material grew at a rapid rate as a result of this and related research. Some of the many ad-

vances include sequencing DNAs and individual genes, identifying genes that have specific functions, synthesizing genes in the laboratory, incorporating genes into cells of other organisms, and producing specific proteins by such genetically altered cells. Determining the complete nucleotide sequence of the human genome represents a vast project that is currently under active investigation.

16.1. SOME BASIC CONCEPTS

16.1.1. Replication, Transcription, and Translation

To store genetic information in newly made cells, chromosomes must be duplicated as cells divide. We call the duplication of genetic material **replication**—the process whereby a DNA molecule leads to synthesis of identical DNA molecules (DNA → DNA). The newly synthesized DNA has the same *base composition* and the same *base sequence* as the original DNA. Replication also refers to the process whereby identical copies of viral RNA are produced in a host cell (RNA → RNA). Viral replication requires the action of either *RNA replicase* or *reverse transcriptase.* Unless otherwise specified, in this chapter replication refers to synthesis of DNA from DNA.

Transmission of the genetic information contained within DNA is accomplished by a copying mechanism called **transcription.** In this process, a DNA strand leads to synthesis of an RNA strand (DNA → RNA). The RNA strand is *complementary* and *antiparallel* to the DNA strand copied (Figure 16.1). Cytosine (C), guanine (G), and thymine (T) in DNA lead to guanine, cytosine, and adenine (A), respectively, in RNA. However, since RNA contains uracil (U) but not thymine, any adenine in DNA leads to uracil in RNA. Synthesis of DNA by copying of the information contained in RNA (RNA → DNA) is called **reverse transcription.**

Key enzymes in DNA replication and transcription are DNA polymerases and RNA polymerases. We name a polymerase according to the product whose formation it catalyzes. **DNA polymerases** catalyze the synthesis of DNA, and **RNA polymerases** catalyze the synthesis of RNA. However, because both DNA and RNA can be produced from *either* DNA or RNA, we include the type of nucleic acid transcribed or replicated in the enzyme name (Table 16.1).

Translation refers to the process whereby an RNA strand directs the synthesis of a polypeptide chain (RNA → protein). Translation constitutes the second stage of *protein biosynthesis* (transcription being the first stage). Because RNA is made from DNA, proteins represent products of specific DNA genes. The protein complement of a cell is an expression of its genetic makeup and is directly controlled by it. This relationship is unique to proteins and nucleic acids. Synthesis of other biomolecules,

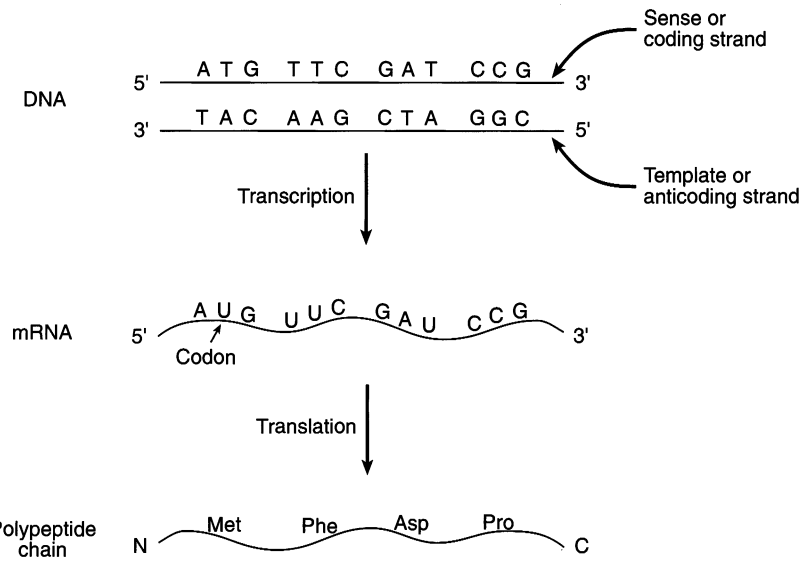


Figure 16.1. The principles of transcription and translation. One DNA strand, the template or anticoding strand, is transcribed into a single-stranded messenger RNA molecule (mRNA), which is then translated into a polypeptide chain. A sequence of three nucleotides (a codon) codes for one amino acid.

Table 16.1. Polymerase Nomenclature

Precise name	Common name
Replication	
DNA-dependent DNA polymerase	DNA polymerase
RNA-dependent RNA polymerase	RNA replicase
Transcription	
DNA-dependent RNA polymerase	RNA polymerase
Reverse transcription	
RNA-dependent DNA polymerase	Reverse transcriptase

like carbohydrates and lipids, does not proceed under direct control of cellular DNA. Rather, these compounds form from a variety of precursors in reactions controlled by specific enzymes.

16.1.2. The Central Dogma of Molecular Biology

The flow of genetic information via replication, transcription, and translation can be described by the **central dogma** of molecular biology. The dogma was enunciated by Francis Crick in 1958. In essence, it states that transfer of information from nucleic acid to nucleic acid or from nucleic acid to protein is possible, but transfer from protein to protein or from protein to nucleic acid is impossible. According to the central dogma, proteins can only be *recipients,* not donors, of genetic information. Proteins do, however, regulate the use of genetic information. In light of our current knowledge, we formulate the central dogma as follows:

DNA can be replicated to form DNA and transcribed to form RNA. RNA can be replicated to form RNA and also can yield DNA by reverse transcription. Lastly, RNA can be translated to yield protein.

16.1.3. Primers, Templates, and Polymerization

In replication, transcription, and translation, one polymer molecule leads to formation of another. In these processes, the original polymer sometimes functions as a primer and sometimes as a template (Figure 16.2).

The term **primer,** or **initiator,** refers to a compound that functions in polymerization by *initiating* the process. It does so by providing a "handle" for the newly synthesized polymer, a point of attachment from which the polymer grows. Because the first building block becomes covalently attached to the primer, the primer is always *covalently linked to the polymerization product.*

The term **template** refers to a compound that functions in polymerization by *controlling* the process, determining which building blocks will be polymerized, and in which order. This control occurs without any building block becoming covalently linked to the template so that the template is only *bound noncovalently to the polymerization product.*

Replication, transcription, and translation can each be divided into four stages:

1. **Initiation:** The first stage of the reaction; synthesis of the beginning of the polymer.
2. **Elongation:** The middle, longest stage during which most of the polymer becomes synthesized.
3. **Termination:** The final stage of the reaction; synthesis of the end of the polymer.
4. **Processing:** Any alteration in the polymer *after* it has been synthesized by means of primers and/or templates. Processing may involve shortening or lengthening of the polymer, or the polymer may be modified chemically in other ways. Depending on the type of polymerization, there occur three types of processing—*postreplicative processing, posttranscriptional processing,* and *posttranslational processing.*

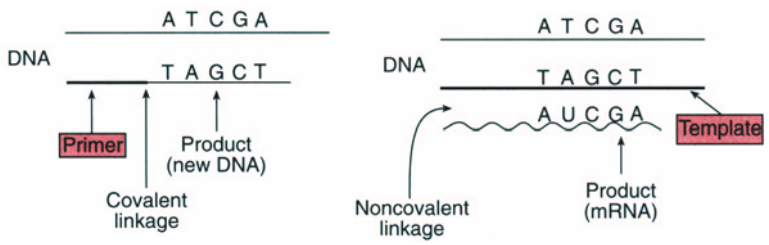


Figure 16.2. Functions of a primer and a template.

16.2. STRUCTURE AND FUNCTION OF RNA

16.2.1. Overview of Protein Biosynthesis

We divide **protein biosynthesis** into two main stages—transcription and translation. During transcription, the genetic information stored in DNA is copied through synthesis of RNA. The RNA strand formed is called **messenger RNA** (**mRNA**); it is complementary and antiparallel to the transcribed DNA strand, except that it contains U instead of T. We call the transcribed DNA strand **template strand** or **anticoding strand.** The nontranscribed DNA strand is termed **coding strand** or **sense strand** because it runs in the same 5′ → 3′ direction as the mRNA; it "reads" the same as mRNA.

The second stage of protein biosynthesis, translation, produces a polypeptide chain via mRNA. The sequence of nucleotides in mRNA determines the sequence of amino acids in the polypeptide chain. Three consecutive nucleotides (a *triplet*) in mRNA specify an amino acid and are called a **codon.** That section of mRNA that codes for the polypeptide chain consists of a sequence of codons in which each codon specifies one amino acid.

Messenger RNA molecules associate with ribosomes—subcellular ribonucleoprotein particles where protein synthesis occurs. As amino acids polymerize, the growing polypeptide chain remains bound to a ribosome (Figure 16.3). Several ribosomes attach to a single strand of mRNA at any given time, forming a complex termed a **polyribosome** or **polysome.** Each ribosome carries the same polypeptide chain but in different stages of completion.

Amino acids are transported to ribosomes by **transfer RNAs (tRNA).** Each tRNA links a specific amino acid covalently, forming an **aminoacyl-tRNA.** In this complex, the amino acid is esterified via its carboxyl group to a ribose hydroxyl group of the 3′-terminal adenine nucleotide of tRNA. Every tRNA contains an **anticodon** in a part of the molecule far from the amino acid attachment site. The anticodon is a trinucleotide, complementary and antiparallel to the codon for the amino acid carried by the tRNA. Aminoacyl-tRNA and mRNA associate at the ribosome by means of hydrogen bonding between the complementary bases of the codon and the anticodon.

16.2.2. Messenger RNA (mRNA)

In prokaryotes, DNA is located in a nuclear region; in eukaryotes, it occurs in the cell nucleus. In both types of or-

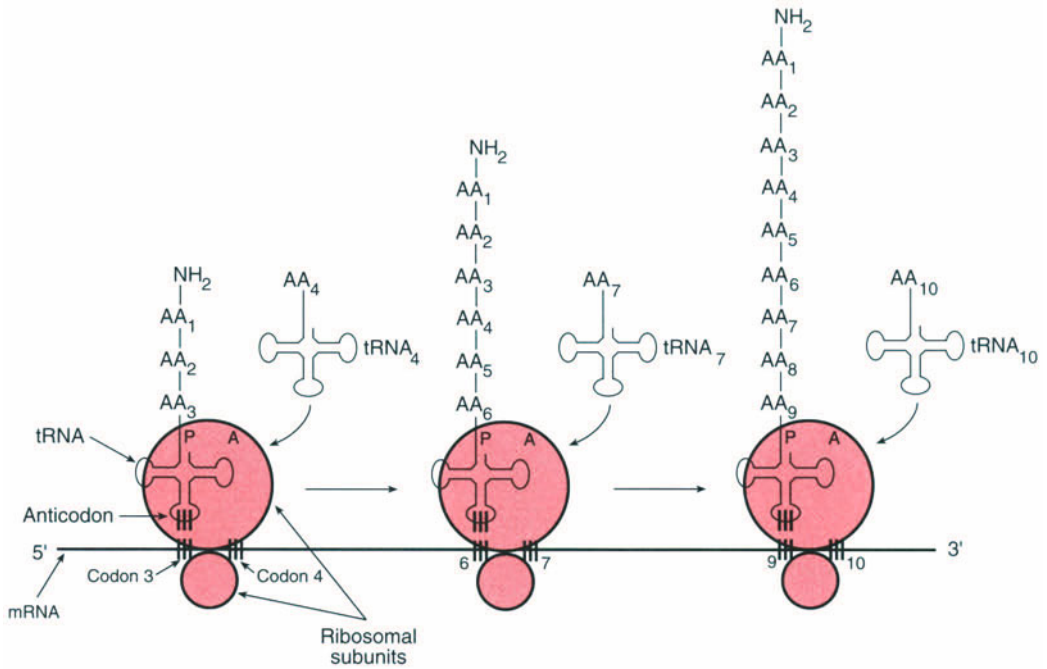


Figure 16.3. Protein synthesis on ribosomes. mRNA is translated in the 5′ → 3′ direction, and the polypeptide chain grows from the N- to the C-terminus. Several ribosomes attach to one mRNA strand, carrying identical polypeptide chains, but in different stages of completion.

ganisms, a large part of protein synthesis takes place on ribosomes located in the cytoplasm. How can nuclear DNA control and determine protein synthesis on cytoplasmic ribosomes?

In 1961, François Jacob and Jacques Monod proposed that a "messenger" conveys the genetic information from DNA to the ribosomes. After these investigators postulated what properties such a messenger should have, a search was undertaken. A single-stranded RNA molecule, *messenger RNA (mRNA),* was ultimately identified as being the link between nuclear DNA and reactions on the ribosome.

As you have seen, messenger RNA is transcribed from one of the strands of double-stranded DNA and is complementary and antiparallel to that strand. Following its synthesis in the nucleus, eukaryotic mRNA passes through pores in the nuclear membrane and associates with ribosomes in the cytoplasm. Prokaryotic mRNA associates with ribosomes as it is produced. Most prokaryotic mRNA has a short half-life (2–3 min) so that its rate of synthesis and degradation constitute important controls of protein synthesis. Half-lives of eukaryotic mRNA vary widely. Some mRNAs are degraded within 30 min of appearing in the cytoplasm, but most have half-lives of hours or days.

Messenger RNA is heterogeneous in size, reflecting the different sizes of genes that code for polypeptide chains of varying lengths. As mentioned, the section of mRNA that codes for a polypeptide chain consists of a sequence of codons in which each codon specifies one amino acid. Thus, a polypeptide containing 100 amino acids requires a messenger RNA having at least 300 nucleotides. (mRNA may also contain regulatory sequences in addition to codons.) The gene of this polypeptide is a double-stranded DNA segment, with each strand consisting of 300 nucleotides.

Messenger RNA represents a small fraction, typically about 5%, of cellular RNA. In *Escherichia coli,* mRNA constitutes about 3% of total cell RNA. To what extent single-stranded mRNA molecules possess intrachain folding (secondary structure) has not been definitively established.

16.2.3. Ribosomes and Ribosomal RNA (rRNA)

Ribosomes represent the sites of protein synthesis. They occur both as free cytoplasmic particles and as particles bound to the *endoplasmic reticulum (ER)* (Figure 16.4). Free ribosomes function in the synthesis of a variety of proteins. Ribosomes bound to the endoplasmic reticulum function in the synthesis of membrane proteins and proteins destined to be secreted from the cell. We term an endoplasmic reticulum studded with many ribosomes a *rough endoplasmic reticulum (RER),* and one that has few or no ribosomes attached a *smooth endoplasmic reticulum (SER).* The number of ribosomes per cell is large; an *E. coli* cell contains about 20,000.

Biochemists recognize five classes of ribosomes. Two large classes consist of cytoplasmic ribosomes from prokaryotes and eukaryotes. The remaining three classes comprise ribosomes of plant mitochondria, animal mitochondria, and chloroplasts. Ribosomes of all classes consist of two unequal and roughly globular subunits (see Figure 19.1), linked together noncovalently by means of magnesium ions and other bonds. Each ribosome has two binding sites (the *A site* and the *P site*) for transfer RNA (see Figure 19.2).

We conventionally describe sizes of ribosomes, ribosomal subunits, and ribosomal RNA by their **sedimentation coefficients (*s*),** expressed in *Svedberg units* (S; see Appendix C). The sedimentation coefficient depends both on molecular weight and on particle shape. Hence sedimentation coefficients are not additive like molecular weights. For example, a bacterial ribosome (70S) forms from two subunits

$$50S + 30S \rightleftharpoons 70S$$

Figure 16.4. Electron micrograph on which *E. coli* ribosomes were first identified. The rod is a tobacco mosaic virus particle serving as a standard for comparison. [Reprinted, with permission, from H. K. Schachman, A. B. Pardee, and R. Y. Stanier, *Arch. Biochem. Biophys.* 38:245–260 (1952).]

We can readily shift the equilibrium between ribosomes and their subunits by changing the magnesium-ion concentration:

$$\text{50S} + \text{30S} \underset{\text{decreasing } [Mg^{2+}]}{\overset{\text{increasing } [Mg^{2+}]}{\rightleftharpoons}} \text{70S}$$

A similar relationship exists for eukaryotic ribosomes, where a 40S subunit and a 60S subunit combine to form an 80S ribosome. The composition of both prokaryotic and eukaryotic ribosomal subunits is about two-thirds RNA and one-third protein as you can see in Table 16.2. In prokaryotes, the small subunit contains 21 different polypeptide chains and one 16S RNA molecule; the large subunit contains 34 polypeptide chains (31 different ones; one chain is present in four copies) and two RNA molecules (5S and 23S). Thus, a prokaryotic ribosome represents a macromolecular aggregate consisting of 55 protein and three RNA molecules. Eukaryotic ribosomes are even larger, consisting of 82 protein and four RNA molecules. Mitochondrial and chloroplast ribosomes differ somewhat from the cytoplasmic ones described in Table 16.2.

Ribosomal proteins (**r-prot**) are difficult to separate; they are basic proteins, like most proteins associated with nucleic acids, and are insoluble in ordinary buffers. We designate proteins of the small and large subunits by the letters *S* and *L*, respectively, and by a number (for example, S4 and L6). Current research attempts to determine the precise location and function of each ribosomal protein. Some ribosomal proteins have enzymatic activity, some bind mRNA or tRNA, and some serve as essential structural components or in some other capacity. The entire ribosomal particle undergoes self-assembly *in vivo*, a process that can be duplicated in the laboratory.

Ribosomal RNA (**rRNA**) consists of a number of different single-stranded molecules. Ribosomal RNA contains both irregularly folded chain sections and double-stranded helical segments, formed by folding of the polynucleotide strand back on itself. Double-stranded RNA segments resemble double-helical A-DNA in helix dimensions (see Section 7.3) and are held together by *in-*

Table 16.2. Physical Properties of Cytoplasmic Ribosomes[a]

	Prokaryotes		Eukaryotes		
Organism	*E. coli*		Rat liver		
Ribosome					
Type	70S		80S		
Mol. wt.	2.5×10^6		4.2×10^6		
Small subunit					
Type	30S		40S		
Mol. wt.	0.9×10^6		1.4×10^6		
Percent RNA	60		50		
Percent protein	40		50		
Proteins					
Number	21		33		
Mol. wt.	8,000–26,000		11,000–42,000		
RNAs					
Type	16S		18S		
Number of nucleotides	1,542		1,874		
Mol. wt.	~500,000		~700,000		
Large subunit					
Type	50S		60S		
Mol. wt.	1.6×10^6		2.8×10^6		
Percent RNA	70		65		
Percent protein	30		35		
Proteins					
Number	34		49		
Mol. wt.	5,000–25,000		12,000–42,000		
RNAs					
Types	5S	23S	5S	5.8S	28S
Number of nucleotides	120	2,904	120	160	4,718
Mol. wt.	~40,000	~1.0×10^6	~40,000	~50,000	~1.7×10^6

[a]Adapted, with permission, from B. Lewin, *Genes,* 3rd ed., Wiley, New York (1987).

trachain hydrogen bonds between complementary bases: A⁼⁼U (2 H-bonds), and G≡≡C (3 H-bonds). Double-stranded RNA shows a hyperchromic effect (see Section 7.6). Ribosomal RNA is the most abundant form of cellular RNA, representing about 80–85% (83% in *E. coli*) of total cell RNA.

16.2.4. Transfer RNA (tRNA)

In 1958, Francis Crick proposed that specific adapters bind amino acids covalently and transport them to ribosomes for polymerization. Researchers subsequently showed that these adapters were RNA molecules. Originally known as adapter RNAs, acceptor RNAs, or soluble RNAs, we now refer to them as *transfer RNAs (tRNAs),* stressing their function in protein synthesis. Transfer RNAs constitute about 10–15% (14% in *E. coli*) of total cell RNA.

Each amino acid has at least one unique tRNA molecule to which it becomes attached. Frequently, several different **isoacceptor tRNAs** can bind and transfer the same amino acid; consequently, the total number of different tRNA species for an organism is typically 50–60 (see Problem 16.1). We designate a tRNA specific for a given amino acid (AA) as $tRNA^{AA}$. Thus, $tRNA^{Gly}$ and $tRNA^{Phe}$ designate tRNAs specific for carrying glycine and phenylalanine. When an amino acid becomes linked to tRNA, the resulting aminoacyl-tRNA designation includes the amino acid abbreviation as a prefix. For instance, Gly-$tRNA^{Gly}$ and Phe-$tRNA^{Phe}$ stand for glycyl-$tRNA^{Gly}$ and phenylalanyl-$tRNA^{Phe}$, that is, glycine and phenylalanine linked covalently to their specific tRNAs.

Transfer RNAs are small single-stranded molecules, composed of 70–90 nucleotides (MW ≈ 20,000–30,000; sedimentation coefficient ≈ 4S). Because of their small size, tRNA molecules were the first nucleic acid molecules to be sequenced. Like ribosomal RNAs, tRNAs have secondary structure wherein the polynucleotide strand folds back on itself forming double-stranded helical segments. In these segments, complementary bases are linked via intrachain H-bonds (A⁼⁼U, G≡≡C). Double-stranded segments and single-stranded sections form a cloverleaf-type structure (Figure 16.5). In this model, hydrogen-bonded segments or *stems* form ladderlike parts, and non-hydrogen-bonded segments, the *loops,* loop out. A stem and a loop combine to form an **arm.** Because of its extensive hydrogen bonding, tRNA shows a significant hy-

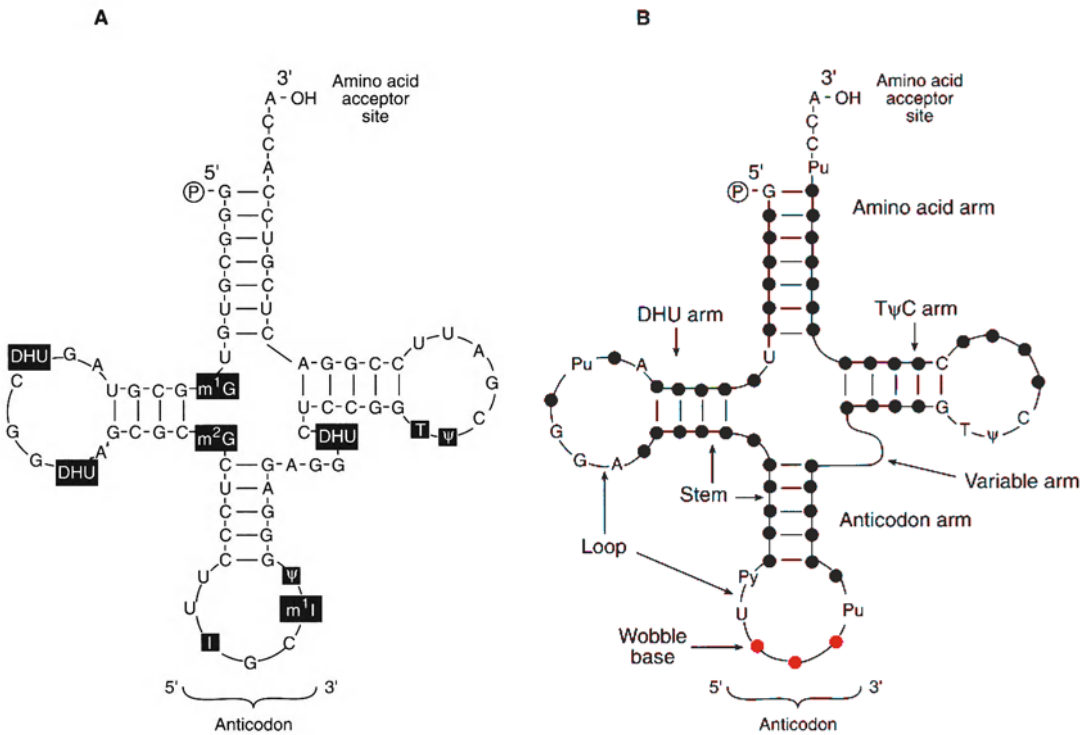


Figure 16.5. The cloverleaf model of transfer RNA. (A) Nucleotide sequence of yeast alanine tRNA. Modified nucleosides are shaded: ψ, pseudouridine; I, inosine; T, ribothymidine; DHU, 5,6-dihydrouridine; m^1I, 1-methylinosine; m^1G, 1-methylguanosine; m^2G, N^2-dimethylguanosine. (B) Generalized structure with common features indicated. Pu, purine; Py, pyrimidine. Typically, 7 nucleotides occur in the TψC loop, 7 in the anticodon loop, and 8–12 in the DHU loop.

perchromic effect and is relatively resistant to digestion by ribonuclease.

All tRNA molecules have four arms. Some have a fifth, **variable arm.** About 60–70% of the bases are linked in helical segments. The cloverleaf model correctly represents the base pairings and secondary structure of tRNA, but not the molecule's three-dimensional structure. The tertiary structure of tRNA, as determined by X-ray diffraction, consists of a *folded cloverleaf* called the **L-shaped structure** (Figure 16.6). It is held together by **tertiary hydrogen bonds,** non-Watson–Crick-type H-bonds formed between various donor and acceptor groups in the bases and in ribose.

Although tRNA is relatively small, it represents a versatile molecule having a number of unusual properties:

- *tRNA participates in two different reactions:*
  1. tRNA functions in **amino acid activation** or **charging.** This process involves a covalent linking of an amino acid to tRNA and is catalyzed by **aminoacyl-tRNA synthetases.** At least 20 different synthetases occur, one for each amino acid. The product of amino acid activation is an aminoacyl-tRNA (a "charged" tRNA).
  2. tRNA functions in peptide bond formation. tRNA molecules transport amino acids to ribosomes in the form of aminoacyl-tRNAs and position them for polymerization into protein.
- *tRNA has at least four different binding sites:*
  1. An **amino acid acceptor site** (CCA-3′) at which the amino acid becomes covalently linked to tRNA (see Figure 16.5).
  2. An **aminoacyl-tRNA synthetase recognition site** that binds at the enzyme's active site (the term *recognition* in molecular biology always refers to a specific binding interaction).
  3. A **ribosome interaction site** by which the aminoacyl-tRNA becomes bound to the ribosome.
  4. An **anticodon site** by which the tRNA H-bonds to a codon in mRNA.
- *tRNA has a large proportion of modified bases.* Transfer RNA contains a large proportion of nonstandard bases and unusual nucleosides. Modified bases include *methylated derivatives* of standard bases and *dihydrouracil* (Figure 7.2). Unusual nucleosides include *inosine* (the ribonucleoside of hypoxanthine; Figure 14.17), *ribothymidine* (ribose linked to thymine), and *pseudouridine* [uracil linked to ribose via C(5) rather than via N(1); Figure 16.7].

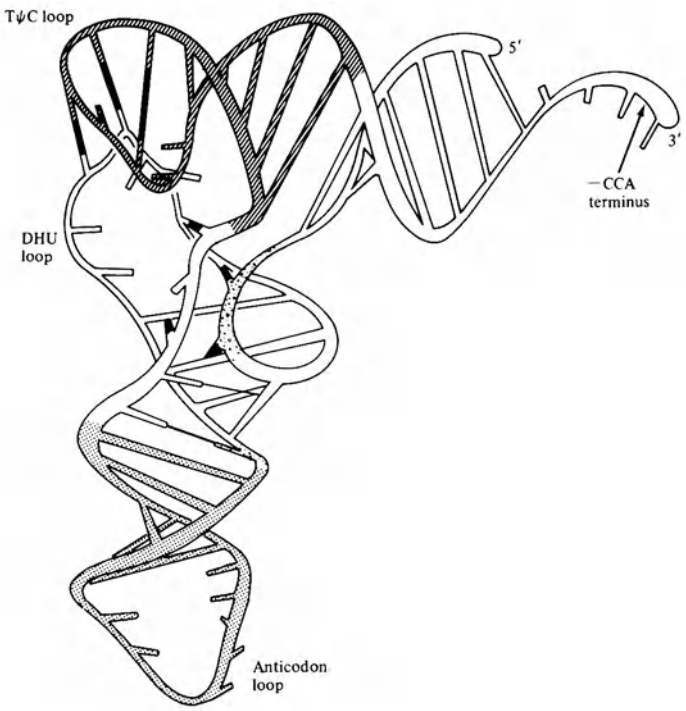


Figure 16.6. The three-dimensional, L-shaped structure of yeast phenylalanine tRNA. [Reprinted, with permission, from G. J. Quigley and A. Rich, *Science* 194:796–806 (1976). Copyright © 1976 American Association for the Advancement of Science.]

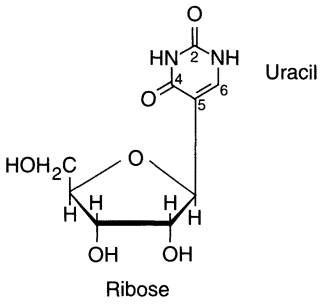


Figure 16.7. Pseudouridine (ψ).

16.2.5. Small RNA

Messenger, ribosomal, and transfer RNA represent the major forms of RNA in both prokaryotes and eukaryotes, and all three RNAs function in protein biosynthesis. In addition, all cells contain **small RNA** molecules. Many small RNAs are associated with proteins and function in posttranscriptional processing of RNA (Section 18.3). In eukaryotic cells, small RNAs are abundant in the nucleus. These **small nuclear RNAs (snRNAs)** vary in size from about 60 to 300 nucleotides and have a highly conserved base sequence from organism to organism. They associate with proteins, forming **small nuclear ribonucleoproteins** (**snRNPs,** pronounced "snurps") that participate in the reactions whereby split genes are converted to functional genes by excision of introns and splicing of exons. Small RNAs occur in the cytosol as well. There, too, they associate with specifc proteins, forming **small cytosolic ribonucleoproteins** (**scRNPs,** pronounced "scurps").

16.3. THE GENETIC CODE

16.3.1. The Alphabet of Genetics

Once it became apparent that nucleic acids serve as repositories of genetic information, researchers wanted to determine what constitutes this information and how it is transferred from nucleic acids to protein. They soon realized that the information content of a nucleic acid must reside within the sequence of its bases since the ribose-phosphate backbone stays the same throughout a polynucleotide strand.

Nucleic acids and proteins have important structural similarities. They are both long and linear (unbranched) polymers; they are constructed from a relatively small number of building blocks; and the building blocks are joined via a single type of bond. Based on these similarities, the notion arose that nucleic acids function as a *template* for the synthesis of proteins. In particular, the idea was that the sequence of bases in a polynucleotide strand specifies or *codes* for a sequence of amino acids in a polypeptide chain. We refer to this relationship as the **genetic code.**

Since only four different bases occur in a nucleic acid and these must code for 20 different amino acids in a protein, we can define the genetic code more specifically as *the manner in which sequences in a polynucleotide strand, constructed of four different bases, code for sequences in a polypeptide chain constructed of 20 different amino acids.*

An analogy to the problem posed by the genetic code would be the task of constructing a dictionary of 20 words using an alphabet of only four letters. George Gamov, a physicist, was the first to suggest, in 1954, that in such a dictionary each word must consist of at least three letters. Similarly, each code word or *codon* for an amino acid must consist of a sequence of at least three bases (a trinucleotide or *triplet*).

Gamov reasoned that if words (codons) are composed of only one letter (base), then four letters (bases) can form only four words (codons), an insufficient number to code for 20 amino acids. If two letters (bases) are used, the number of possible words (codons) increases to $4 \times 4 = 16$, still not large enough to code for all of the amino acids. However, if three letters (bases) are allowed, $4 \times 4 \times 4 = 64$ words (codons) can be formed, more than enough to code for 20 amino acids. In forming these codons, all possible permutations apply, since the sequence of letters (bases) is important. Much as "act" differs from "tca," so 5′-GUA-3′ differs from 5′-AUG-3′.

Following Gamov's proposal, scientists proceeded to decipher the genetic code. This work not only identified which of the 64 possible codons code for which amino acid but also confirmed the triplet nature of the code. Experiments showed that precisely three nucleotides were required to code for an amino acid.

We use several conventions when dealing with the genetic code. First, we commonly employ terms that apply to language (letters, words, reading, etc.) to describe aspects of the genetic code. Second, the term *codon* refers to the nucleotide sequence *in mRNA.* The nucleotide sequence of an anticodon and that of a DNA strand from which the mRNA codon is transcribed are complementary and antiparallel to the mRNA codon:

mRNA 5′ - - - - - - - AUG - - - - - - -3′
codon

tRNA 3′ - - - - - - - UAC - - - - - - -5′
anticodon

Third, a "sequence of bases" actually means a "sequence of nucleotides," since nucleotides, and not bases, are linked together in a nucleic acid strand. Lastly, we always write codons with the 5′-end on the left and the 3′-end on the right.

Because a triplet code provides 64 codons for 20 amino acids, it seems there may be "extra" codons. In fact, investigators have shown that all of the extra codons are used because most amino acids are coded by more than one codon, a phenomenon called **degeneracy.** We call different codons that code for the same amino acid **synonym codons,** and codons that do not code for amino acids **nonsense codons.** There exist only three nonsense codons—UAG, UAA, and UGA—and we refer to them as **termination codons** or **stop codons** because they lead to termination of polypeptide chain synthesis. These codons are also known as *amber, ochre,* and *opal codons,* respectively, based on a laboratory joke. The first term was coined by reference to a scientist by the name of Bernstein (German for "amber"), who helped discover the mutation that changes a **sense codon** (one that codes for an amino acid) to UAG. The other two terms, ochre and opal, came into use because they designate minerals that frequently resemble amber in their color.

Gamov pointed out that, theoretically, there could be two types of genetic codes for "reading" the message; codes could be either overlapping or nonoverlapping (Figure 16.8). In an **overlapping code,** one or two bases from one codon also make up part of a subsequent codon; a given base could be used more than once. In a **nonoverlapping code,** each base is used only once. Both types of codes could be constructed with or without interruptions between codons (*punctuation marks* or *commas*).

The type of code that has evolved in living organisms employs the simplest and most straightforward way of "reading" the genetic information: a triplet, nonoverlapping code without any punctuation. We know of two exceptions to this general way of reading genetic messages. In eukaryotes, many *genes are discontinuous.* Such genes, also called *split genes,* consist of two types of nucleotide sequences, exons and introns. An **exon** (*expresses* genetic information) is a nucleotide sequence that codes for a section of the amino acid sequence of the corresponding protein. An **intron** (*intervening* sequence) is a nucleotide sequence that does not code for amino acids; its function is to separate two exons. We will discuss the mechanism for forming a functional gene from exons and introns in Section 18.3.

The second exception occurs in a number of *bacteriophages,* viruses that infect bacteria. Some bacteriophages (or phages) contain *overlapping genes.* Scientists first demonstrated this phenomenon for ϕX174, a phage that infects *E. coli.* Phage ϕX174 contains a single-stranded circular DNA genome that consists of 5386 nucleotides and codes for the coat proteins of the virus. Based on its length, the DNA could code for a maximum of 1795 amino acids. However, the total length of all the phage coat proteins corresponds to a number of amino acids that exceeds 1795 significantly. Researchers have explained this paradox by showing that some genes overlap so that a given DNA section is used twice. The starting point for one gene is located within another gene, and the two genes are "read" by means of different "reading frames" (Figure 16.9).

Note that in both of these variations on the simple arrangement of genes in the chromosome, the fundamental nature of the code remains unchanged. Within each exon and intron, and within each of the overlapping genes, the code still consists of a triplet, nonoverlapping, and unpunctuated code.

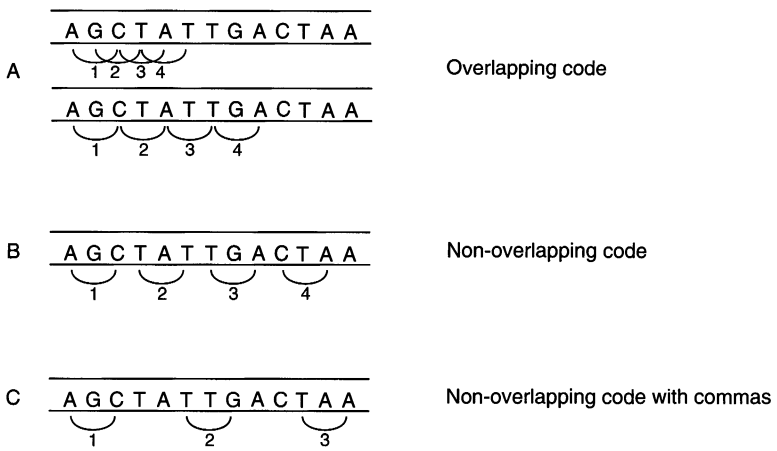


Figure 16.8. Examples of overlapping and nonoverlapping codes. Other codes can be constructed by varying the size and frequency of punctuation marks (commas).

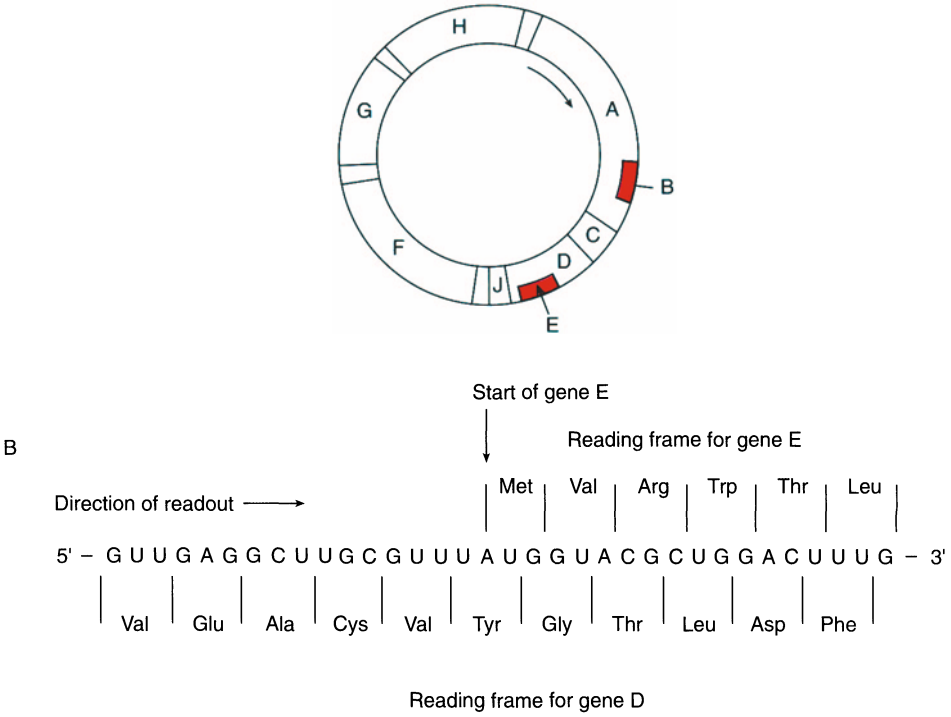


Figure 16.9. Overlapping genes. (A) The DNA of phage ϕX174 contains nine genes (A–H, J), and two genes occur within other genes: B lies within A, and E lies within D. (B) Because an overlapping gene does not start with the 5′-end of a codon, two different reading frames result.

16.3.2. Deciphering the Genetic Code

16.3.2A. Base Composition of Codons.

Once major aspects of the genetic code had been outlined, investigators attempted to decipher the code and to determine precisely which codons code for which amino acid. Breaking the code involved two stages. Initial work resulted in establishing the base composition of codons. Subsequent work determined the actual sequence of the bases within each codon.

The first successful experiments were performed by Marshall Nirenberg (Nobel Prize, 1968) and Héinrich Matthaei in 1961 using *cell-free amino acid-incorporating systems* prepared from *E. coli.* These systems contained ribosomes, enzymes, tRNA, and a synthetically produced RNA that functioned as mRNA. When fortified with energy sources (ATP and GTP) and amino acids, the systems were capable of synthesizing small amounts of protein that could be isolated and analyzed.

The researchers produced a **synthetic mRNA** using ribonucleoside diphosphates (ADP, CDP, GDP, UDP) and **polynucleotide phosphorylase.** Polynucleotide phosphorylase is a bacterial enzyme that catalyzes the polymerization of ribonucleotides in a *random manner,* using one or more types of ribonucleoside diphosphates, to yield a product whose *base composition is like that of the reaction mixture.*

In the simplest such experiment, Nirenberg and Matthaei used the enzyme to produce a **homopolynucleotide,** consisting of a single type of nucleotide. Specifically, they prepared *polyuridylic acid [poly(U)]* from UDP. Since poly(U) contains only one type of codon (UUU), it can code for only one amino acid and must lead to synthesis of a **homopolypeptide.** The investigators used poly(U) and a mixture of 20 amino acids, one of which was ^{14}C-labeled, and determined the radioactivity of the synthetic polypeptide. Of 20 such experiments, using a different ^{14}C-labeled amino acid each time, only one yielded labeled protein. That occurred when the labeled amino acid was phenylalanine. Nirenberg and Matthaei concluded that the codon for phenyalanine must be UUU. Proceeding in a similar fashion, using poly(A), poly(C), and poly(G), they could assign the codons AAA, CCC, and GGG to lysine, proline, and glycine, respectively.

More complex experiments employed synthetic mRNA that was a random copolymer of two or more nucleotides. Occurrence of nucleotides in such a copolymer is a function of probability. For example, in a copolymer

containing 80 mol % U and 20 mol % C [poly(UC) (80:20)], the probability that a given codon consists of three uracil nucleotides (UUU) is given by

$$(0.80)(0.80)(0.80) = 0.51 \text{ or } 51\%$$

There is a 51% chance that any given triplet in this copolymer consists of UUU. The probability that a codon consists of two uracil nucleotides and one cytosine nucleotide is

$$(0.80)(0.80)(0.20) = 0.13 \text{ or } 13\%$$

This probability refers only to *composition* of the codon; the actual sequence of bases in the codon could be UUC, UCU, or CUU. When a copolymer is used as synthetic mRNA, several different amino acids are incorporated into protein. The relative incorporation of these amino acids reflects the relative frequencies with which the corresponding codons occur in the mRNA.

As an illustration, assume that an unknown amino acid (AA_x) is coded for by a codon whose composition is two U and one C [UUC]. Since UUU codes for phenylalanine, it follows that

$$\frac{\begin{matrix}\text{incorporation}\\ \text{of amino acid coded}\\ \text{for by [UUU]}\end{matrix}}{\begin{matrix}\text{incorporation}\\ \text{of amino acid coded}\\ \text{for by [UUC]}\end{matrix}} = \frac{\begin{matrix}\text{incorporation}\\ \text{of Phe}\end{matrix}}{\begin{matrix}\text{incorporation}\\ \text{of } AA_x\end{matrix}}$$

$$= \frac{\text{frequency of [UUU]}}{\text{frequency of [UUC]}} = \frac{51}{13} \approx 4$$

Thus, incorporation of the unknown amino acid should be about one-fourth that of phenylalanine. If you inspect Table 16.3, you will see that serine fulfills this requirement. Hence you can conclude that the unknown amino acid (AA_x) is serine and that it has a probable codon whose composition is [UUC]. You cannot decide from this experiment whether the actual serine codon is UUC, UCU, or CUU. Proceeding in this fashion and using various copolymers as synthetic mRNAs, investigators were able to assign base compositions to codons.

16.3.2B. Base Sequence of Codons. Determining the base sequence of codons was accomplished by means of two types of experiments. The discovery by Marshall Nirenberg and Philip Leder in 1964 that aminoacyl-tRNAs bind to ribosomes in the presence of corresponding trinucleotide codons led to development of a **ribosome binding assay.** This assay measured ribosome binding of labeled aminoacyl-tRNA as a function of added synthetic *trinucleotides of defined sequences.* Any trinucleotide that resulted in binding was identified as a codon

Table 16.3. Amino Acid Incorporation Using Synthetic Messenger RNA[a]

	Incorporation (μmol amino acid/mg ribosomal protein)	
Labeled amino acid	No mRNA	Poly(UC) (80:20)
Phenylalanine	0.05	14.86
Serine	0.05	3.54
Proline	0.02	0.78
Tyrosine[b]	0	0

[a]Adapted, with permission, from J. F. Speyer, Lengyel, P., Basilio, C., Wahba, A. J., Gardner, R. S., and Ochoa, S., *Cold Spring Harbor Symp. Quant. Biol.* 28:559–567 (1963).
[b]Tyrosine serves as a control since it has no codons containing only U and C.

of the amino acid attached to the tRNA. For example, since UUU led to ribosome binding of Phe-$tRNA^{Phe}$ and AAA led to binding of Lys-$tRNA^{Lys}$, UUU and AAA were identified as phenylalanine and lysine codons, respectively.

A second type of experiment derived from the work of Gobind Khorana (Nobel Prize, 1968), who synthesized *polynucleotides of defined sequences.* Investigators used such polynucleotides as synthetic mRNAs in Nirenberg-type cell-free amino acid-incorporating systems and analyzed the polypeptides formed. By determining the amino acid sequence of the polypeptide, one could assign codons to its component amino acids (Figure 16.10). These studies involve two considerations. First, the nature of the

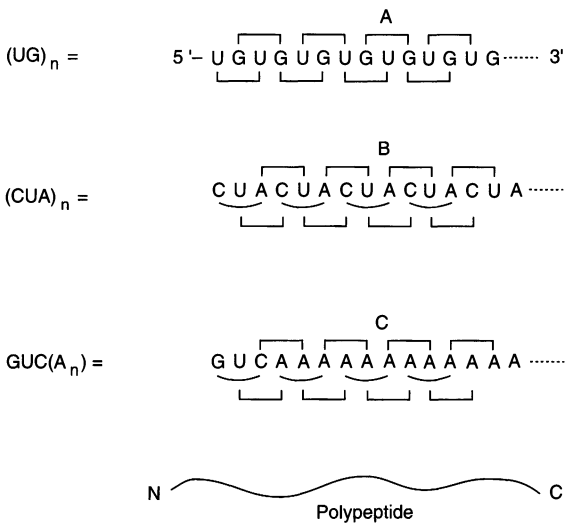


Figure 16.10. Use of synthetic polynucleotides as mRNA in cell-free amino acid-incorporating systems. Depending on the mRNA base at which translation starts, different polypeptides form: (A) one polypeptide chain consisting of an alternating sequence of cysteine (UGU) and valine (GUG); (B) three homopolypeptides, polyleucine (CUA), polytyrosine (UAC), and polythreonine (ACU); (C) polylysine (AAA) having either valine (GUC), serine (UCA), or glutamine (CAA) as N-terminus.

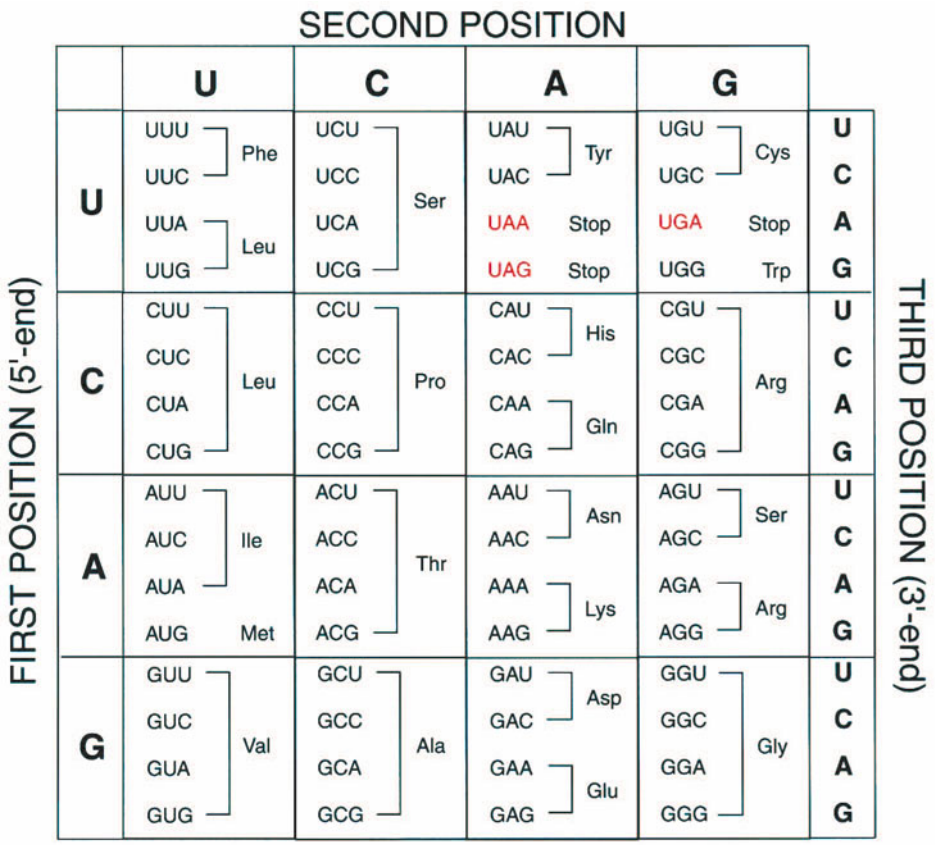


Figure 16.11. The genetic code dictionary. Codons are written in the 5′ → 3′ sense. Stop designates a termination codon.

polypeptide formed depends on the point at which its translation from the oligonucleotide begins; the starting point determines the reading frame for subsequent codons. Second, assigning codons to the amino acids depends on the direction in which the polynucleotides are translated. Because it was known that *mRNA translation proceeds from the 5′- to the 3′-end,* Khorana and co-workers could assign codons unambiguously.

16.3.3. Characteristics of the Genetic Code

Once all of the base sequences of the codons had been determined, the results were compiled into a complete **genetic code dictionary** (Figure 16.11). Analysis of this dictionary, as well as additional studies, revealed key properties of the genetic code.

16.3.3A. Degeneracy of the Code. We still do not fully understand the reasons for the degeneracy of the genetic code. Apparently, in some organisms or under certain conditions, preferential use of some synonym codons over others takes place. However, that degeneracy must have an important biological purpose is apparent from the fact that it is not a completely random phenomenon. When we examine degeneracy carefully, we can discern six unifying principles:

1. *The code is highly degenerate.* In fact, the code is almost totally degenerate. Of 64 codons, 61 code for amino acids; only three do not (UAG, UAA, UGA).

2. *Synonym codons for a given amino acid are related.* Synonym codons of an amino acid can be interconverted by single base changes except for some of the arginine, leucine, and serine synonym codons, whose interconversions require two base changes.

3. *Most of the degeneracy revolves about the third position (3′-end) of the codon.* For most amino acids, the first two codon positions are fixed, but the third one is variable. Arginine, leucine, and serine also show degeneracy in the first position. Variability in the 3′-end involves interconversion of U and C or interconversion of A and G. Thus, degeneracy is expressed by a change of one pyrimidine to another or a change of one purine to another:

X Y [U C] XY [A G]

Mutations resulting in single base changes are **point mutations. Transitions,** involving pyrimidine–pyrimidine or purine–purine conversions, are more likely to occur spontaneously than **transversions,** which require interconversions of purines and pyrimidines. Based on the observed degeneracy, we may conclude that the genetic code reflects the likelihood of spontaneous point mutations that have occurred over time.

4. *Most amino acids that have synonym codons show either a twofold or a fourfold degeneracy* so that they are coded by either two or four synonym codons. The only exceptions to this generalization are isoleucine, which has three synonym codons, and arginine, leucine, and serine, which have six synonym codons each.

5. *Amino acids that occur frequently in proteins have a higher degeneracy* than amino acids that occur infrequently. Leucine and serine are two common amino acids, and each has six synonym codons; methionine and tryptophan are relatively rare in proteins, and each is coded by a single codon.

6. *Related amino acids often have related codons.* Two examples are codons of the acidic amino acids aspartic acid (GAU, GAC) and glutamic acid (GAA, GAG) and codons of the hydroxy amino acids serine (UCU, UCC, UCA, UCG) and threonine (ACU, ACC, ACA, ACG).

16.3.3B. Universality of the Code.

The genetic code is essentially universal: a given codon codes for the same amino acid regardless of whether in an animal, a plant, or a bacterial system. Universality of the code provides strong support for the theory of evolution, since it indicates that all organisms have evolved from some common origin. If that were not so, different genetic codes would most likely have evolved.

However, the code is not completely universal. We know of two exceptions. In ciliated protozoa, AGA and AGG do not code for arginine but rather serve as termination codons. Possibly, this has evolved to provide the organisms with protection against infecting viruses using the "universal" code.

The second exception involves mitochondria that have their own DNA and protein-synthesizing machinery. Mitochondria use a genetic code that differs from the universal one. Actual differences depend on the organism involved. Table 16.4 lists distinctive codons of human mitochondria. In addition to the differences shown in the table, mitochondria achieve an economy by using a single codon for each of the eight four-codon families (Leu, Val, Ser, Pro, Thr, Ala, Arg, and Gly). For example, the leucine codons CUU, CUC, CUA, and CUG are replaced in mitochondria by the single codon GAU. All these changes combine to make the mitochondrial code a simpler code than the universal one. For the universal code, at a minimum, 32 different kinds of tRNA are required to "recognize" all of the 61 sense codons (see Problem 16.1). This calculation takes into account the wobble hypothesis discussed below. For the mitochondrial code, only 22 different tRNAs are required to "recognize" all of the amino acid codons.

Table 16.4. Distinctive Codons of Human Mitochondria

Codon	"Universal" code	Mitochondrial code
UGA	Stop	Trp
UGG	Trp	Trp
AUA	Ile	Met
AUG	Met	Met
AGA	Arg	Stop
AGG	Arg	Stop

In addition to using fewer tRNAs for protein synthesis, mitochondrial DNA requires fewer genes to specify tRNA. Thus, both translation and transcription involve a smaller number of components. Perhaps this reflects the differences in genome size: MW $\approx 1.0 \times 10^7$ for mitochondria and MW $\approx 2.0 \times 10^{12}$ for the cell nucleus (based on an average molecular weight of 700 for a base pair).

16.3.3C. Evolution of the Code.

The number of genetic codes that could theoretically be devised is truly astronomical. Theorists have estimated that one may construct 10^{71}–10^{84} triplet codes, based on 64 triplets coding for 20 amino acids. How then has the one genetic code used by living systems come into existence?

Proposed theories for the evolution of the genetic code fall into two schools of thought. According to the **frozen accident school,** assignment of codons to amino acids resulted from a random event, but once it occurred it bestowed a selective advantage on the translating system and hence became fixed. According to the **specific interactions school,** assignment of codons to amino acids resulted from specific physicochemical relationships between amino acids and their codons, anticodons, or corresponding base sequences in DNA. This theory is supported by the finding that polarity and hydrophobicity of amino acids correlate with the same properties of their anticodon nucleotides. Because most of the degeneracy of the code is in the third position of codons, scientists have suggested that the code has evolved from a doublet code, capable of coding for fewer amino acids, to a triplet code with the capacity to code for all 20 amino acids.

A great deal of survival value is built into the code because of its degeneracy. Resistance to potentially deleterious mutations and evolutionary adaptation to beneficial mutations are both enhanced by code degeneracy. Many mutations in DNA result in formation of corresponding synonym codons in mRNA. Because synonym

codons code for the same amino acid, the protein formed from the mutated mRNA has the same amino acid sequence as the original protein. Such mutations are **silent mutations** and represent changes in DNA that do not become expressed as amino acid changes in proteins.

Frequently, a mutation in DNA results in formation of an mRNA codon for a related amino acid. In such cases, the amino acid sequence of the protein will be altered, and the mutation is a **missense mutation.** However, because the amino acid is often replaced by a related one, the protein may still have some or most of its original structure, properties, and activity. Occasionally, an amino acid replacement may result in an improved protein that provides an evolutionary advantage to the organism.

Lastly, since only three codons out of 64 are termination codons, relatively few mutations will result in formation of termination codons in mRNA. Those that do are **nonsense mutations** and will lead to premature termination of protein synthesis and formation of an incomplete polypeptide chain.

16.3.3D. Uncoded Amino Acids.

As you can see from Figure 16.11, hydroxyproline and hydroxylysine do not have any codons. These amino acids are introduced into proteins by way of their parent amino acids. Proline and lysine become inserted into the growing polypeptide chain in response to their codons in mRNA and are then hydroxylated to form hydroxyproline and hydroxylysine as needed. Such modifications of the polypeptide chain are examples of **posttranslational processing.**

Table 16.5. Allowed Base Pairs According to the Wobble Hypothesis

Anticodon base (5′-end)	Codon base (3′-end)
C	G
A	U
U	A or G
G	U or C
I[a]	U, C, or A

[a]I, Inosine.

16.3.3E. Wobble Hypothesis.

Francis Crick proposed the *wobble hypothesis* in 1966 to account for two experimental observations. One was the finding that the third base (3′-end) of the codon appeared to be less specific and less critical for protein synthesis than the other two positions, as indicated by the extensive degeneracy at the 3′-end. A second observation was that some highly purified tRNAs could "recognize" several codons, indicating that the binding between tRNA anticodon and mRNA codon was not completely "rigid" but had some "play" in it.

We can state the **wobble hypothesis** as follows: The base at the 5′-end of the anticodon is not as spatially confined as the other two and can base-pair with *several* bases at the 3′-end of the codon. Specifically, Crick proposed the base pairings shown in Table 16.5. Subsequent research has provided support for the wobble hypothesis. The base at the 5′-end of the anticodon is the **wobble base.**

SUMMARY

In DNA replication, a DNA molecule leads to synthesis of identical DNA molecules. In transcription, a strand of DNA leads to synthesis of a strand of RNA. Synthesis of DNA from RNA is called reverse transcription. Translation is the process whereby a strand of RNA directs the synthesis of a polypeptide chain.

The central dogma of molecular biology states that both DNA and RNA can be replicated and transcribed and RNA can also be translated to yield protein. Polymerization reactions during replication, transcription, and translation involve primers and templates. Primers are linked covalently to the polymerization product; templates are not so linked.

Three types of RNA participate in protein synthesis—messenger, ribosomal, and transfer RNA (mRNA, rRNA, and tRNA). Messenger RNA is transcribed from one strand of DNA and is complementary and antiparallel to that strand. In eukaryotes, mRNA is synthesized in the nucleus and passes into the cytoplasm, where it associates with ribosomes, the sites of protein synthesis. Several different types of rRNA exist and all form part of the ribosome. Ribosomes consist of two roughly globular subunits. Each subunit contains a large number of protein molecules and one or several RNA molecules. Transfer RNAs attach to amino acids covalently and trans-

port them to ribosomes for polymerization. There exist at least 20 different tRNAs, one for each amino acid. The polynucleotide strand of tRNA folds back on itself to form double-stranded helical segments in which complementary base pairs are H-bonded. In two dimensions, tRNA looks like a cloverleaf. In three dimensions, tRNA is L-shaped.

Genetic information is contained in the sequence of bases in a nucleic acid. The manner in which this information specifies the sequence of amino acids in a polypeptide chain is known as the genetic code. Messenger RNA is transcribed from one strand of DNA. Sequences of three bases in mRNA constitute codons that code for amino acids. Codons do not overlap and are not separated by spacers. Hence, the genetic code is a triplet, nonoverlapping, and unpunctuated code. Techniques for deciphering the genetic code included cell-free amino acid-incorporating systems and ribosome binding assays. The genetic code dictionary contains 64 codons and is degenerate; 61 codons code for amino acids, and 3 are termination codons. Most amino acids have several synonym codons. Degeneracy of the code endows it with a great deal of survival value. The genetic code is essentially universal, but the mitochondrial code differs from the "universal" code. The wobble hypothesis defines the types of base pairs that can form by H-bonding between the anticodon of tRNA and the codon of mRNA.

SELECTED READINGS

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D., *Molecular Biology of the Cell,* 3rd ed., Garland, New York (1994).

Brenner, S., Jacob, F., and Meselson, M., An unstable intermediate carrying information from genes to ribosomes for protein synthesis, *Nature (London)* 190:576–581 (1961).

Crick, F. H. C., Codon–anticodon pairing: The wobble hypothesis, *J. Mol. Biol.* 19:548–555 (1966).

Crick, F. H. C., Central dogma of molecular biology, *Nature (London)* 227:561–563 (1970).

Ferreira, R., Why 4^3 codons? *Z. Naturforsch. C* 50:148–152 (1995).

Glick, B. R., and Pasternak, J. J., *Molecular Biotechnology: Principles and Applications of Recombinant DNA,* ASM Press, Washington, D.C. (1994).

Hill, W. E. (ed.), *The Ribosome: Structure, Function, and Evolution,* American Society for Microbiology, Washington (1990).

Judson, H. F., *The Eighth Day of Creation,* Simon & Schuster, New York (1979).

Khorana, H. G., Nucleic acid synthesis in the study of the genetic code, *Nobel Lectures in Molecular Biology, 1933–1975,* Elsevier, New York (1977).

Morgan, R. A., and Anderson, W. F., Human gene therapy, *Annu. Rev. Biochem.* 62:191–217 (1993).

Nierlich, D. P., and Murakawa, G. J., The decay of bacterial messenger RNA, *Prog. Nucleic Acid Res. Mol. Biol.* 52:153–216 (1996).

Nirenberg, M. W., and Matthaei, J. H., The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, *Proc. Natl. Acad. Sci. USA* 47:1588–1602 (1961).

Rich, A., and Kim, S. H., The three-dimensional structure of transfer RNA, *Sci. Am* 238(1):52–62 (1978).

Tjian, R., Molecular machines that control genes, *Sci. Am.* 272:54–61 (1995).

Watson, J. D, *Recombinant DNA,* 2nd ed., W. H. Freeman, New York (1992).

A. Define each of the following terms:

Elongation	Posttranslational processing
Polysome	Initiation
Isoacceptor tRNAs	Aminoacyl-tRNA
Sense strand	Central dogma
Genetic code	Wobble hypothesis
Degeneracy	Molecular biology
Transfer RNA	Messenger RNA
L-shaped structure	Ribosome binding assay

B. Differentiate between the two terms in each of the following pairs:

Codon/anticodon	Synonym codon/nonsense codon
Silent mutation/missense mutation	Overlapping code/nonoverlapping code
Transition/transversion	Primer/template
Arm/variable arm	snRNA/snRNP

C. (1) What is the wobble hypothesis? Why was it proposed and what is its significance?
(2) Outline the approaches used to decipher the genetic code with respect to (a) base composition of codons and (b) base sequence of codons.
(3) What are the major characteristics of the genetic code?
(4) Discuss the structure and properties of transfer RNA.
(5) What are the differences between: (a) prokaryotic and eukaryotic ribosomes and (b) the universal and mitochondrial genetic codes?

PROBLEMS

16.1.* Show that the minimum number of different tRNAs required to bind to ("recognize") all of the amino acid codons of the "universal" code is 32. Note that methionine (codon AUG) requires two tRNAs for its specification, one for ordinary methionine and one for methionine functioning as initiator of protein synthesis. Base your reasoning on the first four wobble rules listed in Table 16.5.

16.2.* Refer to the previous problem and show that the minimum number of different tRNAs required to bind to ("recognize") all of the amino acid codons in human mitochondria is 22.

16.3. The codon of an amino acid has an anticodon whose base sequence is 5′-CUG-3′. What is the base sequence in DNA from which the amino acid codon (sequence in mRNA) is transcribed?

16.4. What is the base sequence in DNA from which the tRNA anticodon of the previous problem is transcribed?

16.5. What polypeptide products would be obtained in a Nirenberg-type experiment (cell-free amino acid incorporation) using the oligonucleotide $AG(CU)_nGG$ as synthetic mRNA?

16.5. What is the probability that a given triplet in the random copolymer poly(GAU) (25:30:45, in mole percent) would contain one G, one A, and one U?

16.7. Write out the *actual* sequences of all possible triplets in the random copolymer poly(UC) (80:20, in mole percent). What are the probabilities that a given triplet would have any of these specific base sequences?

16.8. In a Nirenberg-type experiment, the oligonucleotide $ACA(GCC)_n$ is used as synthetic mRNA. How can you determine the direction in which the mRNA is being translated from the nature of the products obtained?

16.9. We supply a cell-free amino acid-incorporating system with the synthetic oligonucleotide $UGC(CUA)_n$ and obtain a number of different polypeptides. Among the products, we find $Cys(Leu)_7$ when the incubation time is short and $Cys(Leu)_{20}$ when the incubation period is longer. What does that tell you about the direction of polypeptide chain growth?

16.10.* Scientists have discovered a new bacteriophage and isolated its DNA. The DNA has the following base composition (mol %): A = 21, T = 30, C = 30, G = 19. The DNA is not attacked by either phosphodiesterase I or II (see Table 7.5). What can you conclude from these data?

16.11. The base sequence in the gene (sense strand) coding for an oligopeptide is as follows:

5′ A—C—T—T—T—G—A—C—G—G—A—C—
T—G—C—C—C—T—A—G—A 3′

What is the peptide's amino acid sequence?

16.12. Some mutations lead to insertion or deletion of a base pair in DNA. When this occurs, the normal sequence of amino acids coded for by the DNA is garbled, beginning at the mutation site. Because the reading frame is shifted, these are known as *frameshift mutations*. Write out the amino acid sequence of the peptide formed when (a) an insertion or (b) a deletion of a base occurs in the nucleotide sequence of the previous problem at the points indicated:

insertion of A ↓ deletion of A ↓

5′ A—C—T—T—T—G—A—C—G—G—A—C—
T—G—C—C—C—T—A—G—A 3′

16.11. What products would you obtain in a cell-free amino acid-incorporating system when you use the oligonucleotide $(AGA)_n$ as synthetic mRNA in (a) a cytoplasmic system or (b) a mitochondrial system?

16.14. Prove that the change from normal to sickle-cell hemoglobin (see Section 3.2) can result from a single base mutation.

16.15. How many base pairs are there in the gene coding for human gastrin (see Table 2.5)?

16.16. We grow one batch of *E. coli* cells in a "light" medium, containing ^{12}C and ^{14}N, and a second batch in a "heavy" medium, containing ^{13}C and ^{15}N. We isolate the ribosomes after they become labeled with these isotopes and dialyze separate aliquots of the "light" and "heavy" ribosomes extensively against water. We then combine the two aliquots into one solution and dialyze it against a buffer

containing a high concentration of magnesium ions. What types of ribosomal particles would you expect to find in the solution after dialysis?

16.17.* Tobacco mosaic virus is an RNA-containing virus in which the RNA codes for coat proteins. We can isolate and treat the RNA with HNO_2 to form mutated RNA in which cytosine residues have been deaminated and converted to uracil. We can then use this RNA to produce intact viral particles containing mutated coat proteins. When we compare the mutated proteins with the normal ones we find that, among other changes, some positions that in normal proteins are filled by leucine are now occupied by a different amino acid. Identify this amino acid.

16.18. The gene for a transfer RNA molecule contains the sequence 5′-AATCGTC-3′ in the *nontranscribed* strand. What is the corresponding base sequence in the tRNA?

16.19. Studies of protein amino acid sequences have shown that any amino acid can be preceded by, or followed by, any other amino acid. These data provide strong evidence against the existence of any kind of overlapping code. To illustrate this, assume that the code were overlapping and commaless such that most codon bases were used three times:

```
mRNA:     A—C—U—G—A—U———
Codons:   A—C—U
            C—U—G
              U—G—A
                G—A—U, etc.
```

Such a code would impose restrictions on amino acid sequences. Aspartic acid, for example, would always have to be followed by one of several specific amino acids. What are they?

16.20. Would a single base mutation in the overlapping code of the previous problem be likely to result in one, two, or three amino acid changes in the protein?

16.21. Which of the following pairs of aminoacyl-tRNA and trinucleotide would result in binding to ribosomes?

(a) Seryl-$tRNA^{Ser}$; AGC
(b) Glycyl-$tRNA^{Ser}$; AGU
(c) Seryl-$tRNA^{Gly}$; GGC
(d) Seryl-$tRNA^{Ser}$; GCU

16.22. According to the wobble hypothesis, what codons should be recognized by the following anticodons?

(a) 5′-IGG-3′
(b) 5′-UGA-3′
(c) 5′-GCA-3′

What amino acid residues correspond to these codons?

16.23. Which of the following amino acid changes could be produced by a single point mutation in DNA?

(a) Phe → Ile
(b) Gly → Pro
(c) Arg → Ser
(d) Cys → Val
(e) Ala → Thr
(f) His → Tyr
(g) Leu → Gly

16.24. What is the base composition of an mRNA section whose translation yields polytryptophan? What is the base composition of the gene that codes for this polypeptide?

16.25.* Assume that each of the following amino acid changes results from a single point mutation in DNA. On that basis, deduce the original mRNA codon that specified leucine:

Leu → Pro → Ala → Gly → Trp

16.26. One of the threonine codons is 5′-ACU-3′. What is the sequence in DNA that is transcribed into the anticodon sequence of the corresponding threonine-tRNA?

16.27.* The double-stranded DNA of a bacterium has a length of 4.0×10^6 base pairs (4.0×10^6 nucleotides/strand). The 16S rRNA from the same organism has a length of 1500 nucleotides. The DNA is fragmented, heat-denatured, and then hybridized with the 16S rRNA. A maximum of 0.094% of the DNA can be converted to a DNA–RNA hybrid in the presence of excess 16S rRNA. How many times does the gene for 16S rRNA occur per double-stranded DNA molecule?

16.28.* A student uses the copolymer poly(ACG) (20:30:50, in mole percent) for a Nirenberg-type experiment and finds that the relative incorporation of arginine, proline, and lysine is in the ratio of 6:3:1. Given that the codon for proline is CCC, what are the likely codons for arginine and lysine?

16.29. How long is the gene (in number of base pairs in DNA) that codes for an oligomeric protein composed of four identical subunits? The subunits are held together noncovalently, and each is a single polypeptide chain consisting of 104 amino acids.

Replication—The Synthesis of DNA

17

In their famous, historic paper describing the double helix, Watson and Crick concluded: "It has not escaped our notice that the specific pairing we have postulated immediately suggests a copying mechanism for the genetic material." In a subsequent paper, the authors elaborated on this statement by proposing that the copying mechanism may involve each DNA strand serving as a template for synthesis of a complementary strand.

This proposal has received extensive experimental support and is now accepted as the correct version of the process whereby the genetic information of DNA is copied. The copying process, called **semiconservative replication** (Figure 17.1), requires that the two parental strands unwind and separate and that each strand serves as a *template* for synthesis of a *complementary* and *antiparallel* strand. Separation of this four-strand complex yields two identical *daughter* duplexes. *Each daughter duplex consists of one parental strand and one newly synthesized strand (hence the term semiconservative).* Semiconservative replication occurs in both prokaryotes and eukaryotes, but details of the process vary. An elegant experiment performed by Matthew Meselson and Franklin Stahl in 1957 provided direct experimental proof for semiconservative replication.

17.1. SEMICONSERVATIVE REPLICATION

The **Meselson–Stahl experiment** used *density gradient centrifugation* (see Appendix C), which allows separation of molecules on the basis of differences in their densities (mass/unit volume). In this technique, molecules having a specific density form a band in a density gradient where their density equals that of the gradient.

The Meselson–Stahl experiment consisted of growing *E. coli* cells in a "heavy" medium that contained $^{15}NH_4Cl$ as sole nitrogen source. After several generations (the generation time of *E. coli* is about 30 min), the DNA was fully labeled with the heavy isotope; all of the purine and pyrimidine nitrogens were ^{15}N atoms. Cells were then transferred to a "light" medium, which contained ordinary $^{14}NH_4Cl$ as sole nitrogen source, and allowed to continue their growth. Consequently, any DNA synthesized after the transfer contained only ordinary nitrogen. Aliquots of the "light" medium culture were collected as a function of time, and the DNA was analyzed by density gradient centrifugation.

The gradient had a spread of densities from 1.650 (top of the centrifuge tube) to 1.800 g/cm^3 (bottom of the tube), suitable for detecting the various types of labeled *E. coli* DNA. DNA in which both strands contain ^{15}N constitutes *fully labeled DNA*. This "all heavy" DNA has a

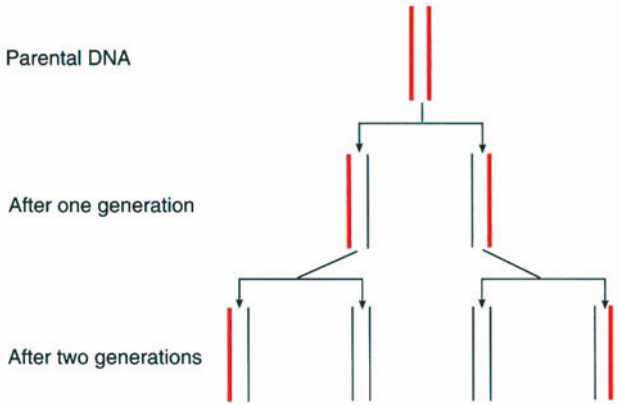


Figure 17.1. Semiconservative replication of DNA.

density of 1.710 g/cm^3 at pH 7.0. DNA in which both strands contain ^{14}N constitutes *unlabeled DNA*. This "all light" DNA has a density of 1.695 g/cm^3 at pH 7.0. Other forms of DNA have densities that are intermediate between these two values.

All of the DNA isolated at the moment of cell transfer, or time zero, banded at a point in the gradient corresponding to the density of "all heavy" DNA (Figure 17.2). As time went on, a new band appeared that corresponded to a density of 1.703 g/cm^3. After exactly one generation, all of the DNA was present in this new band. Since the density of this "hybrid" DNA was intermediate between that of "all heavy" and that of "all light" DNA, it appeared to consist of 50% labeled and 50% unlabeled DNA. DNA isolated from cells grown for more than one generation was characterized by the appearance of yet a third band that increased in intensity with time. This band had a density of 1.695 g/cm^3 and corresponded to "all light" DNA. The appearance of three bands having the above densities is in complete agreement with the semiconservative model of replication.

Meselson and Stahl carried out important follow-up experiments. They heat-denatured the "hybrid" DNA produced after one generation to separate its two strands. Centrifugation of this denatured DNA yielded two bands of equal intensity that corresponded to "all light" and "all heavy" DNA. This indicated that the "hybrid" DNA was a duplex, composed of one light and one heavy strand. Alkaline denaturation of the DNA produced after two generations by suspension in buffer at pH 12 also yielded two bands. One band represented three-quarters of the total DNA and corresponded to "all light" DNA; the other represented one-quarter of the DNA and corresponded to "all heavy" DNA. These findings proved conclusively that replication proceeded via a semiconservative mechanism.

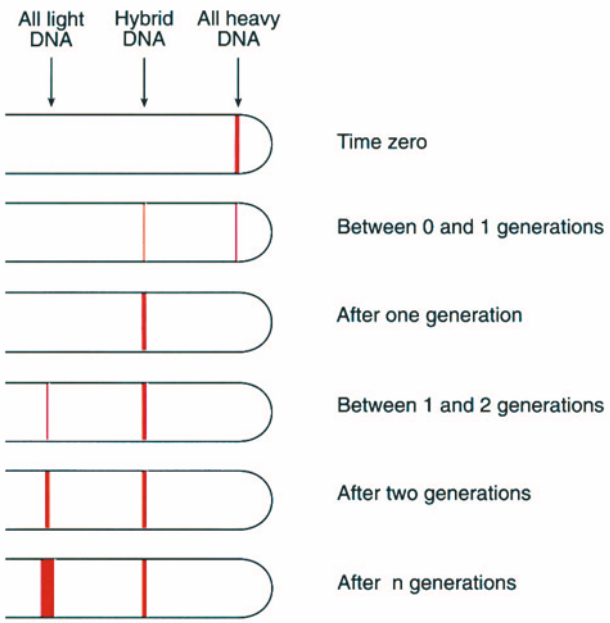


Figure 17.2. The Meselson–Stahl experiment. Determining the distribution of labeled and unlabeled DNA in a density gradient over several generations revealed only three different bands.

17.2. DNA POLYMERASE

DNA polymerase (DNA-dependent DNA polymerase) is the key enzyme of DNA replication. We may schematically represent the polymerization as follows:

$$n(\text{dNTP}) \xrightarrow[Mg^{2+}]{\text{DNA template}} [\text{DNA}] + n(PP_i)$$

$$n(PP_i) \xrightarrow{\text{pyrophosphatase}} 2n(P_i)$$

where dNTP is deoxyribonucleoside triphosphate (dATP, dCTP, dGTP, dTTP). The enzyme requires all four deoxyribonucleotides as well as magnesium ions. Each DNA strand serves as a template for synthesis of a complementary and antiparallel strand (Figure 17.3). The mechanism of the reaction involves a nucleophilic attack by the 3′-OH of the growing chain on the α-phosphate group attached to the 5′-carbon of a deoxyribonucleoside triphosphate. Pyrophosphate is cleaved off as the new deoxyribonucleoside monophosphate becomes incorporat-

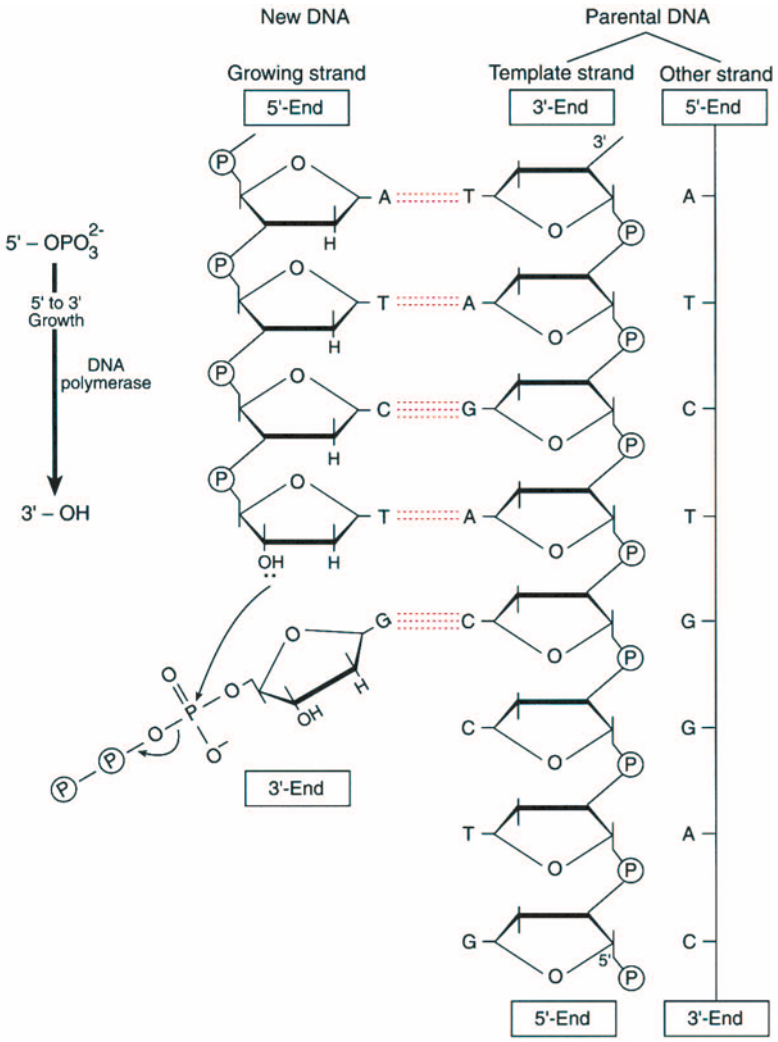


Figure 17.3. Mechanism of action of DNA polymerase. Deoxyribonucleoside 5′-triphosphates serve as substrates, and the new DNA is complementary and antiparallel to the template strand.

ed. We term this mode of polymerization **growth in the 5′→ 3′ direction**. The 5′-end of the strand is synthesized first, and the 3′-end is synthesized last; elongation of the strand proceeds from the 5′- to the 3′-end.

DNA polymerase requires a free 3′-OH group as primer for initiation of polymerization. Like all primers, that for DNA polymerase becomes covalently linked to the reaction product. Because the new DNA strand grows from the 5′- to the 3′-end, the enzyme must move along the template strand by proceeding from its 3′-end to its 5′-end.

The biochemical standard free energy change for the polymerase reaction is actually positive (+2.1 kJ mol^{-1}). Under standard conditions, and in the absence of pyrophosphatase, DNA polymerase catalyzes the conversion of DNA to dNTPs. Thus, the polymerization is dependent on *pyrophosphatase,* which catalyzes hydrolysis of *pyrophosphate (PP_i),* an energy-rich compound, to inorganic phosphate. The highly negative free energy change of PP_i hydrolysis drives DNA synthesis.

17.2.1. DNA Polymerase I in Prokaryotes

DNA polymerase was first isolated from *E. coli* by Arthur Kornberg in 1957 (Nobel Prize, 1959). The isolation was quite a feat—100 kg of bacterial cells yielded 500 mg of purified enzyme. We now refer to this enzyme as **DNA polymerase I (pol I)** or **Kornberg enzyme.**

DNA polymerase I is a *multifunctional enzyme,* consisting of a single polypeptide chain (MW = 103,000) that contains three distinct enzymatic activities: 5′ → 3′ polymerase activity and two *exonuclease* activities, designated 3′ → 5′ exonuclease and 5′ → 3′ exonuclease (see below).

Limited proteolysis of DNA polymerase I with *subtilisin* or *trypsin* cleaves it into two unequal fragments. The larger, C-terminal fragment, called the *Klenow fragment* (MW = 68,000), contains the polymerase and 3′ → 5′ exonuclease activities. The smaller, N-terminal fragment (MW = 35,000), contains the 5′ → 3′ exonuclease activity.

DNA polymerase I is a *processive* enzyme—it catalyzes a number of polymerization steps without dissociating from the template. Researchers believe that the enzyme's **processivity** results from a deep cleft present in its structure. The cleft is large enough to accommodate a section of B-DNA and represents the active site of the enzyme. During replication, the cleft is almost completely closed, holding the DNA in position for successive enzymatic steps.

17.2.2. Exonuclease Activities in Prokaryotes

The exonuclease activities associated with DNA polymerase I have different properties and different functions. **3′ → 5′ Exonuclease** catalyzes removal of mismatched nucleotides at the 3′-end of the growing strand (Figure 17.4). The enzyme catalyzes removal of one nucleotide at a time and, because of its 3′ → 5′ specificity, acts in a direction opposite to that of the polymerase. Despite the polymerase's great fidelity, occasionally its action may lead to incorporation of a wrong nucleotide. Because all four deoxyribonucleoside triphosphates serve as enzyme substrates, a wrong base pair may occasionally form, much as the anticodon and the codon can become linked in nonstandard fashion (see the wobble hypothesis in Section 16.3). If polymerase action has resulted in such wrong incorporation of one or more nucleotides (*misincorporation*), the exonuclease catalyzes excision of the wrong nucleotides. Accordingly, the 3′ → 5′ exonuclease performs a **proofreading** or **editing function.**

5′ → 3′ Exonuclease catalyzes a reaction that proceeds in the same direction as that of the polymerase but

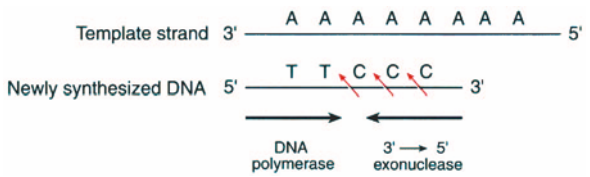


Figure 17.4. Mode of action of 3′ → 5′ exonuclease, the proofreading function of DNA polymerase. The enzyme catalyzes removal of misincorporated nucleotides, one at a time.

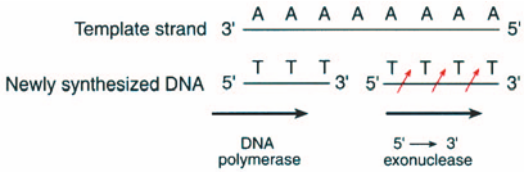


Figure 17.5. Mode of action of 5′ → 3′ exonuclease. The enzyme catalyzes removal of mono- or oligonucleotides (ribo- or deoxyribonucleotides).

takes place ahead of the polymerase (Figure 17.5). The enzyme functions to clear the path for the polymerase. 5′ → 3′ Exonuclease catalyzes the cleavage of both mono- and oligonucleotides (1–10 nucleotides), which can be *either* deoxyribonucleotides or ribonucleotides. Removing the ribonucleotides H-bonded to a template DNA strand constitutes the major function of 5′ → 3′ exonuclease in DNA replication (Section 17.4).

When the polymerase and 5′ → 3′ exonuclease activities of DNA polymerase I function at comparable rates, we refer to the combined effect as **nick translation** (Figure 17.6). In this process, a nick in an otherwise intact DNA proceeds down the double strand without any net synthesis of DNA. Researchers use nick translation to prepare radioactively labeled DNA for *in vitro* studies.

17.2.3. DNA Polymerases II and III in Prokaryotes

Biochemists first thought that DNA polymerase I was the enzyme responsible for DNA replication in *E. coli.* However, in 1969 John Cairns isolated an *E. coli* mutant that could grow normally but was deficient in DNA polymerase I activity. This stimulated a search for additional DNA-polymerizing activities. Because DNA poly-

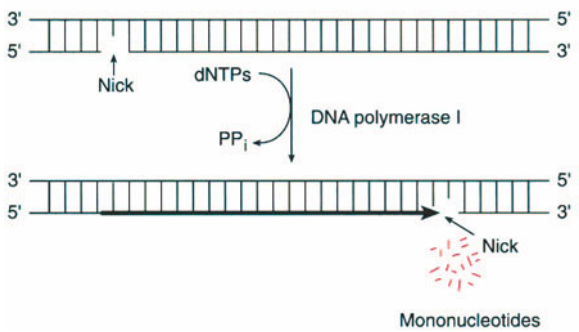


Figure 17.6. Nick translation. Combined action of DNA polymerase and 5′ → 3′ exonuclease causes a nick to move down a double strand without net DNA synthesis.

Table 17.1. Properties of DNA Polymerases from *E. coli*

Property	Pol I	Pol II	Pol III
5′ → 3′ Synthesis	Yes	Yes	Yes
Need free 3′-OH	Yes	Yes	Yes
Template-directed	Yes	Yes	Yes
3′ → 5′ Exonuclease	Yes	Yes	Yes
5′ → 3′ Exonuclease	Yes	No	No
Molecular weight	103,000	90,000	130,000
Molecules/cell	400	100	10–20
Nucleotides polymerized per minute per enzyme molecule (37°C)	600	30	9,000
Processivity	3–200	10,000	500,000
Major function	DNA repair, DNA replication	DNA repair	DNA replication

merase I generally has a very high level of activity, its absence from the cell made it possible to detect the existence of two other prokaryotic DNA polymerases—**DNA polymerase II** and **DNA polymerase III** (**pol II** and **pol III**). Although DNA polymerase I participates in DNA replication, the major replication enzyme is DNA polymerase III. In addition to its role in replication, DNA polymerase I serves as a key enzyme of DNA repair. DNA polymerase II also functions in DNA repair. Table 17.1 summarizes some properties of the three polymerases.

DNA polymerase III functions *in vivo* as part of a large multisubunit complex, called **DNA polymerase III holoenzyme.** The complex is a dimer with a molecular weight of about 900,000. Each monomer consists of 10 different types of subunits (Table 17.2). The complex formed by the α, ϵ, and θ subunits constitutes the "core enzyme." Two β subunits encircle the DNA and function like a sliding clamp, holding the enzyme to the template and increasing its processivity tremendously.

Table 17.2. Subunit Structure of DNA Polymerase III Holoenzyme[a]

Subunit	Molecular weight	Function
α	130,000	Polymerase
ϵ	28,000	3′ → 5′ Exonuclease/processivity
θ	10,000	Processivity
τ	71,000	Dimerization
γ	48,000	DNA binding/processivity
β	41,000	DNA binding/processivity
δ	34,000	DNA binding/processivity
χ	12,000	Processivity
ψ	14,000	Processivity
δ′	32,000	Processivity

[a]Adapted, with permission, from C. S. McHenry, *Annu. Rev. Biochem.* 57:519–550 (1988). Copyright © 1988 by Annual Reviews, Inc.

Both the processivity and the efficiency of DNA polymerase III exceed those of DNA polymerase I. These differences seem appropriate when you consider the roles that the enzymes have in DNA replication. DNA polymerase III catalyzes synthesis of long sections of DNA, while DNA polymerase I functions in closing short gaps between successive DNA fragments.

17.2.4. Eukaryotic DNA Polymerases

Scientists have identified five different DNA polymerases in eukaryotes (Table 17.3): **DNA polymerases α, β, γ, δ,** and **ε** (**pol α, β, γ, δ,** and **ϵ**). These enzymes, like prokaryotic DNA polymerases, use deoxyribonucleoside triphosphates and catalyze template-directed synthesis of DNA in the 5′ → 3′ direction.

Studies employing inhibitors have been particularly helpful in elucidating the functions of eukaryotic DNA polymerases, since the enzymes differ in their susceptibility to particular inhibitors. For example, *aphidicolin* (Figure 17.7), a fungal product that has a steroidlike structure, is a specific inhibitor of DNA polymerases α, δ, and ϵ. By contrast, *2′,3′-dideoxyribonucleotides* (Section 7.7) strongly inhibit DNA polymerases β and γ.

DNA polymerase α plays a similar role in eukaryotes as DNA polymerase III does in prokaryotes; it serves as the major replication enzyme and is tightly associated with *primase* (Section 17.4). The rate of polymerization with DNA polymerase α is about 500 nucleotides/min, or about 1/20 the rate with DNA polymerase III. DNA polymerase δ also functions in replication of nuclear DNA, but the two enzymes have different roles. DNA polymerase α catalyzes the *discontinuous synthesis* of the *lagging strand* (explained in Section 17.4), while DNA polymerase δ catalyzes the *continuous synthesis* of the *leading strand.*

Three of the polymerases (γ, δ, and ϵ) possess a

Table 17.3. Properties of Eukaryotic DNA Polymerases

Property	Pol α	Pol β	Pol γ	Pol δ
5′ → 3′ Synthesis	Yes	Yes	Yes	Yes
Need free 3′-OH	Yes	Yes	Yes	Yes
Template-directed	Yes	Yes	Yes	Yes
3′ → 5′ Exonuclease	Yes	No	Yes	Yes
5′ → 3′ Exonuclease	No	No	No	No
Molecular weight	250,000	39,000	200,000	170,000
Sensitivity to aphidicolin	High	Low	Low	High
Processivity	Moderate	Low	High	Low
Major function	Replication of nuclear DNA (lagging strand)	Repair of nuclear DNA	Replication of mitochondrial DNA	Replication of nuclear DNA (leading strand)
Location	Nucleus	Nucleus	Mitochondria	Nucleus

3′ → 5′ exonuclease activity with proofreading capability. An accessory protein probably serves as the editing function for DNA polymerase α. It appears that 5′ → 3′ exonuclease activities are not part of the polymerases but exist as separate molecular entities.

17.3. CHARACTERISTICS OF THE REPLICATING SYSTEM

Since the best understood replicative system is that of *E. coli,* we will concentrate on it for the rest of this chapter. While we emphasize prokaryotic replication, we will discuss eukaryotic systems as appropriate.

17.3.1. The Polarity Problem

Semiconservative replication of DNA, as exemplified by the Meselson–Stahl experiment, implies that double-stranded DNA must open up during replication. Synthesis of a nucleotide strand complementary to one strand of parental DNA requires that the two parental strands separate. Only then can the necessary base pairing between the new strand and the template strand take place. However, because the two strands are intertwined (*plectonemic coiling*), they cannot separate without unwinding the helix; unwinding constitutes a prerequisite of replication.

Simultaneous copying of the two unwinding parental strands represents a fundamental dilemma. One strand runs in the 5′ → 3′ direction while the other runs in the opposite direction, 3′ → 5′ (Figure 17.8). How can these two strands be replicated simultaneously?

Because all DNA polymerases catalyze synthesis of DNA in the *5′ → 3′ direction* only, the polarity of synthesis dictates that the *enzyme must move in opposite directions* along the two template strands (Figure 17.8). Only then can both strands be replicated simultaneously and can the two new strands both grow in the 5′ → 3′ di-

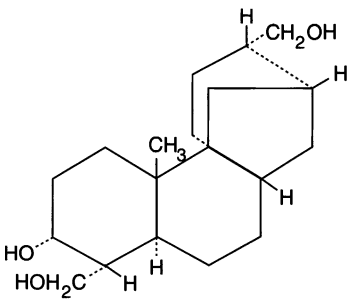


Figure 17.7. Aphidicolin.

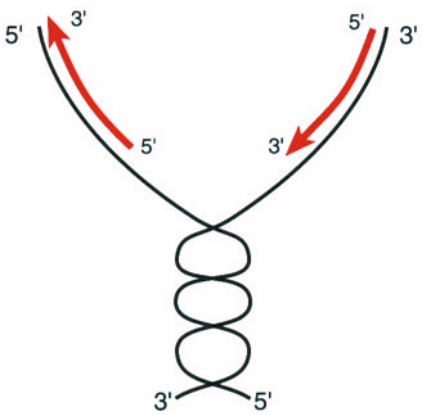


Figure 17.8. A replication fork. Because DNA polymerase catalyzes synthesis in the 5′ → 3′ direction only, it must move in opposite directions along the two parental strands.

rection. As we will see, in *E. coli* the polymerase (pol III) occurs as a dimer, with one subunit replicating one template strand and the second subunit replicating the other strand.

17.3.2. Initiation Points

Initiation of replication does not occur randomly along the DNA molecule. Instead, there exist specific sites, called **origins of replication (*ori*),** at which replication is initiated. In *E. coli,* there occurs one such site per molecule of DNA, called the *ori*C site. The site has a unique base sequence of 245 base pairs.

The *ori*C site is rich in A/T base pairs, which facilitates unwinding of the helix. Binding of a specific protein (*dnaA protein*) to *ori*C initiates the process. Unwinding of the helix and synthesis of the new strands results in a Y-shaped structure called the **replication fork** (see Figure 17.8).

The number of replication forks per molecule of DNA varies among different replicative systems. Mitochondrial DNA and some viral DNAs are replicated by specialized mechanisms (discussed in Section 17.6) that involve one replication fork. Replicating bacterial DNA has two replication forks, and eukaryotic DNA has many replication forks. Multiple starting points are needed in eukaryotes because eukaryotic DNA is much longer and its rate of replication much slower than in prokaryotes. *E. coli* DNA replicates at a rate of about 60,000 bp/min per replication fork. In other words, each strand becomes extended at the rate of 60,000 bases per minute. Eukaryotic DNA replicates at a rate of about 4000 bp/min per replication fork.

We refer to the functional unit of DNA replication as a **replicon.** A replicon contains a unique origin at which replication begins and a termination site at which it ends. An entire replicon is replicated once in the life cycle of a cell. In bacteria, the entire chromosome functions as a single replicon. In eukaryotes, each chromosome consists of many replicons.

17.3.3. Movement of Replication Forks

Replicative systems differ not only in the number of replication forks per molecule but also in the movement of these forks. Replication can be **unidirectional,** one replication fork moving in one direction, or **bidirectional,** two replication forks moving in opposite directions. Unidirectional replication occurs in the D-loop mechanism of mitochondria and the rolling circle mechanism of viruses (see Section 17.6). Bidirectional replication occurs in all other prokaryotic and eukaryotic systems.

Because bidirectional replication of circular bacterial DNA results in a structure that resembles the Greek letter theta, we also call it **theta replication** (Figure 17.9). Eukaryotic DNA replication, involving multiple replication forks moving in opposite directions, leads to formation of "eyes" or "bubbles" in the DNA (Figure 17.10).

Occurrence of a theta structure in the replication of bacterial DNA was demonstrated by John Cairns and his co-workers in 1963. These researchers grew *E. coli* cells in the presence of ^{3}H-labeled thymidine so that newly synthesized DNA was labeled with tritium. They then used *autoradiography* (see Appendix C) to detect the incorporated label; because tritium is a weak β emitter, the film was exposed to the labeled DNA for two months. In this milestone experiment, DNA was "caught in the act" of replication (Figure 17.11).

That DNA replication in *E. coli* is bidirectional was confirmed by other workers who also used autoradiography and ^{3}H-labeled thymidine. They isolated DNA from cells that were first grown in a medium containing slightly labeled thymidine and were then transferred to

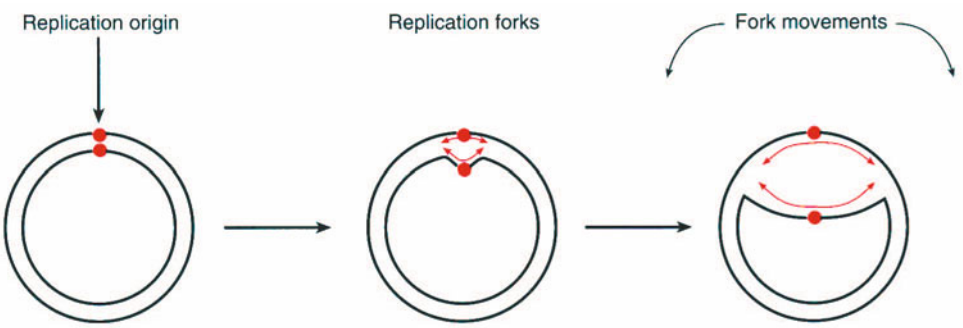
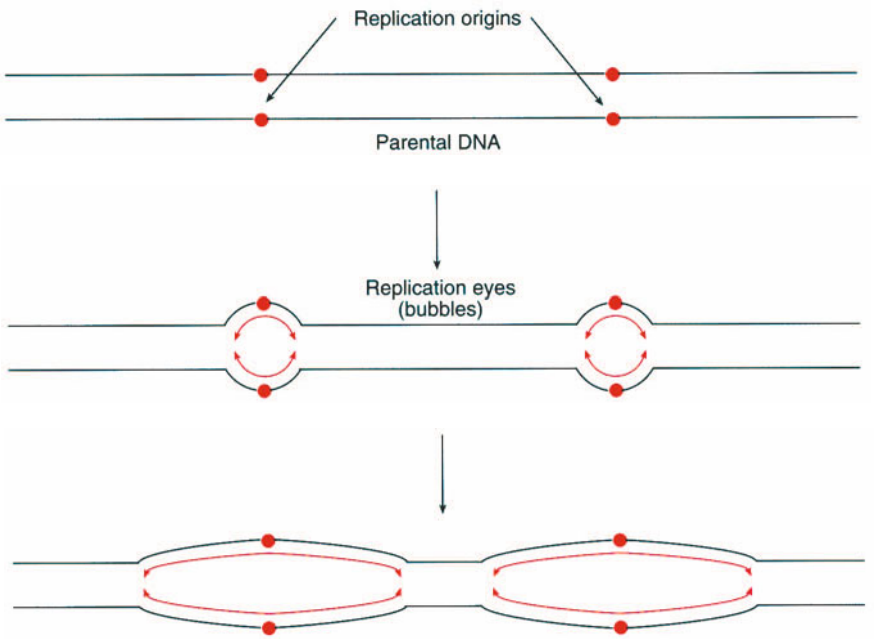


Figure 17.9. Bidirectional replication of prokaryotic DNA.



Bidirectional replication of eukaryotic DNA.

a medium containing extensively labeled thymidine. DNA containing slightly labeled thymidine leads to slight exposure of the photographic film and produces a track of low grain density. On the other hand, DNA containing extensively labeled thymidine leads to extensive exposure of the film and produces a track of high grain density. Because a track of low grain density was found to be bracketed by two tracks of high grain density (Figure 17.12), the experiment showed that replication was bidirectional.

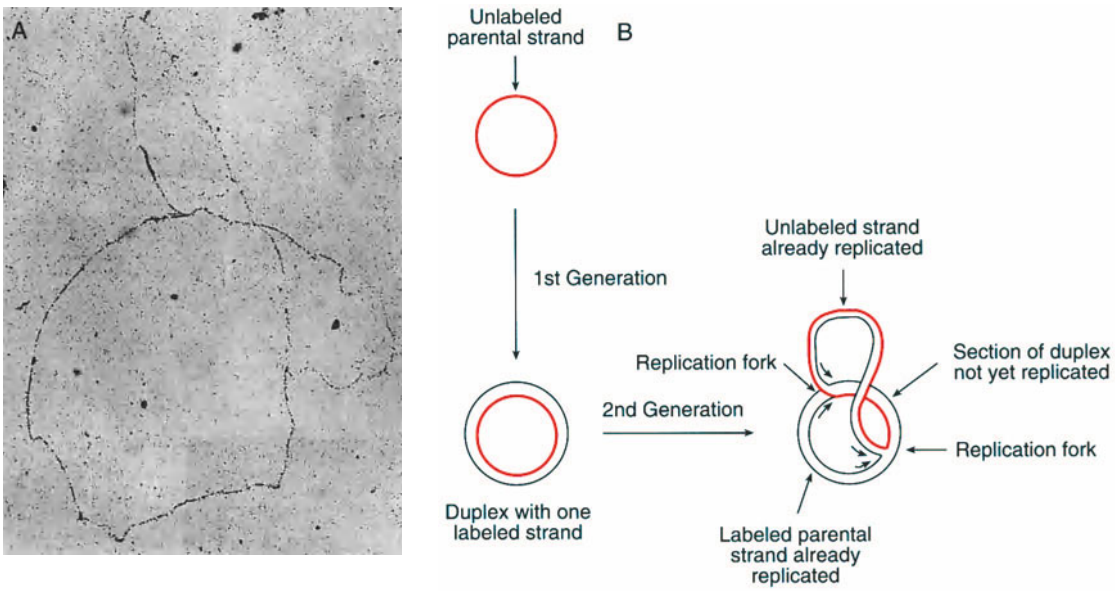


Figure 17.11. Replication of *E. coli* DNA in cells grown in a medium containing [^{3}H]thymidine. (A) Autoradiograph. [Reprinted, with permission, from J. Cairns, *Cold Spring Harbor Symp Quant. Biol.* 28:43–46 (1963)]. (B) Interpretative drawing. A duplex, formed in the previous generation, contains one labeled and one unlabeled strand and undergoes bidirectional replication. The labeled parental strand already replicated contains two labeled strands and has twice the grain density (number of dark spots/unit distance) of the other sections.

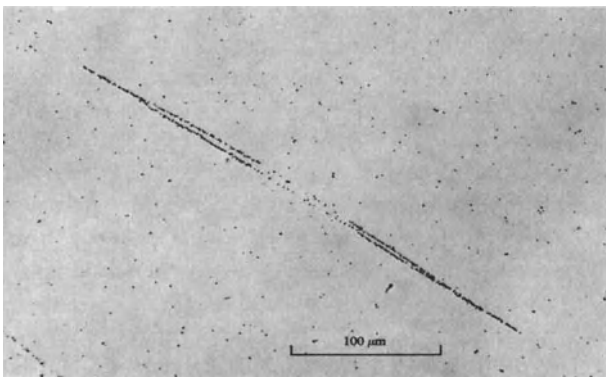


Figure 17.12. Autoradiograph showing bidirectional replication of *E.coli* DNA. [Reprinted, with permission, from D. M. Prescott and P. L. Kuempel, *Methods Cell Biol.* 7:147–156 (1973).]

17.4. MECHANISM OF DNA REPLICATION

17.4.1. Unwinding of the Double Helix

The double helix begins to unwind when dnaA protein binds to three A/T-rich, tandemly repeated segments (13 bp long) in the *ori*C site. At least three other classes of proteins then join dnaA protein in the unwinding process: helicases, single-strand binding proteins, and topoisomerases.

Helicase (*dnaB protein*) is an ATPase that catalyzes the breakage of hydrogen bonds linking two complementary bases. The enzyme binds *ahead* of the replication fork and catalyzes an ATP-dependent unwinding of the duplex. Two molecules of ATP are required for every base pair broken.

Following the action of helicase, **single-strand binding proteins (SSB)** bind to the two separated strands and prevent them from reassociating. SSB form tetramers (MW = 19,000/subunit) that exhibit strong cooperativity and that bind to a single strand regardless of its base composition or base sequence. Many copies of SSB bind to DNA in a highly cooperative fashion. SSB do not bind to double-stranded DNA.

Unwinding of the two strands at the replication fork generates a twist that is relieved by *topoisomerase II* or **DNA gyrase.** This enzyme binds just *ahead* of the replication fork and functions as a swivel so that only a short segment of the strand rotates. The enzyme catalyzes an ATP-dependent conversion of positively supercoiled DNA to relaxed DNA, followed by ATP-independent conversion of relaxed to negatively supercoiled DNA (Figure 17.13).

The torsional problem inherent in unwinding the helix is formidable. You might consider the process by visualizing a two-stranded rope unwinding. In the absence of any swivel, or point of attachment, as one strand becomes unwound, the other would flail around wildly.

For a circular chromosome, such rotation would lead to rapid accumulation of positive supercoils in DNA. In *E. coli,* for example, a replication fork result in incorporation of about 120,000 nucleotides (60,000 bp) per minute. Because there are 10 bp per turn of helix (see Table 7.6), unwinding the DNA would produce positive supercoils at the rate of about 6000 per minute. By comparison, note that a speed of 6000 rpm exceeds that of an engine in a car traveling at 70 mph! At this rate, the DNA would become so tightly *overwound* in a very short time that any further unwinding would be impossible, and replication would come to a halt. The tremendous torsional strain inherent in DNA replication is prevented by action of DNA gyrase, which relieves developing tension by inducing negative supercoiling in the DNA.

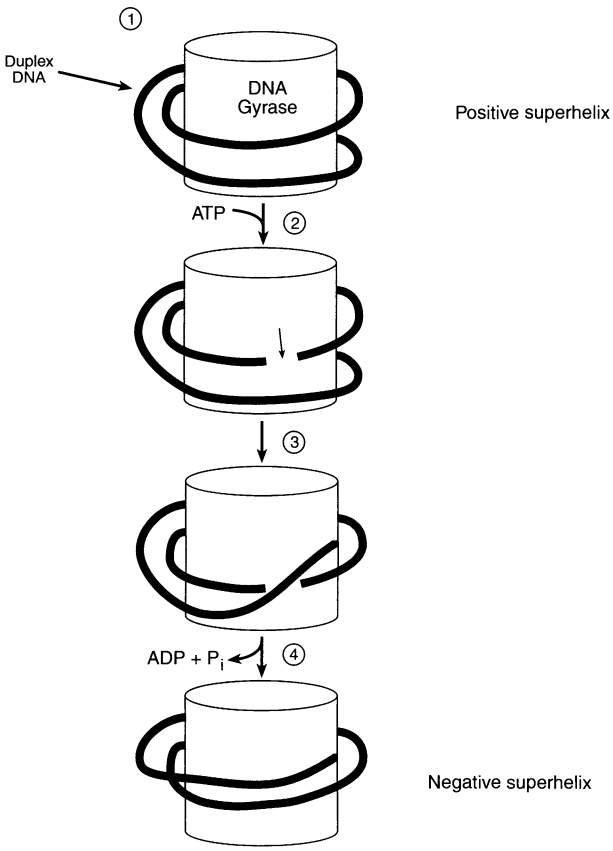


Figure 17.13. Mechanism of action of DNA gyrase: (1) DNA coils around the enzyme; (2) gyrase catalyzes an ATP-dependent double-strand scission; (3) unbroken DNA passes through the broken duplex; (4) DNA ligase catalyzes sealing of the cut strands.

17.4.2. Discontinuous DNA Synthesis

Synthesis of new DNA along the two parental template strands differs in a number of ways. We have already stressed the polarity aspect, which demands that synthesis must proceed in opposite directions in order to comply with the polymerase specificity and provide for growth in the 5′ → 3′ direction.

Now consider what happens as the replication fork moves, opening up a new section of DNA for replication. Synthesis along one template strand can proceed continuously by simply adding a deoxyribonucleotide to the 3′-OH formed in the preceding step. We call the strand synthesized in this fashion the **leading strand** (Figure 17.14). This new strand, once it has been started, *becomes extended continuously and grows in the direction of fork movement.*

The second template strand, however, requires *discontinuous synthesis.* After synthesizing a given DNA fragment, the polymerase must start all over again by beginning synthesis of the next fragment at a different point, corresponding to a later position of the replication fork. We call the strand synthesized in this fashion the **lagging strand.** Synthesis of the lagging strand requires multiple starts as the replication fork moves and results in production of multiple short DNA fragments that must be linked later to form one continuous strand. Thus, growth of the lagging strand is *discontinuous, and each fragment grows in a direction opposite to that of fork movement.*

Reiji Okazaki has provided experimental support for this replication model. He showed that a significant fraction of newly synthesized DNA is present in the form of short fragments during the early stages of replication. As time goes on, the average length of newly synthesized DNA increases, because shorter fragments become linked to form longer ones. In honor of their discoverer, we call these short DNA fragments **Okazaki fragments.** They are synthesized by DNA polymerase III and are typically 1000–2000 nucleotides long in prokaryotes but only 100–200 nucleotides long in eukaryotes. Because synthesis of the lagging strand is discontinuous, we refer to the mechanism of DNA replication as **discontinuous** (more precisely, *semidiscontinuous*) **replication.**

Note that DNA polymerase can catalyze extension of an Okazaki fragment right up to the 5′-end of a preceding fragment, but it cannot catalyze the joining together of two Okazaki fragments. Linking two DNA fragments requires the action of a different enzyme, **DNA ligase,**

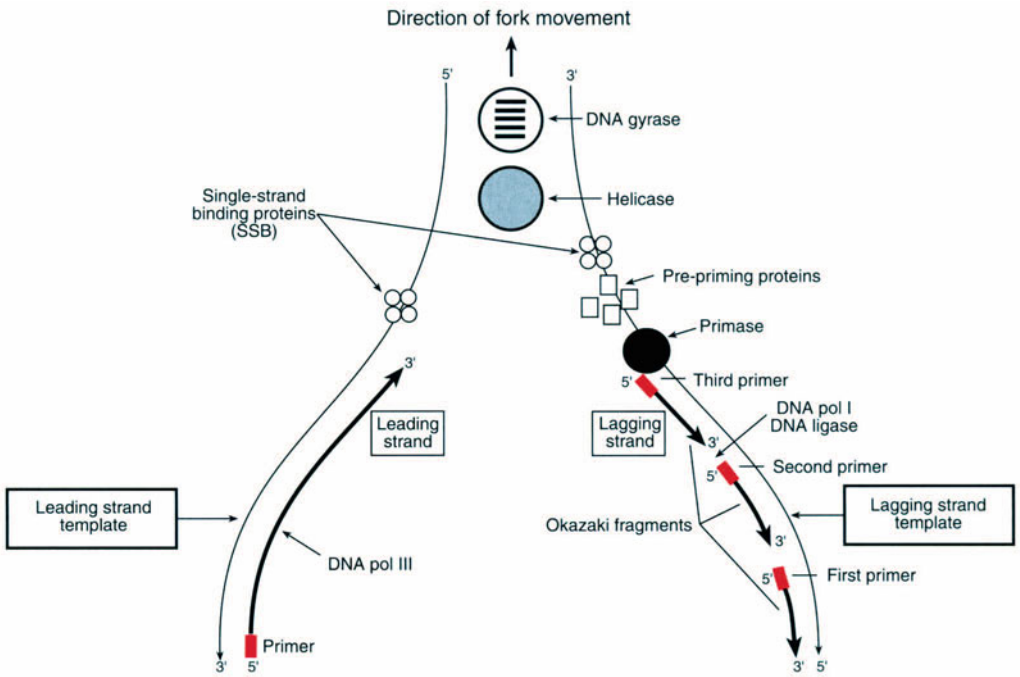


Figure 17.14. Schematic illustration of events at the replication fork in discontinuous DNA replication.

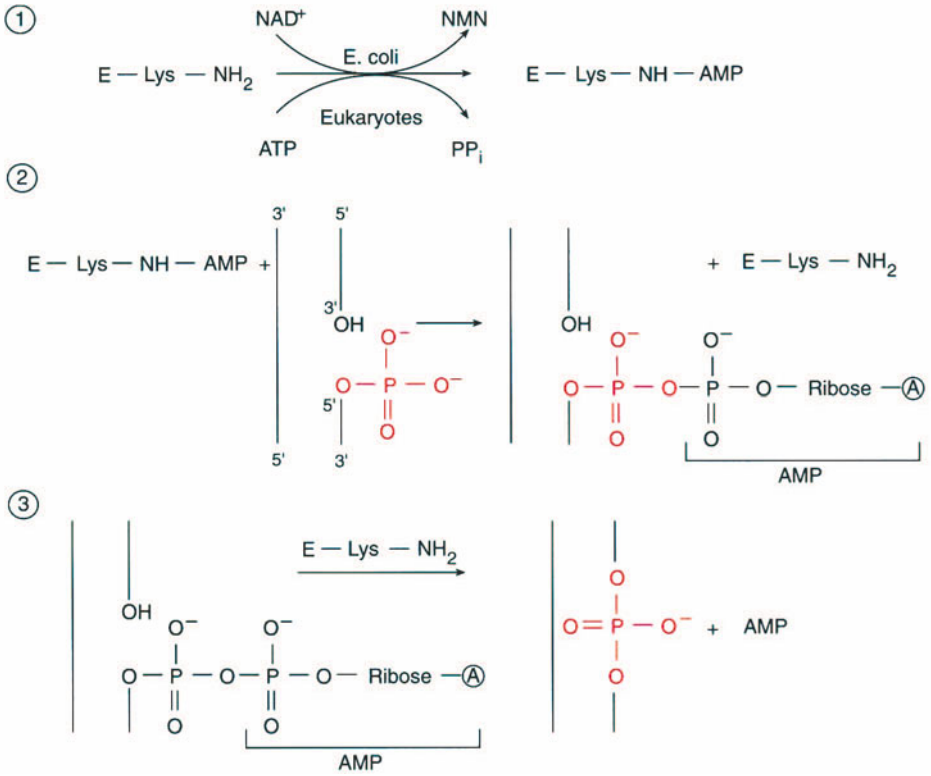


Figure 17.15. Mechanism of action of DNA ligase. (1) An AMP moiety, derived from either NAD^+ or ATP, becomes linked to the ϵ-NH_2 group of a lysine residue in the enzyme (E). (2) The AMP residue is transferred to the 5′-end of a DNA strand. (3) Nucleophilic attack by a 3′-OH of a second strand on the activated 5′-phosphorus atom forms a 3′,5′-phosphodiester bond between the two strands.

which catalyzes the *linking of two single-stranded ends of a duplex:*

NO ↓	YES ↓
ssDNA (linear or circular)	dsDNA (linear or circular)

The DNA ligase reaction requires formation of an *enzyme adenylate* as *common intermediate* of the first two steps in the mechanism (Figure 17.15). The third step involves a nucleophilic attack by the 3′-OH group on the activated 5′-phosphorus atom. The enzyme adenylate consists of AMP linked via a nitrogen–phosphorus bond to the ϵ-NH_2 group of lysine in the enzyme. In prokaryotes, NAD^+ provides the AMP; in eukaryotes and T4 phage, ATP supplies the AMP. The DNA ligase reaction constitutes an endergonic reaction, driven by expenditure of two energy-rich bonds. One bond is spent during formation of the enzyme adenylate by cleavage of either NAD^+ or ATP. The second energy-rich bond is spent either in regenerating NAD^+ from NMN and ATP or in pyrophosphatase-catalyzed hydrolysis of PP_i to 2 P_i.

17.4.3. RNA Primers

In addition to direction of synthesis and continuous versus discontinuous synthesis, reactions along the leading and lagging strands differ in yet another important aspect.

Because synthesis of new DNA is catalyzed by DNA polymerase, a primer having a free 3′-OH group is required for initiation of polymerization. In characterizing the nature of this primer, a surprising discovery was made: short *RNA* fragments serve as primers for synthesis of *DNA!* Synthesis of the leading strand requires only one such **RNA primer.** Once the initial small RNA fragment has been provided, synthesis of the strand is continuous. Synthesis of the lagging strand requires multiple primers, one for each DNA fragment made. Every time DNA polymerase produces another Okazaki fragment, a new RNA primer must be synthesized first.

RNA primers are produced by a *DNA-dependent RNA polymerase* called **primase.** Primase consists of a single polypeptide chain (MW = 60,000), and there are about 50 enzyme molecules per *E.coli* cell. Like the RNA polymerase that functions in transcription, primase is a

self-priming enzyme; it does not require a primer. Primase produces an RNA strand from ribonucleoside triphosphates. The RNA strand grows in the 5′ → 3′ direction and is complementary and antiparallel to the DNA template strand. Since RNA contains the pyrimidine uracil (as opposed to thymine in DNA), an adenine in the template DNA leads to uracil in the RNA.

RNA primers synthesized by means of primase constitute short fragments of about 5–10 nucleotides. Once an RNA primer has been synthesized, DNA polymerase III extends it by addition of deoxyribonucleotides to its 3′-end. The 5′ → 3′ exonuclease activity of DNA polymerase I then removes the RNA primers, and the polymerase activity of the enzyme fills the resulting gaps with sections of DNA. Lastly, DNA ligase catalyzes the linking of all DNA fragments.

Primase is active only in the presence of proteins called **pre-priming proteins.** Accordingly, binding of pre-priming proteins to the replication fork precedes the binding of primase. These proteins appear to bind specifically to the lagging strand template, where the presence of primase is required repeatedly. The complex of pre-priming proteins and primase has been dubbed **primosome.** Researchers believe that the leading strand is primed through the combined action of primase and the RNA polymerase that mediates transcription. The role of RNA primers in DNA replication supports the theory that the evolution of RNA preceded that of DNA (see Section 1.1).

17.4.4. Sequence of Events at the Replication Fork

DNA replication has been studied most extensively in *E. coli,* but even for this system (Table 17.4) our understanding of replication is not yet complete. We refer to the array of enzymes and other proteins that bring about replication as a **replisome.** It consists of the primosome plus SSB, DNA polymerases I and III, DNA ligase, helicase, and DNA gyrase. A review of Figure 17.14 suggests that we can describe DNA replication by the following steps, not necessarily in precise chronological order:

1. Replication begins when dnaA protein binds at the replication origin, *ori*C, leading to an unwinding of three tandemly repeated A/T-rich segments.
2. Helicase binds ahead of the replication fork and produces a local unfolding of the double helix.
3. Single-strand binding proteins (SSB) bind to both template strands and prevent them from reassociating.
4. Pre-priming proteins bind to the lagging strand template and ensure the access of primase.
5. Primase binds to pre-priming proteins on the lagging strand template, thereby forming the primosome. The primosome catalyzes synthesis of an RNA primer, complementary and antiparallel to the DNA template.
6. Combined action of primase and the RNA polymerase that mediates transcription leads to synthesis of a primer for replicating the leading strand template.
7. Two replisomes assemble at the *ori*C and proceed to replicate the DNA bidirectionally. A dimer of DNA polymerase III holoenzyme binds to the primosome. One monomer of DNA polymerase III catalyzes replication of the lagging strand, and the other catalyzes replication of the leading strand. Investigators have proposed that this occurs by looping the lagging strand template so that both parental strands pass through the polymerase sites in the same direction (Figure 17.16).
8. DNA polymerase III catalyzes extension of both RNA primers. Growth is in the 5′ → 3′ direction. Errors the enzyme makes are corrected by means of its associated proofreading function (3′ → 5′ exonuclease). DNA polymerase III represents the main synthetic enzyme of DNA replication.

Table 17.4. Protein Components of the DNA-Replicating System in *E. coli*

Protein	Function	Size (MW)	Molecules/cell
dnaA protein	Starts unwinding of duplex	52,000	20
Pre-priming proteins	Facilitate primase access	25,000–50,000	?
Single-strand binding proteins (SSB; tetramer)	Prevent strand annealing	74,000	300
Primase	Synthesizes RNA primers	60,000	50
Helicase (hexamer)	Unwinds the double helix	300,000	20
DNA pol I	Removes RNA primers and fills gaps	103,000	300
DNA pol III (dimer)	Synthesizes DNA	~900,000	20
DNA gyrase	Relieves unwinding tension	400,000	250
DNA ligase	Joins DNA fragments	74,000	300

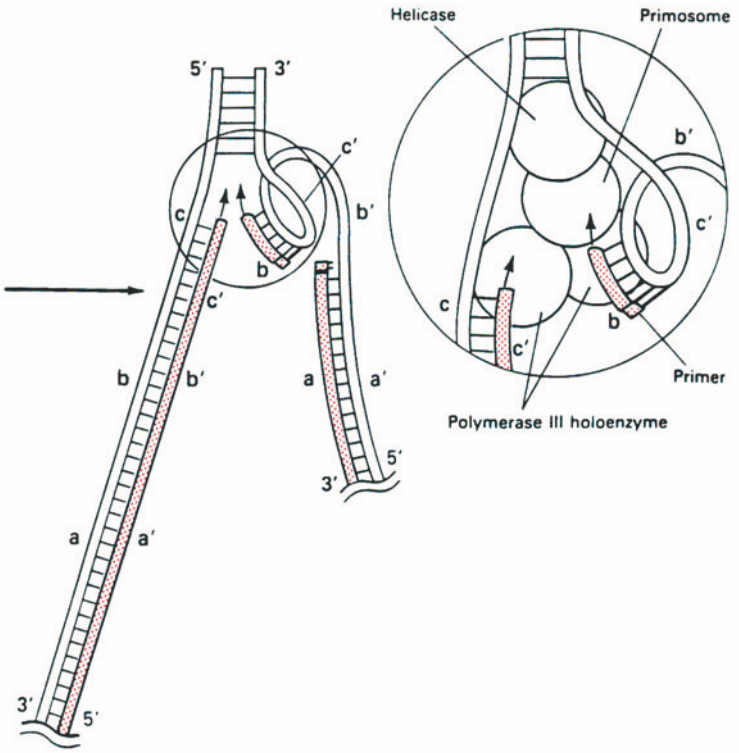


Figure 17.16. Model showing a proposed looping of the lagging strand template at the replication fork. Looping enables the two subunits of dimeric DNA polymerase III to move along both DNA strands in the same direction, one subunit along the lagging strand, and the other subunit along the leading strand. (Courtesy of Dr. A. Kornberg.)

9. Once primed, *the leading strand template can be copied continuously* through action of the dimeric DNA polymerase III.

10. Along the lagging strand template, the primosome starts repeatedly at different points corresponding to later positions of the replication fork. Each new RNA primer synthesized becomes extended by a DNA fragment in a reaction catalyzed by DNA polymerase III. Each RNA primer, with its Okazaki DNA fragment attached, is H-bonded to the lagging strand template. *The lagging strand template is copied discontinuously.*

11. DNA polymerase I catalyzes an excision of the RNA primer, attached to the preceding Okazaki fragment, by means of its associated $5' \rightarrow 3'$ exonuclease activity. As the *ribonucleotides* of the RNA primer are removed by action of the exonuclease, they are replaced with *deoxyribonucleotides* by action of the polymerase activity of DNA polymerase I. As in nick translation, polymerase activity and exonuclease activity of DNA polymerase I function at comparable rates. In this fashion, DNA polymerase I catalyzes an extension of the new Okazaki fragment and a narrowing of the gap between this and the preceding fragment. As DNA polymerase I displaces DNA polymerase III, the latter initiates synthesis of the next Okazaki fragment.

12. DNA ligase catalyzes the joining of the two adjacent Okazaki fragments.

13. DNA gyrase binds ahead of the replication fork and relieves the strain produced by the unwinding DNA.

14. Replication is terminated at a specific termination site and involves a protein (MW = 36,000), called **terminator utilization substance (tus),** that binds at this site. Tus inhibits the helicase of the replisome and thereby prevents the replication forks from passing through the termination site. In bidirectional replication of a circular chromosome, the termination site is located where the two replication forks approach each other, opposite the replication origin. The termination site also contains specific DNA sequences that function in separating the daughter DNA molecules.

15. Following termination of replication, DNA becomes subject to **postreplicative processing,** which consists primarily of two groups of reactions. One group involves methylations of bases, catalyzed by *modification methylases* and designed to protect DNA against digestion by *restriction endonucleases* (Section 7.6). The other group of reactions includes a variety of DNA repair mechanisms (see below).

There exists an apparent anomaly in prokaryotic replication; it takes less time for the cell to divide than it

takes for its DNA to replicate. Cells of *E. coli,* for example, can double every 20 minutes but require some 40 minutes for their DNA replication. How can a cell divide when its DNA has not yet been fully replicated? The answer is that in rapidly dividing cells, a replicating DNA molecule can initiate a second round of replication from the origin before the first replication has been completed. Because initiation of replication occurs more frequently than once every 40 minutes, the effective replication time of the chromosome decreases to 20 minutes. How this is regulated remains to be unraveled.

17.4.5. Replication in Eukaryotes

DNA replication in eukaryotes involves many of the same factors and principles discussed for prokaryotes. Semiconservative replication, leading and lagging strands, Okazaki fragments, DNA polymerases, DNA ligase, DNA unwinding, and so on are part of eukaryotic replicative systems. However, there also exist many differences between the two types of systems. Replication in eukaryotes is even more complex than in prokaryotes. The eukaryotic genome is much larger than the bacterial chromosome and contains numerous replication origins. For example, the largest chromosome of the fruit fly (*Drosophila melanogaster*) consists of 62,000 kbp and contains some 3000 replication origins, resulting in 6000 replication forks. A mammalian cell having 46 chromosomes and 10^3–10^4 replication origins must have 20–200 origins per chromosome. Because of the large number of replication origins, eukaryotic chromosomes are replicated rapidly even though the rate of movement of each replication fork is only about ½₀ the rate in prokaryotes. If the *Drosophila* chromosome contained only one replication origin, its replication would take more than 16 days. The actual time required for replication, accomplished by 3000 replication origins, is less than three minutes.

Differences between DNA replication in eukaryotes and prokaryotes also include the form in which DNA occurs in the cell. Whereas bacterial replication has ready access to "naked" DNA, eukaryotic systems must deal with DNA packaged with histones in the form of nucleosomes (Section 7.5). Researchers believe that the binding of histones to DNA is responsible, at least in part, for the slower rate of movement of eukaryotic replication forks.

The occurrence of DNA in nucleosomes also requires synthesis of histones as DNA replication proceeds. It appears that histones of the parental DNA remain bound as the DNA replicates, and newly synthesized histones bind to the daughter strands as they form. The number of histones required for nucleosome construction doubles with each round of DNA replication. Extensive histone synthesis is made possible by multiple copies of histone genes. The number of gene copies varies and appears not to be related to the size of the genome. Birds and mammals have 10–20 copies of each of the five histone genes; the fruit fly has about 100; and the sea urchin has several hundred. In most organisms, histone genes are the only protein genes that occur as multiple copies. Histone genes also differ from other protein genes in that almost all lack introns.

Regulation of histone synthesis occurs not only at the level of transcription but also at the level of mRNA degradation. Histone mRNAs have relatively short cytoplasmic half-lives. Consequently, when DNA synthesis stops, histone mRNAs are rapidly degraded. Both DNA replication and histone synthesis occur during the synthesis stage (S phase) of the cell cycle—the sequence of events from the beginning of one cell division to the beginning of the next.

17.5. INTEGRITY OF DNA STRUCTURE

All cells possess elaborate machinery to eliminate damage to DNA and restore the integrity of the double helix. This machinery operates both during and after replication. DNA can be damaged by means of *replication errors* introduced during replication and by means of *mutations* introduced after the DNA has been synthesized.

17.5.1. Fidelity of DNA Replication

DNA replication constitutes the most accurate of all known information transfer processes. High fidelity of replication is necessary so that an organism's genetic information can be transmitted accurately from generation to generation without producing changes in the species. Because DNA is synthesized only once during each cell cycle—as opposed to RNA and proteins, which are synthesized repeatedly—it is particularly important that DNA replication be carried out with great fidelity.

Despite these requirements, a wrong base may occasionally become incorporated into the growing DNA strand. Such **misincorporation** leads to formation of nonstandard base pairs. Researchers have estimated that misincorporation in *E. coli,* in the absence of the proofreading function, would amount to about one base per 10^4–10^5 bp incorporated. They have also estimated that by means of the $3' \rightarrow 5'$ exonuclease activity and postreplicative repair mechanisms the actual level of misincorporation decreases to about one base per 10^9 or 10^{10} bp.

Misincorporation of one base per 10^{10} bp represents

a very high degree of fidelity, an almost perfect copying process of the genetic material. You can get a feel for this number by recalling that the *E.coli* genome contains about 10^7 nucleotides (see Table 7.8). Hence, an error of 1 in 10^{10} corresponds to an erroneous incorporation of one base per about 1000 bacterial cells per generation. DNA replication in eukaryotes also has a high degree of accuracy; only one base in 10^9–10^{11} is incorrectly incorporated.

Three factors account for the high fidelity of DNA replication in both prokaryotes and eukaryotes: exonuclease activity, replisome complexity, and RNA primers.

The proofreading function of $3' \rightarrow 5'$ exonuclease of DNA polymerase represents the major factor responsible for the fidelity of DNA replication. The enzyme serves to excise wrong bases erroneously incorporated by the polymerase. Significantly, both DNA polymerase I and III possess this proofreading function.

A second contributing factor arises from the complexity of the replisome, an aggregate that consists of a large number of different enzymes and proteins. The various components are interdependent in their properties and activities. The complexity of this network helps to ensure that the overall process is carried out with great accuracy.

Lastly, fidelity of DNA replication increases because RNA primers, rather than DNA primers, serve in the process. We think that errors in base incorporation are most likely to occur at the very early stages of replication, when the first few nucleotides become incorporated. That these first few nucleotides are ribonucleotides, rather than deoxyribonucleotides, is therefore beneficial. Any errors that occur at this stage will be eliminated as the ribonucleotides are excised and replaced with deoxyribonucleotides by the $5' \rightarrow 3'$ exonuclease and polymerase activities of DNA polymerase I.

17.5.2. Mutations

Mutations of DNA can occur both spontaneously and as a result of exposure to chemical or physical *mutagens.* Biochemists have estimated that the rate of spontaneous mutations (number of mutations per round of DNA replication) is about 10^{-7} in T4 phage, 10^{-10} in *E. coli,* and 10^{-5} in a human cell. Mutations involving a single nucleotide (base) are *point mutations* and can be of four types: transitions, transversions, deletions, and insertions (Figure 17.17).

Transitions involve the change of one purine to another or the change of one pyrimidine to another. *Transversions* involve the change of a purine to a pyrimidine, or vice versa. Both transitions and transversions result in formation of a new base pair. **Deletions** and **insertions** are mutations involving either the removal or the additon of a base pair. Because deletions and insertions lead to changes in the reading frame, we call them **frameshift mutations.**

Physical agents that bring about mutations in DNA include X rays, which cause **nicks** or breaks in a single strand, and ultraviolet radiation, which leads to formation of pyrimidine dimers. A **pyrimidine dimer,** such as a *thymine dimer,* forms when two consecutive pyrimidines in one strand become covalently linked to each other (Figure 17.18).

Various chemicals (Figure 17.19) interact with DNA and bring about mutations. Many of these **mutagens** are

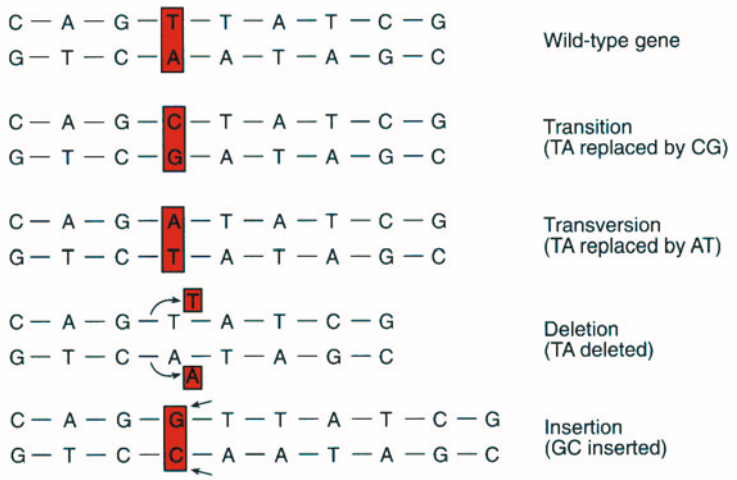


Figure 17.17. Different types of point mutations.

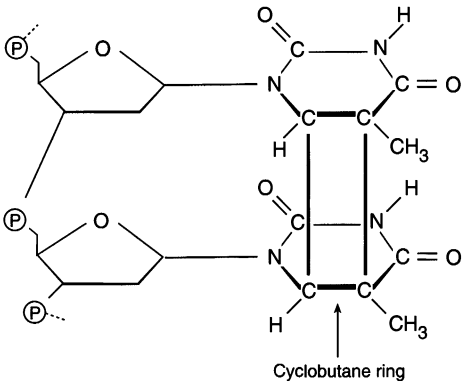


Figure 17.18. A thymine dimer produced by ultraviolet irradiation of DNA. Two adjacent thymines in one strand become linked covalently.

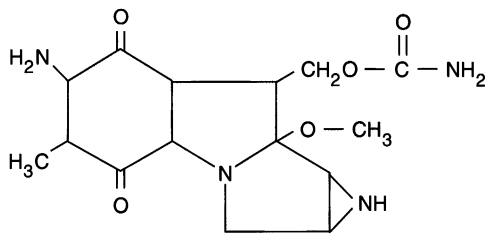
also *carcinogens,* agents capable of producing cancer. Chemical mutagenesis may result from modification of bases (for example, oxidative deamination by nitrous acid or alkylation by alkylating agents), *cross-linking* of DNA strands (for instance, by mitomycin C, an antibiotic), or alteration of the helix by *intercalating agents,* planar molecules that become inserted between adjacent base pairs and lead to insertion or deletion of one or more base pairs (for example, ethidium bromide and acridine dyes).

17.5.3. DNA Repair Mechanisms

Even in relatively simple organisms, such as *E. coli,* there exist several mechanisms for repairing damaged DNA. Major DNA repair processes are similar in both prokaryotes and eukaryotes and may be divided into five general categories: direct reversal, excision, mismatch repair, recombination repair, and SOS repair.

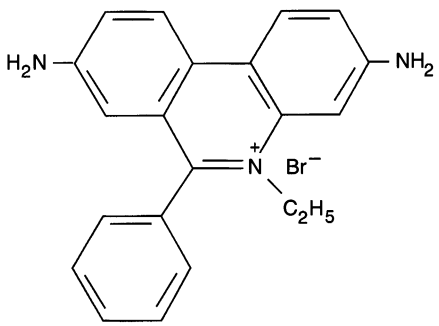
17.5.3A. Direct Reversal of Damaging Reaction. This mechanism of DNA repair consists of a direct reversal of the reaction that led to the damage in DNA. Some alkylations, for instance, can be reversed by means of a specific enzyme that catalyzes transfer of the alkyl group from a base in DNA to an amino acid residue in the enzyme. Repair by means of *photolyase* also constitutes a direct reversal of DNA damage. Photolyase is activated by absorption of visible light (300–500 nm) and catalyzes the cleavage of the C—C bonds that link two adjacent pyrimidines in a pyrimidine dimer, thereby restoring them to their normal structures.

17.5.3B. Excision Repair. Excision repair employs some of the enzymes of DNA replication. The mechanism involves removing a damaged DNA segment and replacing it with the normal base sequence. Damage resulting

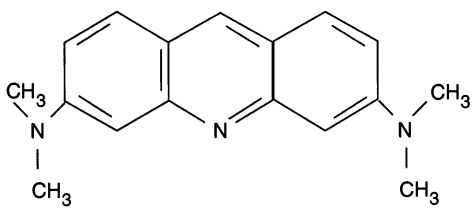


Mitomycin c

Nitrous acid



Ethidium bromide



Acridine orange

Figure 17.19. Structures of some chemical mutagens.

from formation of pyrimidine dimers can be repaired by excising the dimer and filling the gap with normal bases. A nuclease catalyzes the excision, and DNA polymerase I fills the resulting gap. DNA ligase then catalyzes the joining of the original and newly synthesized strands. Figure 17.20 illustrates the mechanism as it occurs in *E. coli.*

A different excision repair mechanism allows for removal and replacement of an incorrect or altered base. An incorrect base may be uracil or hypoxanthine, formed by spontaneous deamination of cytosine or adenine. An altered base may be one that has accidentally been methylated or changed in some other way. Cells contain **DNA glycosylases** that catalyze cleavage of the glycosidic bond of an incorrect or altered base. Action of the enzyme removes the base and produces a site at which only the deoxyribose-phosphate moiety remains. Because such sites lack either a purine or a pyrimidine, we call them **apurinic** or **apyrimidinic sites (AP sites).**

AP sites are "recognized" by **AP endonuclease,** a specific nuclease that catalyzes the cleavage of the phosphodiester bond on the 5′-side of the AP site. Deoxyribose (at times together with a few nucleotides) is subsequently removed by an exonuclease, either the 5′ → 3′ exonuclease of DNA polymerase I or some other cellular exonuclease. Action of DNA polymerase I results in a filling of the gap, and the strands become linked by means of DNA ligase (Figure 17.21).

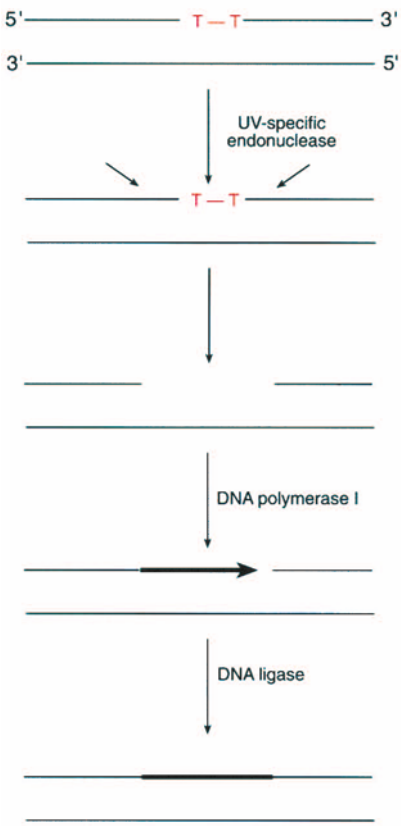


Figure 17.20. Excision repair of a thymine dimer (T—T) in *E. coli* DNA. Endonuclease-catalyzed removal of the dimer, together with eight nucleotides on its 5′-side and four or five nucleotides on its 3′-side, is followed by polymerase and ligase action.

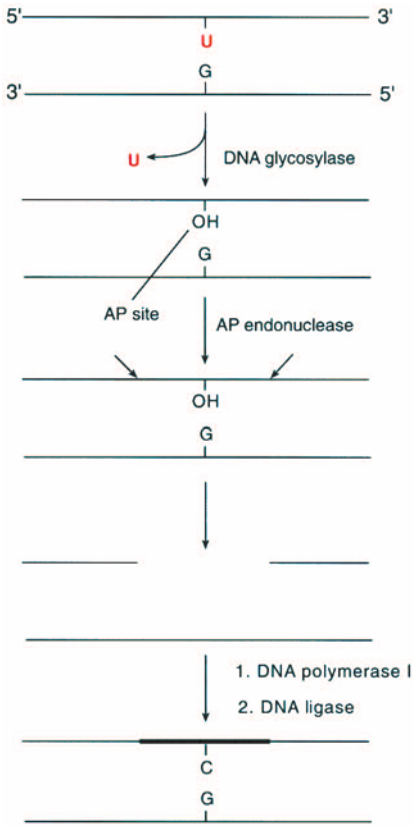


Figure 17.21. Excision repair by means of DNA glycosylase.

Thymine dimers can also be repaired by means of a DNA glycosylase mechanism. This occurs in T4 phage and *Micrococcus luteus.* A glycosylase catalyzes cleavage of the glycosidic bond at the 5′-end of the thymine dimer and forms an apyrimidinic site. The AP site is recognized by AP endonuclease, excised, and repaired.

Defective DNA repair mechanisms can lead to severe pathological conditions. A rare hereditary disease in humans, called *xeroderma pigmentosum,* is believed to be due (at least in some of the afflicted individuals) to a deficiency of the endonuclease that introduces nicks into damaged sections of DNA, such as those containing thymine dimers. Patients afflicted with the disease exhibit abnormal sensitivity to sunlight. Thymine dimers, produced by ultraviolet light in exposed skin areas (arms and face), are not repaired and result in continuous death of

skin cells. Sufferers develop marked skin changes in infancy, such as dryness and freckling, and these worsen with time. Eventually, they develop skin cancers that are often fatal.

17.5.3C. Mismatch Repair. When DNA polymerase catalyzes the incorporation of an incorrect base during replication, a non-Watson–Crick base pair forms. If the proofreading functions of DNA polymerase I and III do not correct such errors, they can be eliminated by **mismatch repair,** a mechanism based on the occurrence of methylated bases in DNA (Figure 17.22). During replication, parental DNA is fully methylated and contains these modified bases wherever called for. However, newly synthesized daughter DNA strands remain undermethylated relative to the parental DNA for a brief period of time, because DNA methylation lags behind DNA synthesis.

The mismatch repair system has the capacity to "recognize" both specific unmethylated sequences and mismatched base pairs in the newly synthesized daughter strand. Recognition of unmethylated sequences serves to target that strand, rather than the parental strand, for repair. The repair mechanism involves excising a segment of the daughter strand that is adjacent to the unmethylated sequence and includes the mismatched base. The excised segment is then replaced by a new segment containing the correct base.

In *E. coli,* the mismatch repair apparatus consists of a multienzyme system that includes helicase II, single-strand binding protein, and several other proteins. The DNA site that becomes methylated consists of the sequence —GATC— in which adenine occurs as N^6-methyladenine (Figure 7.2).

17.5.3D. Recombination Repair. At times, damaged DNA may undergo replication before the spe-

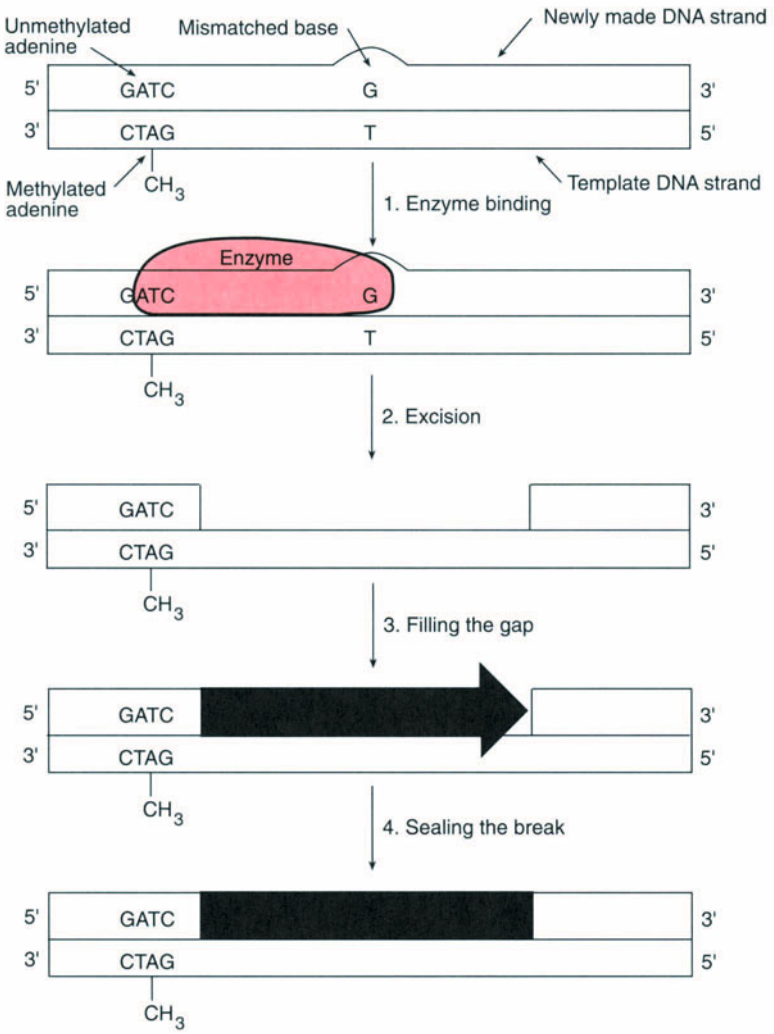


Figure 17.22. Mismatch repair. Enzyme "recognition" of an unmethylated sequence and a mismatched base in a newly synthesized strand leads to excision of a section containing the mismatched base. The resulting gap is filled and sealed by means of DNA polymerase I and DNA ligase.

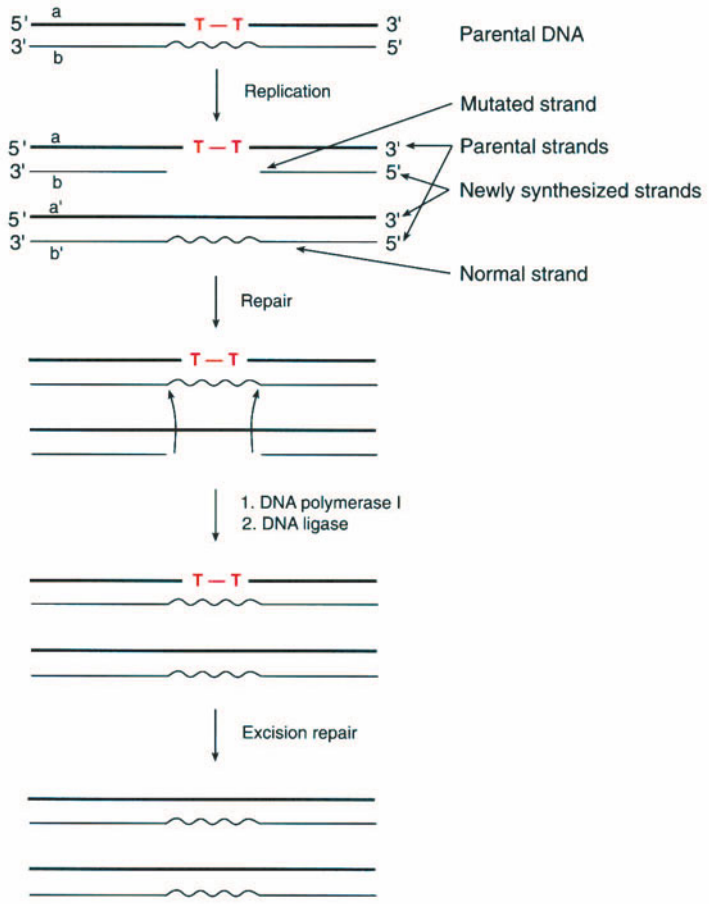


Figure 17.23. Recombination repair of DNA, illustrated by replication of a thymine dimer-containing molecule. Before the dimer can be eliminated by excision repair, it must face a complementary template rather than a gap. The missing template section is transferred into the gap from a normal sister strand formed during replication.

cific lesion has been repaired by the mechanisms described above. In that case, replication of the distorted strand becomes interrupted at the lesion and resumes at some point past it so that the newly synthesized daughter strand has a gap in its base sequence. This defect can be corrected by **recombination repair,** a mechanism that resembles genetic recombination (see below). In recombination repair, two daughter molecules interact (Figure 17.23). The gap in one strand of one daughter molecule becomes filled by insertion of the corresponding segment from the normal sister strand of the second daughter molecule. Because the resulting gap in the normal sister strand lies opposite an undamaged strand, it can be repaired by means of DNA polymerase and DNA ligase.

17.5.3E. SOS Repair. A complex set of cellular changes, termed the **SOS response,** is induced in bacteria subjected to potentially lethal mutagens such as ultraviolet irradiation, alkylating agents, or cross-linking agents. As part of the SOS response, cells cease to divide and increase their capacity to repair damaged DNA. SOS repair constitutes an *error-prone repair.* Damaged DNA strands form even though they are mutagenic. Survival of the cell with mutations is deemed preferable to cell death. Repair of damaged DNA is achieved through the coordinated induction of several enzymes.

17.5.4. Genetic Recombination

In addition to deleterious alterations of DNA resulting from replication errors and mutations, chromosomes undergo alterations that are part of their normal biological functions. These alterations include genetic recombination and transposition.

Genetic recombination leads to production of progeny that derives some of its genes from one parent and some from another, genetically different parent. In prokaryotes, genetic recombination occurs via three mechanisms—transformation, conjugation, and transduction. In **transformation,** a DNA fragment becomes incorporated into the chromosome of a recipient bacterial cell. Transformation of R-type *Pneumococcus* cells to S-type cells (see Section 7.5) is an example. **Conjugation** occurs in bacteria and other unicellular organisms. It resembles sexual reproduction and

involves DNA transfer between cells of opposite mating types, associated side by side. **Transduction** involves transfer of a DNA segment from one bacterial cell to another via a bacteriophage (phage). The DNA segment first becomes incorporated into the phage DNA and from there is transferred to the DNA of the recipient cell.

In eukaryotes, genetic recombination occurs via **crossing over,** a process whereby genetic material is exchanged between homologous chromosomes during meiosis. *Homologous chromosomes* are two chromosomes, one derived from the male and one derived from the female parent, that contain the same linear gene sequence so that each gene is present in duplicate. In 1964, Robin Holliday proposed a general model for crossing over between homologous segments of two duplex DNA molecules (Figure 17.24).

17.5.5. Transposition

Transposition involves physical movement of DNA segments from one chromosomal locus to another. The notion of mobile genetic elements was first advanced in the early 1950s by Barbara McClintock (Nobel Prize, 1983), based on her genetic analysis of the pigmentation pattern of maize (Indian corn). Her proposal went counter to the then accepted principles of genetics and was met with widespread rejection. Some 20 years passed before additional evidence emerged from studies of *E. coli,* supporting the same "heretical" notion. Today the principle of mobile genetic elements, or **transposable elements,** has become well established. Genes, or groups of genes, can move from one locus to another on the same chromosome, or even to a locus on a different

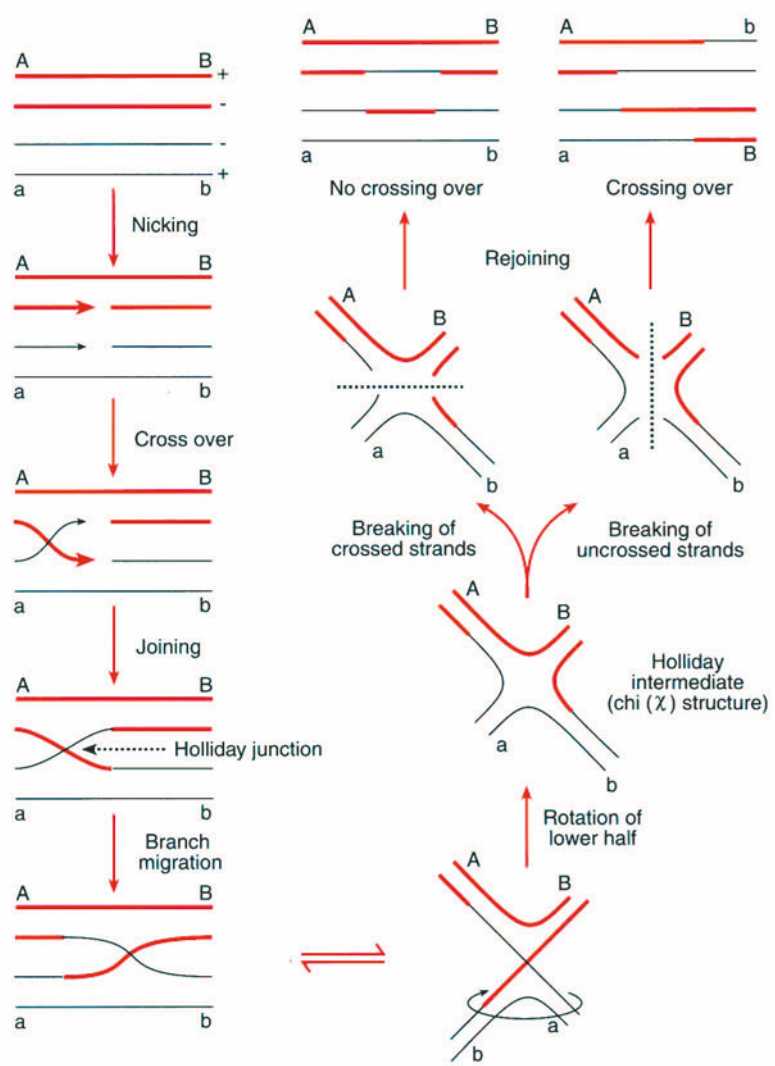


Figure 17.24. The Holliday model of general recombination. Letters denote two different genes located on homologous chromosomes. After crossing over, nicked strands are joined by action of DNA ligase. Branch migration is based on unwinding and rewinding of the double helix. The Holliday intermediate can lead to two sets of duplex molecules by reactions that require nicking and sealing of the strands. The reactions have equal probability but differ in the locations of the nicks.

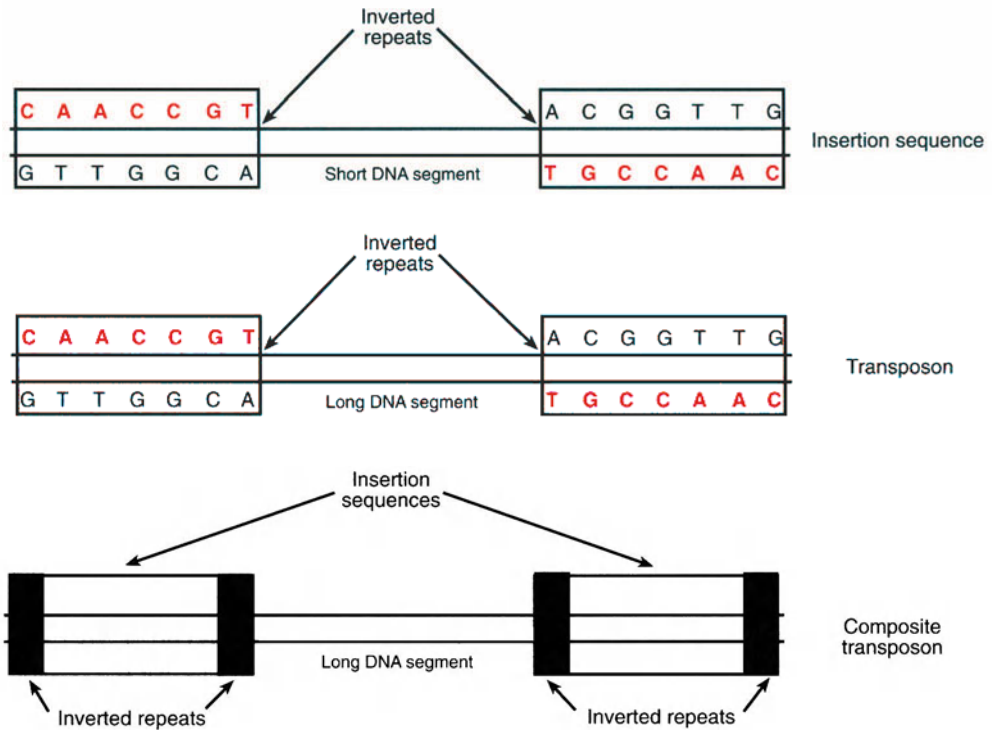


Figure 17.25. Major classes of transposable elements.

chromosome. The phenomenon occurs in both prokaryotes and eukaryotes.

Transposable elements can be divided into three classes, insertion sequences (IS), transposons, and composite transposons (Figure 17.25). An **insertion sequence (IS)** represents the simplest transposable element and consists of a small double-stranded DNA segment that generally contains fewer than 2000 bp. The sequence contains a gene that codes for *transposase,* an enzyme essential for transposition, and occasionally a regulatory gene related to the process of transposition, but no other genes. An insertion sequence is flanked by *inverted repeats.*

Transposons and **composite transposons** are larger transposable elements that contain not only the gene for transposase but also a number of other genes not related to the process of transposition. Transposons are flanked by inverted repeats, but composite transposons are flanked by insertion sequences.

Transposition promotes changes in DNA such as those involving insertions and deletions. It also plays an important role in the capacity of bacteria to mutate at a rapid rate. Lastly, transposition accounts for the development of bacterial resistance to antibiotics, since transposons and composite transposons frequently carry genes that confer resistance to one or more antibiotics. Following the widespread use of antibiotics, such transposons have accumulated in bacterial plasmids by transposition. When a previously susceptible bacterium acquires plasmids of this type, the bacterium suddenly develops resistance to the antibiotic(s). It is of great medical concern that many pathogenic bacteria are becoming increasingly resistant to treatment with antibiotics because of these processes.

Many eukaryotic transposons have base sequences similar to those of retroviruses. Investigators believe that these transposons, called **retrotransposons,** represent degenerated retroviruses. Retrotransposons are transposed by being first transcribed into RNA. Action of *reverse transcriptase* (Section 18.4) produces a DNA complementary to the RNA. This DNA, which constitutes a copy of the original retrotransposon, becomes inserted into the host DNA.

17.6. SPECIAL MECHANISMS OF REPLICATION

17.6.1. Mitochondrial DNA

Mitochondrial DNA is a circular, double-stranded molecule that does not associate with histones. There are about

5–10 copies per mitochondrion. The DNA codes for components of mitochondrial protein synthesis as well as for proteins that function in the electron transport system and oxidative phosphorylation. As you already know, the genetic code of mitochondrial DNA varies somewhat from the "universal" code. Replication of mitochondrial DNA proceeds via a mechanism called **displacement loop** or **D-loop replication** (Figure 17.26).

Mitochondrial DNA has two origins of replication, one for each strand. Synthesis begins at the replication origin of the leading strand template and proceeds *unidirectionally* around the circular chromosome. As the leading strand becomes replicated, it displaces the lagging strand template, forming a **displacement loop (D-loop)** that can be seen with the electron microscope.

When replication of the leading strand has progressed about two-thirds around the chromosome, the D-loop passes and exposes the origin of replication of the lagging strand template. Replication of the lagging strand now begins and proceeds *unidirectionally* in the opposite direction. Because of the delay in initiating the replication of the lagging strand, this strand is only about one-third replicated by the time that replication of the leading strand has been completed. Synthesis of each strand is initiated by means of an RNA primer, and each strand grows in a *continuous* fashion.

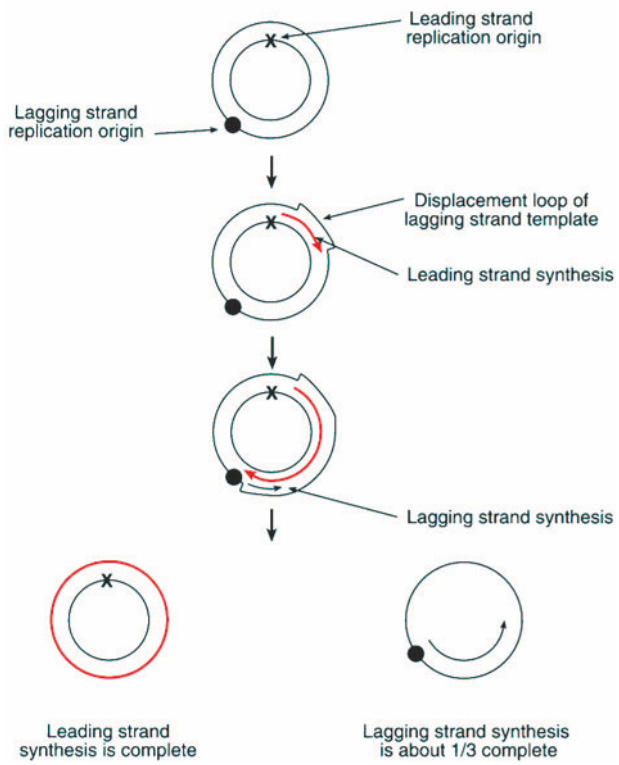


Figure 17.26. Displacement loop (D-loop) model for replication of mitochondrial DNA. Each strand has a replication origin. Replication of the leading strand is about two-thirds complete before replication of the lagging strand is initiated.

17.6.2. Single-Stranded DNA

Some bacteriophages have their genome in the form of circular, single-stranded DNA. Phage ϕX174, which infects *E. coli,* is an example. Replication of these viral DNAs occurs via a mechanism termed **rolling circle replication.** Details of the mechanism vary with the phage involved, but its general principles are diagrammed in Figure 17.27. Phage DNA, after entering a host cell, serves as a template for synthesis of a complementary and antiparallel circular strand, thereby producing a double-stranded circular molecule, called *replicative form (RF).* We designate the original phage strand *plus (+)* and the newly synthesized strand *minus (−).* The latter strand serves as a template for transcription into viral mRNA. These designations are based on a convention that applies to both viral DNA and viral RNA: we consider any strand that has the same polarity as the viral mRNA to be (+), and the complementary and antiparallel strand to be (−). Accordingly, we designate double-stranded nucleic acids as (±).

17.6.3. Viral RNA

Many viruses contain RNA, rather than DNA, as their genetic material. We can group RNA viruses into four classes (Figure 17.28) based on the mechanism for production of viral mRNA after infection of the host cell.

Class I or **positive-strand viruses** (for example, poliovirus) contain single-stranded RNA (+RNA) that serves as a template for synthesis of a complementary RNA (−RNA) that is then transcribed into viral mRNA. **Class II** or **negative-strand viruses** (such as rabies virus) contain single-stranded RNA (−RNA) that serves as a template for synthesis of a complementary RNA (+) that functions directly as viral mRNA. **Class III viruses** (for instance, reovirus) contain double-stranded RNA (±) as their genome and direct the synthesis of an RNA complementary to the (−) strand. The resulting +RNA functions as viral mRNA. Reoviruses are animal viruses that occur in the enteric (intestinal) and respiratory tracts but are not commonly associated with disease. Because such viruses have no links to disease symptoms, we refer to them as *orphans,* hence the term **reovirus** (contraction of *r*espiratory *e*nteric *o*rphan *virus*). **Class IV viruses** are the *retroviruses,* which are discussed in Section 18.4.

Viruses that belong to Classes I, II, and III contain as part of their genome the information necessary to synthe-

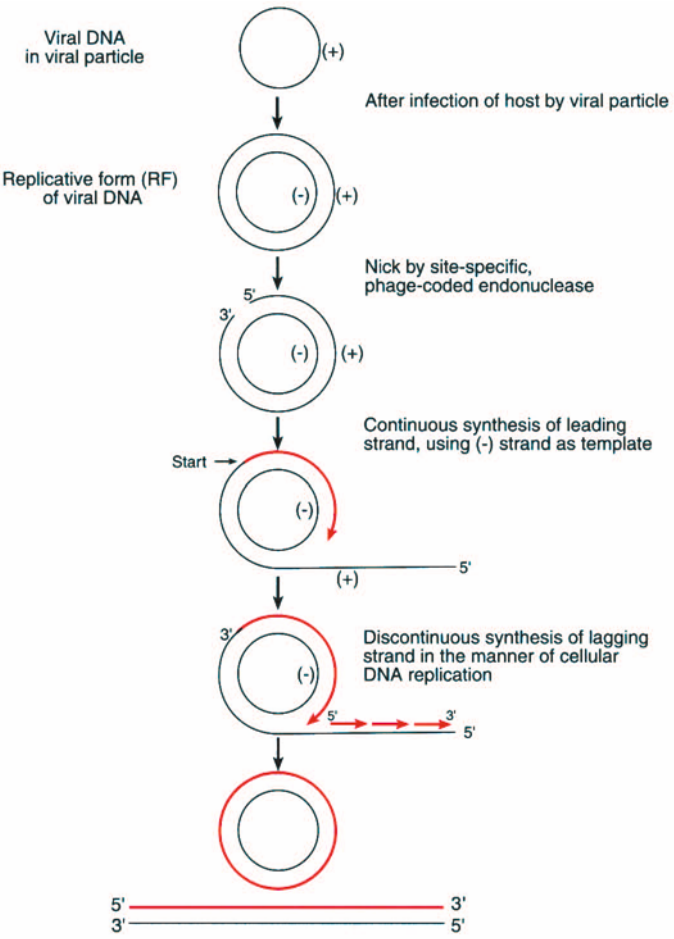


Figure 17.27. Rolling circle model for replication of single-stranded viral DNA. In some systems, the (+) strand does not replicate but becomes incorporated directly into a new viral particle.

size an *RNA-dependent RNA polymerase* called **RNA replicase.** RNA replicase catalyzes the synthesis of an RNA strand complementary and antiparallel to a template RNA strand. Synthesis proceeds in the 5′ → 3′ direction and requires ribonucleoside triphosphates as substrates. RNA replicase is a self-priming enzyme (like primase), but, unlike other DNA and RNA polymerases, it is template specific; it catalyzes replication of the specific viral RNA but not that of the host cell RNA. This enzyme property explains why RNA viruses replicate preferentially inside a host cell.

All known RNA replicases lack a proofreading function such as that associated with DNA-dependent DNA polymerases. Consequently, viral RNA replication is more error-prone than DNA replication, and RNA viruses readily undergo mutational changes. The capacity of

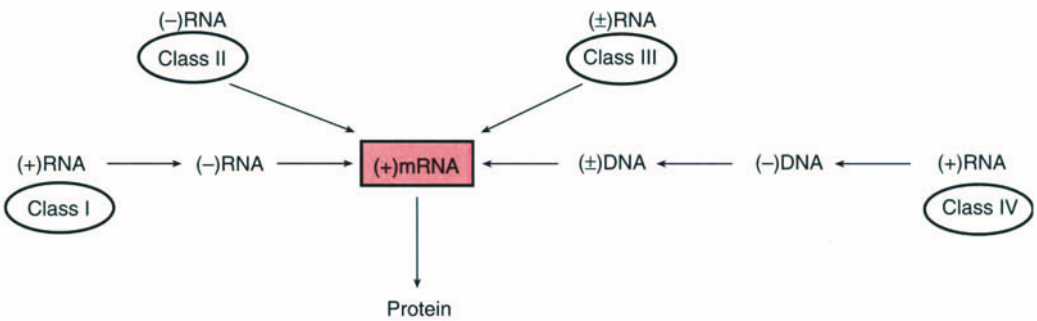


Figure 17.28. Classification of RNA viruses based on the mechanism involved in forming viral mRNA.

viruses to change rapidly aids them in avoiding defense mechanisms of animals and plants and makes it difficult to counteract viruses with synthetic drugs, hence the intense pathogenicity of viral infections. The occurrence, during almost every winter season, of a new and potent strain of the influenza (flu) virus is a well-known case in point.

SUMMARY

DNA replication proceeds via a semiconservative mechanism in which the two parental strands separate and each serves as a template for synthesis of a complementary and antiparallel daughter strand. Experimental proof for this mechanism was provided by the Meselson–Stahl experiment, in which DNA was labeled with ^{15}N and analyzed by density gradient centrifugation over several generations of cell growth.

DNA synthesis involves DNA polymerase, which uses deoxyribonucleoside triphosphates as substrates, cleaving out PP_i for every nucleotide incorporated. DNA polymerase requires a free 3′-OH as primer, and growth of the new strand is in the 5′ → 3′ direction. Three DNA polymerases (I, II, III) occur in prokaryotes; eukaryotes have five such enzymes (α, β, γ, δ, ϵ). In prokaryotes, DNA polymerase III is the major replication enzyme; DNA polymerase I functions in both replication and DNA repair. All three prokaryotic DNA polymerases have an associated 3′ → 5′ exonuclease that catalyzes excision of nucleotides misincorporated by the polymerase. DNA polymerase I and III also have a 5′ → 3′ exonuclease that catalyzes removal of nucleotides H-bonded to the template ahead of the polymerase.

DNA replication requires unwinding of the double helix to form a replication fork. Unwinding is catalyzed by dnaA protein and helicase. Single-strand binding proteins bind to the two template strands and prevent their reassociation. DNA gyrase relieves the twist generated by the unwinding DNA.

Initiation of replication takes place at an origin of replication. *E. coli* DNA has one such origin, but eukaryotic chromosomes contain many. *E. coli* DNA replication proceeds via two replication forks moving bidirectionally; eukaryotic replication is also bidirectional but involves multiple replication forks per chromosome.

Because of the specificity of DNA polymerase and the antiparallel nature of DNA, movement of the polymerase along the two template strands is in opposite directions. Initiation of replication along both strands occurs by means of RNA primers synthesized by primase, a self-priming enzyme. Replication of the leading strand requires one RNA primer; that of the lagging strand requires multiple primers. Each RNA primer becomes extended with a DNA segment by means of DNA polymerase III. The leading strand grows continuously and in the direction of fork movement. The lagging strand grows discontinuously in the form of short fragments (Okazaki fragments) and in a direction opposite to that of fork movement. DNA polymerase I catalyzes excision of the RNA primers by means of its 5′ → 3′ exonuclease activity and catalyzes filling of the gaps between Okazaki fragments by means of its polymerase activity. DNA ligase catalyzes the linking of Okazaki fragments.

DNA can be damaged through errors introduced during replication and by mutation. Replication fidelity is very great, largely due to the 3′ → 5′ exonuclease. Postreplicative repair involves five major mecha-

nisms—direct reversal of damaging reactions, excision repair, mismatch repair, recombination repair, and SOS repair. The genetic material also undergoes a number of changes as part of its normal biological function. Two important mechanisms are those of genetic recombination and transposition. Replication of mitochondrial DNA, single-stranded viral DNA, and single-stranded viral RNA proceeds via special mechanisms.

SELECTED READINGS

Campbell, J. L. (ed.), DNA replication, *Methods in Enzymology* Vol. 262, Academic Press, San Diego (1995).

Coverley, D., and Laskey, R. A., Regulation of eukaryotic DNA replication, *Annu. Rev. Biochem.* 63:745–776 (1994).

Friedberg, E. C., Relationships between DNA repair and transcription, *Annu. Rev. Biochem.* 65:15–42 (1996).

Gruss, C., and Knippers, R., Structure of replicating chromatin, *Prog. Nucleic Acid Res. Mol. Biol.* 52:337–365 (1996).

Joyce, C. M., and Steitz, T. A., Function and structure relationships in DNA polymerases, *Annu. Rev. Biochem.* 63:777–822 (1994).

Kelman, Z., and O'Donnell, M., DNA polymerase III holoenzyme: Structure and function of a chromosomal replicating machine, *Annu. Rev. Biochem.* 64:171–200 (1995).

Kornberg, A., and Baker, T., *DNA Replication,* 2nd ed., W. H. Freeman, New York (1991).

Kunkel, T. A., DNA replication fidelity, *J. Biol. Chem.* 267:18251–18254 (1992).

Lohman, T. M., and Bjornson, K. P., Mechanisms of helicase-catalyzed DNA unwinding, *Annu. Rev. Biochem.* 65:169–214 (1996).

Lohman, T. M., and Ferrari, M. E., *Escherichia coli* single-stranded DNA-binding protein: Multiple DNA-binding modes and cooperativities, *Annu. Rev. Biochem.* 63:527–570 (1994).

Modrich, P., and Lahue, R., Mismatch repair in replication, fidelity, genetic recombination, and cancer biology, *Annu. Rev. Biochem.* 65:101–133 (1996).

Ogawa, T., and Okazaki, T., Discontinuous DNA replication, *Annu. Rev. Biochem.* 49:421–458 (1980).

Sancar, A., DNA excision repair, *Annu. Rev. Biochem.* 65:43–81 (1996).

Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L., The structural basis of specific base-excision repair by uracil-DNA glycosylase, *Nature (London)* 374:487–493 (1995).

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. S., Viable offspring derived from fetal and adult mammalian cells, *Nature (London)* 385:810–813 (1997).

Wood, R. D., DNA repair in eukaryotes, *Annu. Rev. Biochem.* 65:135–167 (1996).

REVIEW QUESTIONS

A. Define each of the following terms:

Processivity	Nick translation
Okazaki fragment	DNA glycosylase
Insertion sequence	Proofreading function
Transposable element	Postreplicative processing
Pyrimidine dimer	AP endonuclease
SSB	Theta replication
SOS response	AP site

B. Differentiate between the two terms in each of the following pairs:

Replisome/replicon	Primase/primosome
Transposon/retrotransposon	Pol I/pol III
DNA gyrase/DNA ligase	Misincorporation/mismatch repair
Leading strand/ lagging strand	Excision repair/recombination repair
5′ → 3′ exonuclease/ 3′ → 5′ exonuclease	Positive-strand virus/negative-strand virus
RNA replicase/helicase	Insertion/deletion

C. (1) Outline the Meselson–Stahl experiment and explain why it supports the concept of semiconservative replication.

(2) How can DNA be damaged and what are the mechanisms for repairing it?

(3) Describe the sequence of events at the replication fork during discontinuous DNA replication in prokaryotes.

(4) Outline the steps of (a) D-loop replication and (b) rolling circle replication.

(5) List all of the proteins required for replication of *E. coli* DNA. What is the specific function of each protein?

PROBLEMS

17.1. Why is it advantageous for replication that the *ori*C site of *E. coli* contains three consecutive A/T-rich segments rather than other segments?

17.2. Why does the role of RNA primers in DNA replication support the notion that RNA developed before DNA in the course of evolution?

17.3. The following represents a partial sequence in a lagging strand template: 3′-ACTGTCGATGGAC-CA-5′. What is the sequence of the corresponding (a) RNA primer, (b) Okazaki fragment, and (c) RNA primer if the template cytosines were first deaminated to uracils?

17.4.* Assume that replication has resulted (at time t_1) in the production of 20 Okazaki fragments, each having a length of 1000 nucleotides. Assume further that, at time t_2, one-half of these fragments have become linked to produce one continuous DNA segment. What is the average molecular weight of the DNA of the lagging strand at time t_1 and time t_2 in terms of the number of nucleotides?

17.5. A misincorporation rate of one base per 10^{10} bases corresponds to an erroneous incorporation of one base per how many cells of *Mycoplasma hominis* per generation? (See Table 7.3.)

17.6. As a graduate student, you propose to determine whether compounds other than acridine dyes and ethidium bromide can be made to intercalate with DNA. Which of the following compounds might you consider for your study? (a) heme; (b) coenzyme A; (c) NAD^+; (d) testosterone. Why?

17.7. How many Okazaki fragments, each 1000 nucleotides long, would be required to cover the entire lagging strand template of *E. coli?* (See Table 7.3.)

17.8. DNA replication in eukaryotes proceeds at a rate of about 4000 bp/min per replication fork. How long would it take to replicate lungfish DNA (refer to Table 7.3), assuming that all 19 chromosomes replicate simultaneously and that each chromosome replicates via two replication forks moving bidirectionally from a single origin (as in *E. coli*)?

17.9. Why is a molecule of DNA better suited for storing genetic information than a molecule of RNA?

17.10. Beginning with fully labeled DNA ("all heavy" DNA), what will be the distribution of products after precisely three generations in a Meselson–Stahl experiment?

17.11. An *in vitro* assay of DNA polymerase I uses dTTP labeled with ^{32}P and measures the amount of labeled DNA formed. The phosphate group closest to the sugar ring is designated α, the next β, and the farthest γ. A student prepares a number of dTTP forms labeled in: (a) α; (b) β; (c) γ; (d) α, β; (e) β; γ; (f) α, β, γ. Which of these compounds represent suitable substrates for the assay? Why?

17.12. To determine the movement of the replication fork in an unknown bacterium, you first grow cells in a medium containing thymidine that is extensively labeled with tritium. You then transfer the cells to a medium containing thymidine that is slightly labeled with tritium and isolate and analyze the DNA. Sketch the autoradiograph you would expect if the DNA were replicated (a) unidirectionally or (b) bidirectionally.

17.13.* Refer to Tables 7.3 and 7.6 to calculate the number of turns that must be unwound as the B-DNA of the *E. coli* chromosome undergoes replication. How many turns would have to be unwound if the chromosome consisted of Z-DNA?

17.14. How long does it take for the gene of ribonuclease (104 amino acids) to be replicated during cell division in *E. coli,* given that the rate of replication is 60,000 bp/min per replication fork? How long would it take to replicate the same gene in eukaryotes, assuming the absence of introns and a leader sequence? Use 4000 bp/min per replication fork as the rate of replication in eukaryotes.

17.15.* The rate of replication in *E. coli* is 60,000 bp/min per replication fork, and 10 bp of B-DNA represent a length of 3.4 nm of double helix (see Table 7.6). What is the rate of DNA synthesis in *E. coli* in units of centimeters per hour?

17.16.* Consider the ATP requirement for unwinding the double helix by helicase. On this basis alone, how many molecules of glucose would have to be degraded under aerobic conditions to supply the energy necessary for unwinding the double helix of *E. coli* DNA?

17.17. What are the fractions of "all heavy," "all light," and "hybrid" DNA in a Meselson–Stahl experiment carried out for precisely four generations?

17.18.* A line of mammalian cells in tissue culture has 1.2 meters of duplex DNA per cell. These cells divide every five hours. If the rate of DNA duplex growth is 16 microns per minute per replication fork, how many replication forks must be operating during chromosome replication? (1.0 micron = 10^{-6} meter)

17.19.* In a Meselson–Stahl experiment, a sample of cell culture contains 5 mg of "hybrid" DNA and 45 mg of "all light" DNA. If you allow this sample of cells to grow for two additional generations, what would be the amount of "hybrid" DNA in units of milligrams and its concentration in units of percent by weight?

17.20.* A positive-strand virus is treated with nitrous acid, which deaminates the cytosines of the viral RNA to uracils. The mutated virus is used to infect host cells and leads to synthesis of viral proteins. What amino acids in the mutated protein are likely to be found

in some of the positions occupied by serine in the normal viral protein?

17.21.* Under optimal conditions, the entire chromosome of *E. coli* is replicated in 40 min. Use data from Table 7.3 to calculate the distance traversed by one replication fork in terms of: (a) base pairs per minute; (b) microns (μm) per minute.

17.22. In *E. coli,* which enzymes are common to both DNA replication and excision repair of DNA?

17.23. Adenine is spontaneously deaminated to hypoxanthine, and the latter can form a base pair with cytosine. If this occurs, what mutated base pair would be formed in DNA from an A·T base pair?

17.24. Give the complete structure of the following inverted repeat by replacing [X] with the correct bases:

5′ G—A—T—C—A—T—X—X—X—X—X—X 3′
3′ X—X—X—X—X—X—X—X—X—X—X—X 5′

17.25. Why is the ribonucleotide 5-fluorouridine triphosphate likely to act as an inhibitor for DNA replication?

17.26. Replication of nuclear DNA by nuclear DNA polymerase proceeds with much greater fidelity than replication of mitochondrial DNA by mitochondrial DNA polymerase. Why can a lower fidelity of replication be tolerated for mitochondrial DNA but not for nuclear DNA?

17.27. A scientist discovers that a novel bacterium contains an unusual DNA ligase that uses neither NAD^+ nor ATP for the ligation reaction. Which of the following compounds might possibly serve as energy sources for the reaction? (a) FAD; (b) FMN; (c) GTP; (d) CoA—SH; (e) $NADP^+$; (f) AMP; (g) CoQ. Why?

17.28.* A density gradient is set up in a tube. The density changes in a linear fashion from 1.600 to 1.800 g/cm^3 over a distance of 3.00 cm. A synthetic oligonucleotide derivative bands in this gradient at a point 2.00 cm below the top of the gradient. What is the density of the oligonucleotide?

Transcription—The Synthesis of RNA

18

Synthesis of messenger RNA by transcription of the genetic information contained in DNA constitutes the first stage of protein biosynthesis. In the second stage (translation), mRNA serves as the template that directs amino acid polymerization. The term *transcription* is well chosen: a sequence of *deoxyribonucleotides* in DNA becomes *transcribed* into a sequence of *ribonucleotides* in RNA. Transcription retains the genetic information in the form of nucleotide sequences but alters the type of nucleotide. In addition to generating mRNAs, transcription produces ribosomal, transfer, and small RNAs. The basic mechanism of transcription is similar in all cells and requires catalysis by **RNA polymerase (DNA-dependent RNA polymerase).** We may schematically represent the polymerization as follows:

$$n(\text{NTP}) \xrightarrow[\text{Mg}^{2+}]{\text{DNA template}} [\text{RNA}] + n(\text{PP}_i) \xrightarrow{\text{pyrophosphatase}} 2n(\text{P}_i)$$

where NTP is ribonucleoside triphosphate (ATP, CTP, GTP, and UTP). The enzyme requires all four ribonucleoside triphosphates, as well as magnesium ions, and functions as a self-priming enzyme; it *does not require a primer* for initiation of polymerization. Only one of the two DNA strands (the template or anticoding strand) is transcribed. Pyrophosphatase-catalyzed hydrolysis of pyrophosphate to inorganic phosphate drives RNA synthesis, much as it drives DNA synthesis.

The reaction mechanism of RNA polymerase is analogous to that of DNA polymerase (see Figure 17.3) except that ribonucleoside triphosphates replace deoxyribonucleoside triphosphates, and synthesis involves formation of the following complementary base pairs:

DNA:	A	T	C	G
	⋮	⋮	⋮⋮	⋮⋮
RNA:	U	A	G	C

RNA polymerase action yields an RNA that is complementary and antiparallel to the DNA strand transcribed. RNA synthesis also proceeds in the 5′ → 3′ direction so that the RNA polymerase must move along the template strand from its 3′-end to its 5′-end. Note that the RNA polymerase of transcription differs from the primase of DNA replication.

18.1. INITIATION OF TRANSCRIPTION

18.1.1. RNA Polymerases

Both prokaryotic and eukaryotic RNA polymerases are oligomeric proteins. *E. coli* RNA polymerase consists of five subunits linked noncovalently. The pentamer ($\alpha_2\beta\beta'\sigma$) has a molecular weight of about 500,000. An aggregate of four subunits ($\alpha_2\beta\beta'$) constitutes the **core enzyme** and possesses the catalytic activity. The **sigma subunit** (σ) has no catalytic activity and serves to guide and bind the enzyme to the **promoter.** A promoter is a locus on DNA at which RNA polymerase becomes bound and at which transcription is initiated.

The sigma subunit attaches to the core enzyme prior to initiation of transcription, and the entire complex (the *holoenzyme*) binds first nonspecifically to the DNA. Following this initial binding, the complex drifts along the DNA to become bound specifically to the promoter locus for initiation of transcription. After about four nucleotides have become polymerized, the sigma subunit dissociates from the core enzyme, and the core enzyme continues with polymerization by moving along the DNA. We call this reaction sequence the **sigma cycle** (Figure 18.1).

Eukaryotic RNA polymerases have molecular weights in the range of 500,000–700,000. Five different types occur, three in the nucleus and one each in mitochondria and chloroplasts. Nuclear RNA polymerases differ in their intranuclear location and in their specificity for RNA synthesis: RNA polymerase I is found in the nucleolus and catalyzes the synthesis of most ribosomal RNAs; RNA polymerase II occurs in the nucleoplasm (the protoplasm of the nucleus) and catalyzes the synthesis of messenger RNA and several other RNAs; RNA polymerase III also occurs in the nucleoplasm and catalyzes the synthesis of transfer RNA, 5S ribosomal RNA, and a variety of small RNAs. The RNA polymerases of mitochondria and chloroplasts catalyze the synthesis of the different types of RNA present in these subcellular organelles.

Nuclear RNA polymerases can be distinguished by their sensitivity to α-*amanitin,* a toxic bicyclic octapeptide derived from the poisonous mushroom *Amanita phalloides.* RNA polymerase I is very resistant to α-amanitin; it takes a concentration greater than $10^{-3}M$ to produce any effect. RNA polymerase II is very sensitive to the compound and is inhibited by concentrations as low as 10^{-9}–$10^{-8}M$. The effect on RNA polymerase III is intermediate; the enzyme is inhibited by concentrations of α-amanitin of 10^{-5}–$10^{-4}M$.

18.1.2. Promoters

Biochemists customarily represent promoters and other DNA segments by base sequences in the *sense strand* or *coding strand.* We choose this strand, rather than the template strand, because it runs in the same direction (5′ → 3′) as the RNA strand being synthesized. We assign numbers to the base pairs in the promoter region, indicating their positions relative to the site of transcription initiation. To that site we assign the number +1. Base pairs preceding this site are given negative numbers and are not transcribed. Base pairs following the initiation site are given positive numbers. We do not assign zero. Proceeding toward the 5′-end of the sense strand is termed mov-

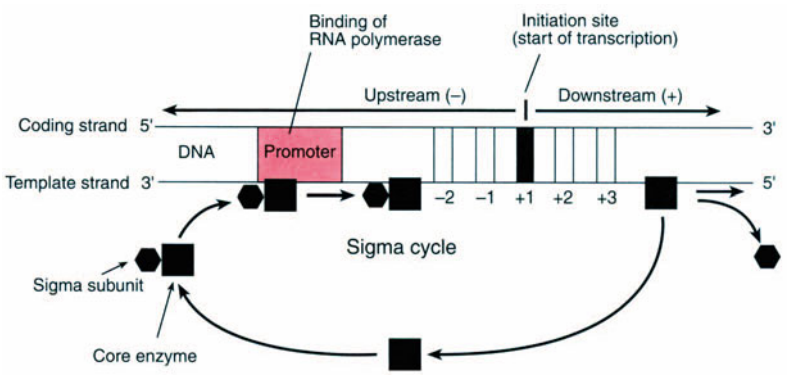


Figure 18.1. The sigma cycle in transcription.

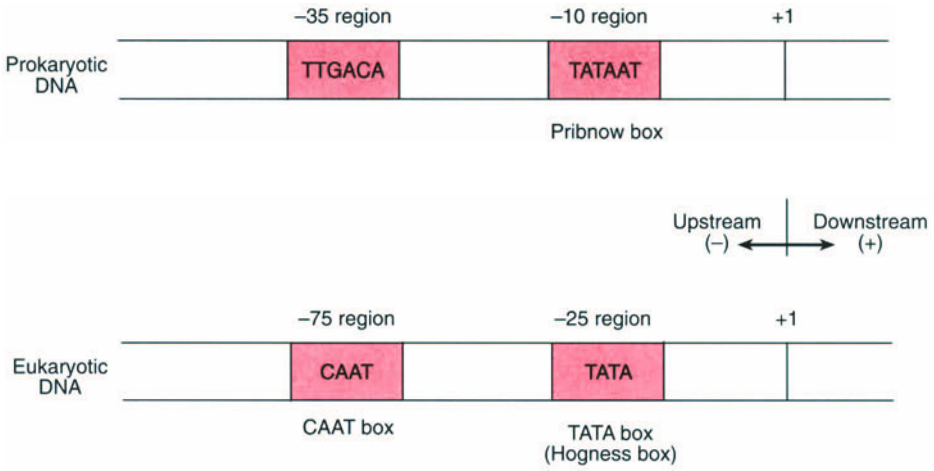


Figure 18.2. Consensus sequences in the promoter regions of prokaryotic and eukaryotic DNA.

ing *upstream;* the opposite direction is *downstream* (see Figure 18.1).

In dealing with these conventions, keep in mind that either strand of DNA can, in principle, be transcribed. Some genes may be transcribed from one strand, some from the other strand. Regardless of which strand is transcribed, RNA synthesis proceeds in the 5′ → 3′ direction, and the RNA is complementary and antiparallel to the transcribed strand.

Researchers have determined promoter sequences for numerous *E. coli* genes and found that certain nucleotides are *conserved*—they occur in a large number of cases. We call a conserved nucleotide sequence a **consensus sequence.** Most promoters differ from a consensus sequence by at most one or two nucleotides. Both prokaryotic and eukaryotic promoters usually contain two consensus sequences. For prokaryotes, these sequences occur at about −10 and −35 and are named the −10 region (*Pribnow box*) and the −35 region, respectively (Figure 18.2). For eukaryotes, consensus sequences occur at about −25 and −75 and are named the −25 region (*TATA box* or *Hogness box*) and the −75 region (*CAAT box*), respectively.

In both prokaryotes and eukaryotes, the consensus sequence close to the start of transcription consists of an A/T-rich region. This allows local unwinding of double-stranded DNA, a prerequisite for transcription much as it is for replication. However, since the RNA polymerase system of transcription is less complex than the replisome system of replication, DNA unwinding and rewinding occur without expenditure of energy.

In addition to being A/T-rich, consensus sequences of promoters are asymmetric; they are not palindromes or inverted repeats. Accordingly, a consensus sequence ensures that transcription proceeds only in one direction, and along one strand, in that section of DNA.

18.1.3. Mechanism of Initiation

Based on the foregoing, we can summarize the initiation of transcription in *E. coli* as follows (Figure 18.3):

1. A sigma subunit binds to the core enzyme, forming the RNA polymerase holoenzyme.
2. The holoenzyme binds electrostatically and nonspecifically to some region of the DNA upstream from the promoter.
3. The RNA polymerase holoenzyme moves downstream and binds specifically to the promoter. The binding interaction involves the consensus sequence of the promoter. Binding of the holoenzyme causes partial unwinding of a segment of DNA, stretching from about −9 to +2 (about 11 bp).
4. The enzyme binds either ATP or GTP in a *self-priming step.* The bound nucleoside triphosphate constitutes the 5′-end of the RNA transcript and is retained throughout the transcription process. Binding a nucleoside triphosphate generates a free 3′-OH to which the next nucleotide becomes esterified.
5. About four nucleotides become polymerized to the ATP (GTP) as the holoenzyme continues downstream.
6. The sigma subunit dissociates from the holoenzyme, and the core enzyme continues by itself. DNA rewinds as the holoenzyme passes. The newly synthesized RNA strand is complementary and antiparallel to the DNA strand transcribed and grows in the 5′ → 3′ direction.

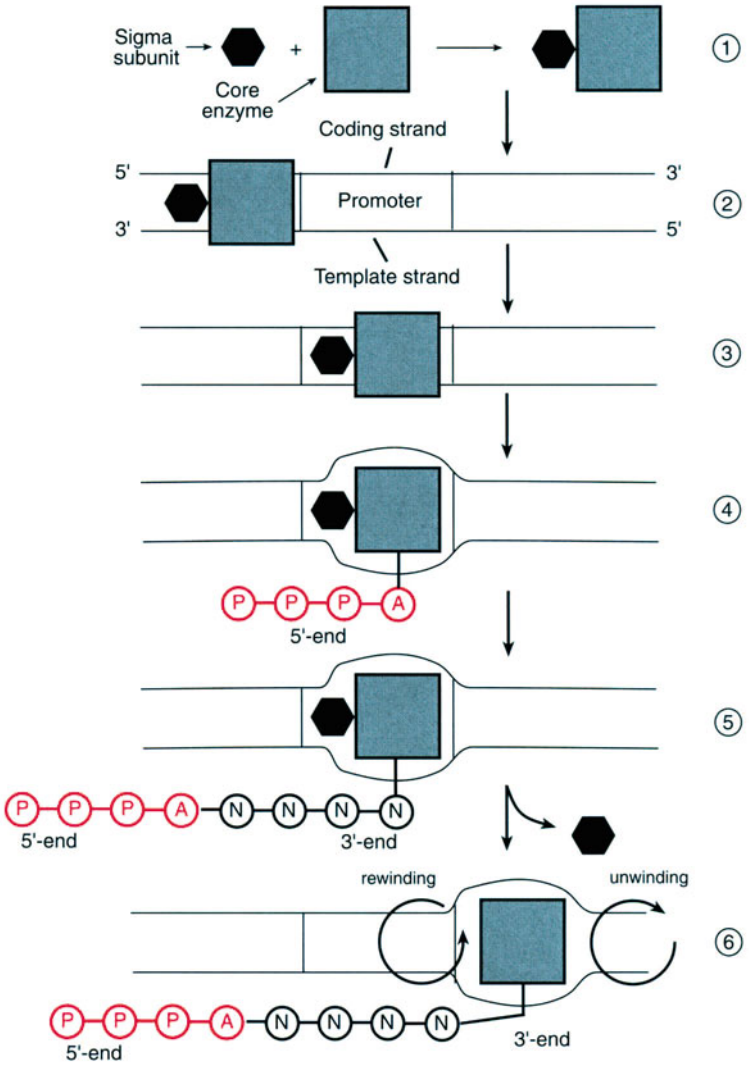


Figure 18.3. The initiation stage of transcription in prokaryotes. Numbers refer to the stages described in the text.

18.2. ELONGATION AND TERMINATION

After the initial ATP (or GTP) has become bound to the RNA polymerase holoenzyme, a second nucleotide becomes incorporated via formation of a 3′,5′-phosphodiester bond with the 3′-OH of ATP (or GTP). The second nucleotide generally derives from either UTP or CTP. Incorporation of the second and subsequent nucleotides proceeds via the basic reaction mechanism of RNA polymerase. Repeated formation of 3′,5′-phosphodiester bonds between incorporated nucleotides constitutes the *elongation stage* of transcription, during which the bulk of the RNA molecule becomes synthesized. As RNA polymerase proceeds along the template, DNA rewinds behind the enzyme to re-form the double helix. The elongation stage can be visualized by means of the electron microscope (Figure 18.4).

DNA contains specific sites at which transcription comes to a stop. In *E. coli,* termination involves two mechanisms, termed *rho-independent* and *rho-dependent.* Rho (ρ) is an oligomeric protein composed of six identical subunits (MW = 46,000 each).

Rho-independent termination is based on the occurrence of an A/T-rich segment in DNA, preceded by a G/C-rich segment. The G/C-rich segment is such that the RNA transcript formed from it can fold back on itself to produce a *hairpin (stem and loop)* type structure. The hairpin ends with several U residues, complementary to A residues in the template (Figure 18.5). Researchers believe that as the RNA transcript is produced, the hairpin

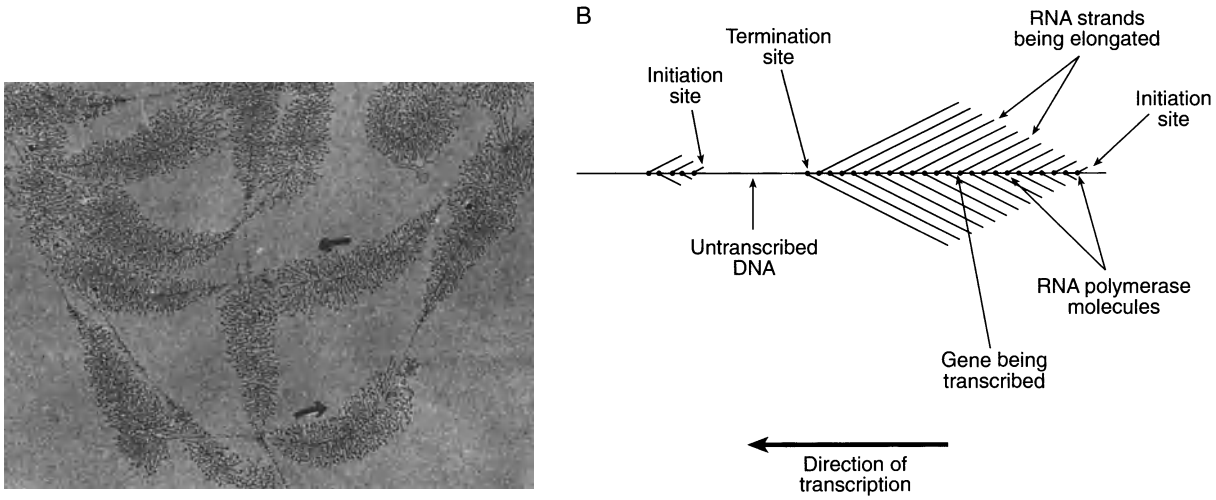


Figure 18.4. Visualization of transcription. (A) Electron micrograph of DNA from *Triturus viridescens* undergoing transcription. [Reprinted, with permission, from O. L. Miller and B. R. Beatty, *J. Cell. Physiol.* 74:225–232 (1969), Supplement. Copyright © 1969 by Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc.] (B) Interpretive drawing. Many RNA polymerase molecules transcribe a gene simultaneously; they carry identical mRNA strands but in various stages of completion.

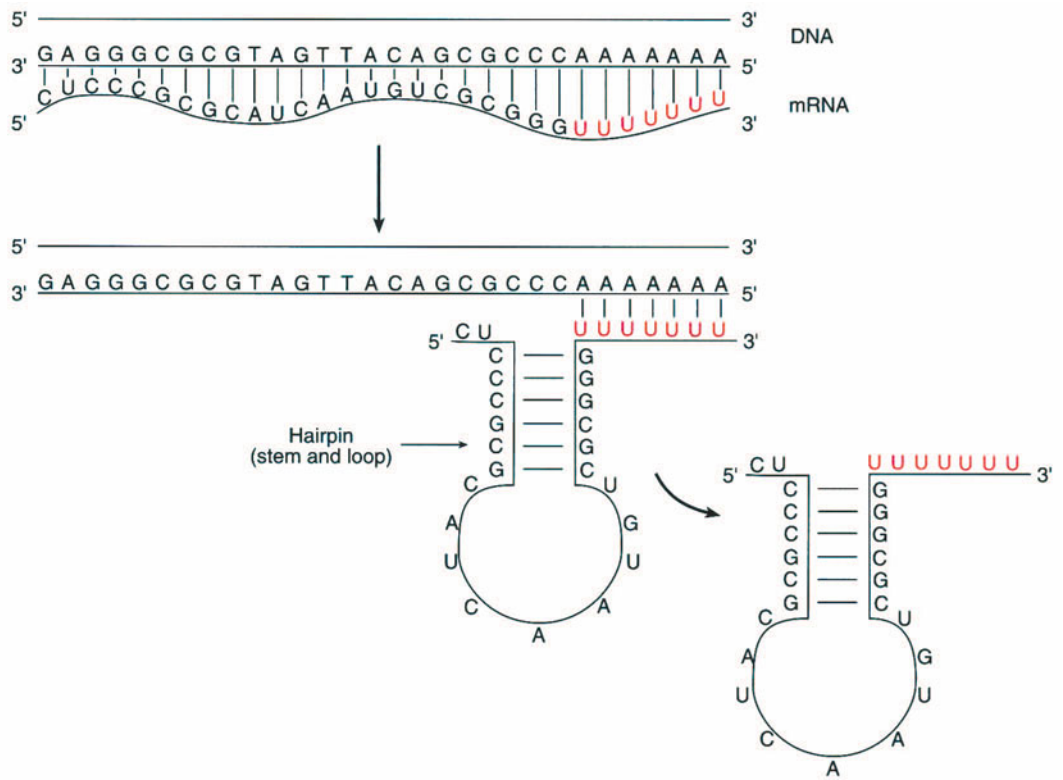


Figure 18.5. Rho-independent termination of transcription. Hairpin formation causes partial displacement of mRNA from the template. Disruption of H-bonding between the template's poly(A) and the mRNA's poly(U) causes complete displacement.

forms and forces the polymerase to pause momentarily. Some have suggested that the stopped polymerase undergoes a conformational change that leads to displacement of the poly(U) tail from the poly(A) segment of the template. This displacement is relatively easy, since A and U are linked via only two H-bonds. Detachment of the poly(U) segment of the RNA from the DNA template signals the termination of transcription.

Rho-dependent termination is called into play at sites that do not have A/T- and G/C-rich segments and requires participation of the rho protein. Scientists believe that rho catalyzes an unwinding of the RNA–DNA hybrid, presumably as a result of helicase activity associated with the rho protein.

18.3. POSTTRANSCRIPTIONAL PROCESSING

Many RNAs are transcribed initially in forms that do not have full activity. Such **primary transcripts** require **posttranscriptional processing** to convert them to active and **mature RNAs.** Posttranscriptional processing includes three general types of molecular alterations: (a) removal of nucleotides from the primary transcript; (b) addition of nucleotides to the transcript; and (c) modification of specific bases in the transcript.

18.3.1. Processing of Ribosomal and Transfer RNAs

Transcription of ribosomal RNA of both prokaryotes and eukaryotes yields large precursor molecules (**pre-rRNA**). Prokaryotic pre-rRNAs have a sedimentation coefficient of about 30S and contain one copy each of 5S, 16S, and 23S rRNA. The primary rRNA transcript may also contain several tRNA precursors. Eukaryotic pre-rRNAs have sedimentation coefficients of about 35–47S and contain one copy each of 5.8S, 18S, and 28S rRNA. The fourth eukaryotic rRNA, 5S RNA, is transcribed and processed separately. Specific nucleases, acting in several steps, cleave large pre-rRNAs to produce mature ribosomal RNAs (see Table 16.2).

Transcription of transfer RNA of both prokaryotes and eukaryotes also yields large precursors (**pre-tRNA**). Many eukaryotic pre-tRNAs contain a short intron next to the anticodon and extra nucleotides at their 3′- and 5′-ends. Specific nucleases, acting sequentially, cleave the primary transcripts to form mature tRNAs. An unusual enzyme, *ribonuclease P,* catalyzes the initial cleavage of most pre-tRNAs to produce the 5′-end of the molecule (pG). Ribonuclease P consists of one RNA molecule (377 nucleotides) and one protein (MW ≈ 20,000). Both components are required for full activity, but the *catalytic subunit is the RNA* and *not* the protein. The protein serves to maintain the proper folding of the RNA and to maximize its catalytic activity. Ribonuclease P constitutes a **ribozyme,** a *ribo*nucleic acid that has en*zyme*-like catalytic activity.

Both prokaryotic and eukaryotic mature tRNAs require the sequence CCA-OH at their 3′-end. Because some primary transcripts lack this sequence while others lose it due to exonuclease cleavage, most tRNAs require posttranscriptional formation of the 3′-end. The enzyme *nucleotidyl transferase* catalyzes formation of the CCA-OH sequence from ATP and 2 CTP.

The many unusual bases that constitute part of tRNA's structure form by modifications of standard bases. These modifications are introduced into the tRNA after the "parent" molecule, containing the standard bases, has been transcribed.

18.3.2. Processing of Messenger RNA

Prokaryotic mRNA does not undergo processing. The primary mRNA transcript serves directly for translation. In fact, transcription and translation are coupled and proceed simultaneously. Coupling is made possible because both processes occur in the same cellular compartment (cytoplasm). In eukaryotes, coupling of transcription and translation is not possible because the two processes occur in separate cellular compartments (nucleus and cytoplasm), separated by the nuclear membrane. Eukaryotic primary transcripts of mRNA (**pre-mRNA**) are subject to extensive processing that includes *exon splicing, capping,* and *polyadenylation.*

18.3.2A. Exon Splicing. Many eukaryotic genes occur as **discontinuous genes** or **split genes.** Such genes consist of two types of base sequences. Some sequences carry genetic information and code for segments of mRNA. We call these DNA sequences *exons* because they lead to gene *ex*pression. Other base sequences of the gene do not code for mRNA but may contain some regulatory sequences. We call these *in*tervening sequences *introns.* Introns separate one exon from another. Note that we use the terms exon and intron also for *the corresponding base sequences in pre-mRNA.*

The primary transcript of a discontinuous gene consists of a large nonfunctional mRNA that contains copies of both exons and introns. This transcript must be processed so that a functional mRNA can be assembled

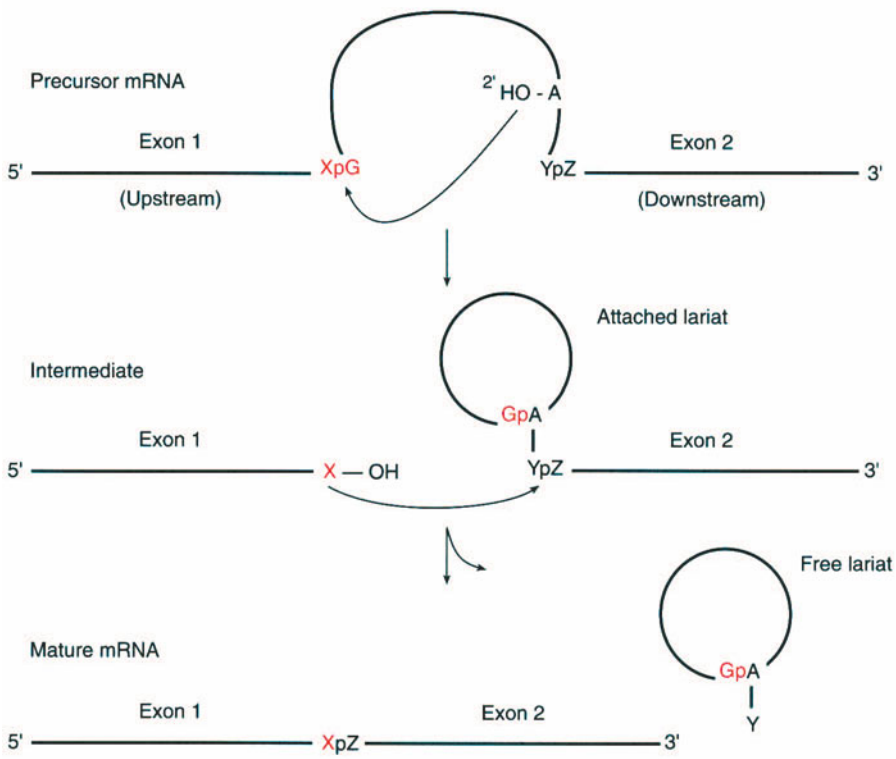


Figure 18.6. Proposed mechanism for exon splicing in eukaryotic pre-mRNA. A 2′-OH of an internal AMP residue in the intron attacks the phosphodiester bond linking the intron's 5′-end to the upstream exon. The upstream exon's 3′-OH then attacks the junction between the intron and the downstream exon. The two exons join, and the intron is released as a free lariat.

from its component parts. Conversion of pre-mRNA to mature mRNA requires a cutting out, or *excision,* of introns and a joining, or *splicing,* of exons.

Exon splicing is catalyzed by complexes consisting of *small nuclear RNAs (snRNA)* that combine with proteins to form *small nuclear ribonucleoproteins (snRNP,* pronounced "snurp"). Aggregates of snRNP form poorly characterized complexes, dubbed **spliceosomes** (about 40–60S). Investigators believe that a spliceosome "recognizes" a splicing junction by complementary base pairing between its own snRNA and that of the transcript.

Splicing involves two consecutive nucleophilic attacks by hydroxyl groups on phosphodiester bonds. The intron is first converted to an attached *lariat* (tailed circle), followed by release as a free lariat (Figure 18.6) that subsequently undergoes degradation in the nucleus.

18.3.2B. Capping. Eukaryotic mRNA undergoes two other modifications—attachment of a *methylated cap* at its 5′-end and attachment of a *poly(A) tail* at its 3′-end. **Capping** at the 5′-end represents the only modification of RNA that occurs while transcription is still in progress; it constitutes **cotranscriptional processing.** Attachment of the methylated cap is enzyme-catalyzed and occurs in four steps before the RNA is more than 20 nucleotides long:

1. Removal of the terminal 5′-phosphate group by phosphatase:

$$\underset{5'\ \text{mRNA}\ 3'}{\text{ppp}\text{———}} + H_2O \xrightarrow{Mg^{2+}} \underset{5'\ \text{mRNA}\ 3'}{\text{pp}\text{———}} + P_i$$

2. Transfer of a GMP residue from GTP to the ϵ-NH_2 group of lysine in a *capping enzyme* (E):

$$\text{E-Lys} + \text{GTP} \longrightarrow \text{E-Lys-GMP} + PP_i$$

3. Transfer of the GMP residue from the enzyme to the 5′-pyrophosphoryl group of mRNA:

$$\text{E-Lys-GMP} + \underset{5'\ \text{mRNA}\ 3'}{\text{pp}\text{———}} \longrightarrow \underset{5'\ \text{mRNA}\ 3'}{\text{Gppp}\text{———}} + \text{E-Lys}$$

4. Methylation by *S-adenosylmethionine* (*SAM;* see Figure 13.4) at N(7) of the 5′-terminal guanine and at C(2) of the ribose of one or two nucleotides at the 5′-end (Figure 18.7). We believe

that the cap serves to protect the 5′-end of mRNA against degradation by nucleases and to aid in positioning mRNA on the ribosome.

Cap-1 designates a structure in which the first nucleoside following 7-methylguanosine carries a methyl group. Cap-2 indicates that both the first and the second nucleoside carry a methyl group, and Cap-0 indicates that neither nucleoside is methylated. Cap-1 is the predominant form in multicellular organisms, and Cap-0 predominates in unicellular eukaryotes.

18.3.2C. Polyadenylation. Attachment of a poly(A) tail at the 3′-end of eukaryotic mRNA (**polyadenylation**) involves the sequential addition of a large number of adenylic acid residues (50–200) by means of *poly(A) polymerase.* A nuclease first cleaves pre-mRNA at a site some 10–30 nucleotides downstream from a highly conserved sequence of AAUAAA. After this cleavage, poly(A) polymerase catalyzes the addition of a poly(A) tail (Figure 18.8). We think that the poly(A) tail serves to protect the 3′-end of mRNA against degradation by nucleases and to aid in the transfer of mRNA from the nucleus to the cytoplasm.

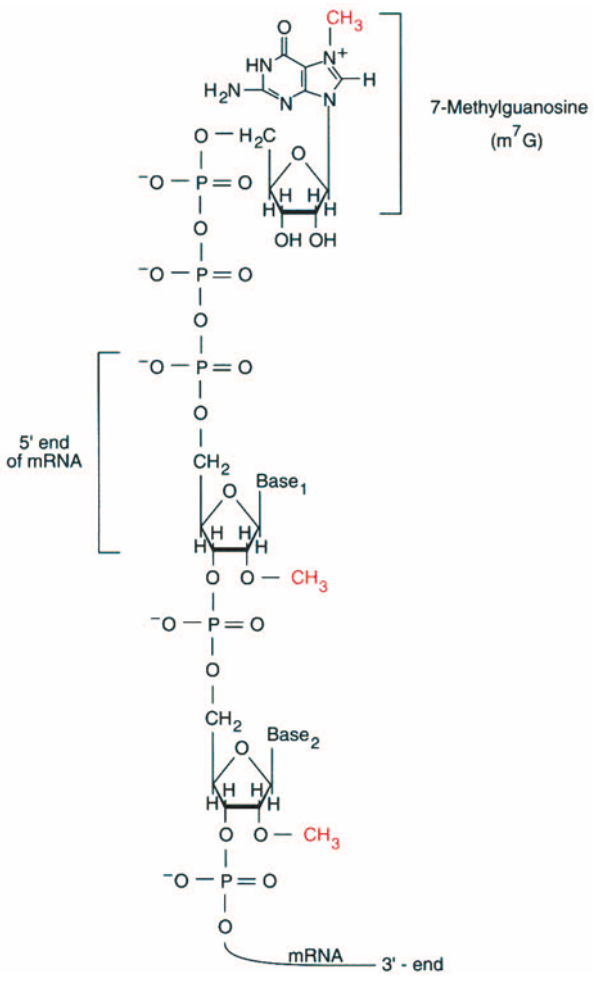


Figure 18.7. Methylated caps at the 5′-end of eukaryotic mRNA.

18.3.3. Self-Splicing RNAs

In 1982, the surprising discovery was made that some RNA transcripts are **self-splicing;** they can remove their own introns without the aid of protein enzymes. Because these RNAs have catalytic activity, we consider them *ribozymes.* T. R. Cech first demonstrated this phenomenon for ribosomal RNA of the protozoan *Tetrahymena.* The splicing mechanism resembles that outlined in Figure 18.6 except that the first step is initiated when a 3′-OH of a free guanine nucleotide attacks an upstream phosphodiester bond (Figure 18.9).

The occurrence of ribozymes, such as self-splicing RNAs and ribonuclease P, has revolutionized our concept of biological catalysis. The existence of ribozymes also suggests that the earliest living organisms may have been based entirely on RNA and that DNA and proteins evolved at a later stage (see Section 1.1).

18.4. REVERSE TRANSCRIPTION

18.4.1. Reverse Transcriptase

Most viruses infect and kill susceptible cells; normal cell growth is incompatible with virus replication. These viruses have a genome that consists of either DNA (for example, pox virus) or RNA (for example, poliovirus). However, some RNA-containing viruses *transform* affected eukaryotic cells rather than killing them outright. Because these viruses convert normal cells to malignant ones, we refer to them as **oncogenic** (cancer-producing) **RNA viruses.** Cells infected by such viruses survive and continue to divide but also form new virus particles. An example of an oncogenic RNA virus is the *Rous sarcoma virus,* which contains single-stranded RNA and produces cancer in chickens.

Because a cell's genome determines all cellular properties, transformation of a normal cell to a malignant one must result from a change in the cell's genome. But how can the *DNA* genome of a eukaryotic cell be altered by infection with an *RNA*-containing virus? The answer to this question remained a mystery for a long time until Howard Temin proposed, in 1962, that oncogenic RNA viruses contain an enzyme that can catalyze the *transcription of RNA to DNA.* His proposal met with strong

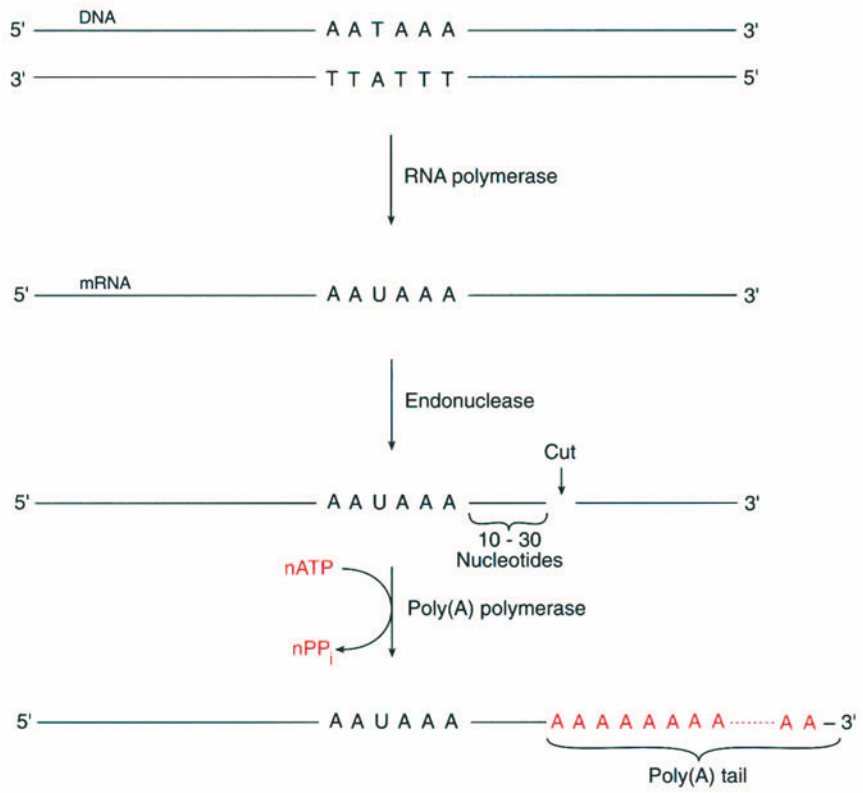


Figure 18.8. Formation of a poly(A) tail at the 3′-end of eukaryotic mRNA.

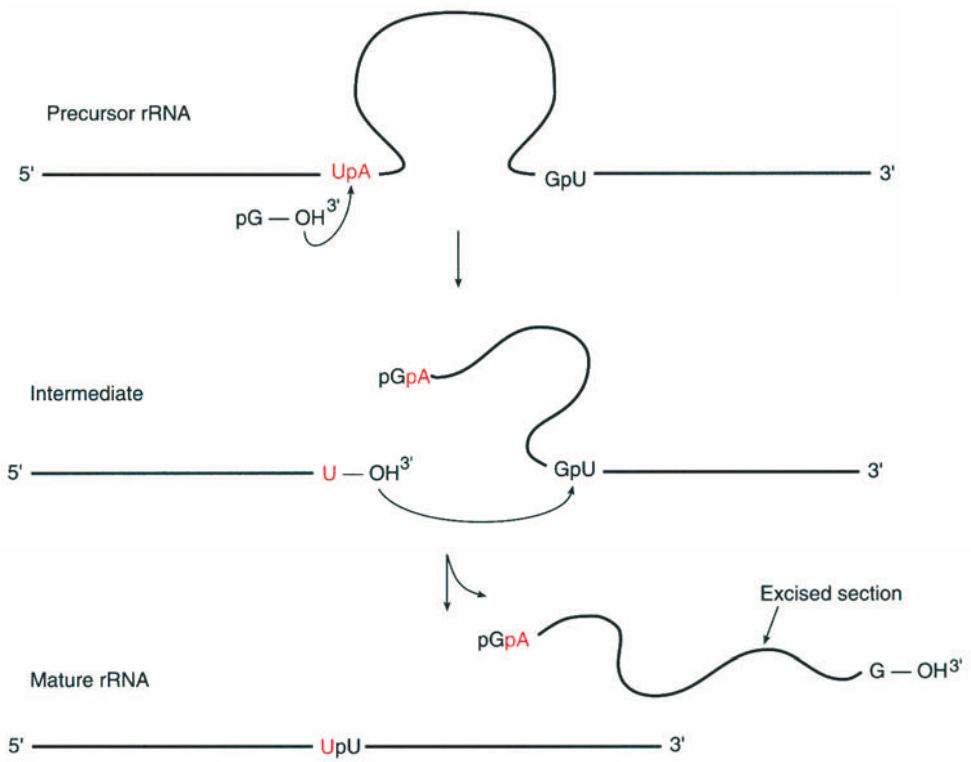


Figure 18.9. Proposed mechanism for pre-rRNA self-splicing. The mechanism resembles exon splicing (Figure 18.6), except that it is initiated by attack of a 3′-OH group of a free guanine nucleotide.

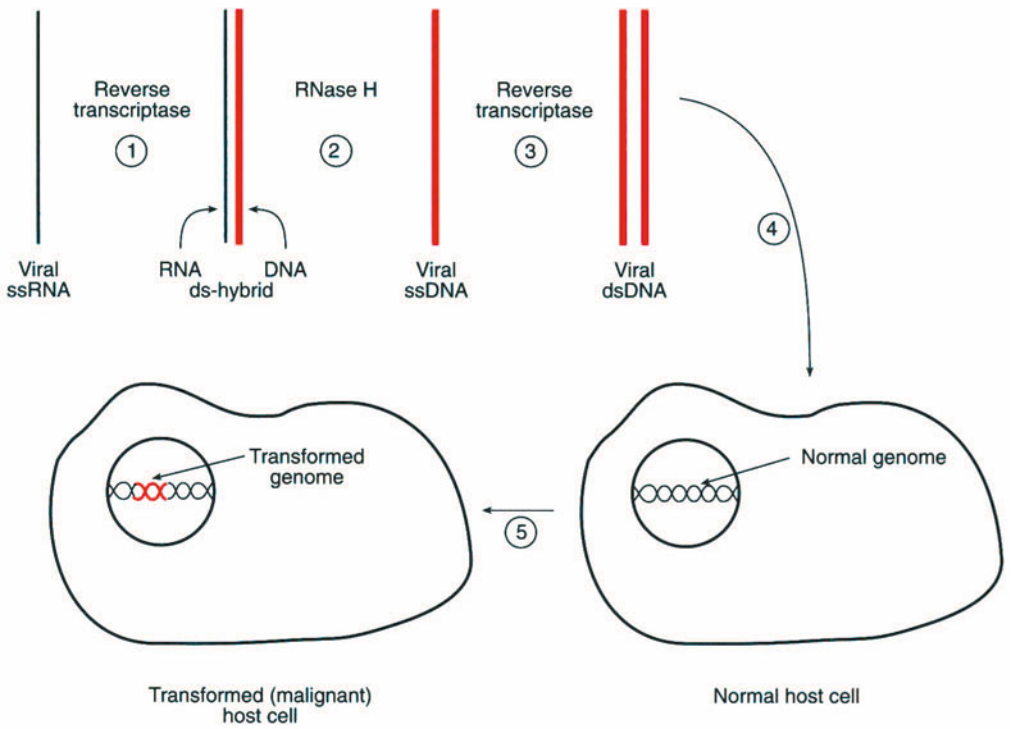


Figure 18.10. Mechanism of action of reverse transcriptase. Numbers refer to the steps described in the text.

opposition, since it went against the then accepted form of the central dogma by postulating a "backward" flow of information, from RNA to DNA. By 1970, however, Temin and David Baltimore had conclusively demonstrated the presence of just such an enzyme in oncogenic RNA viruses. We refer to this *RNA-dependent DNA polymerase* as **reverse transcriptase.** Because of the "backward" flow of information, we call oncogenic RNA viruses **retroviruses.**

Reverse transcriptase can use either DNA or RNA as a template for synthesis of a complementary and antiparallel DNA strand. Like other DNA polymerases, reverse transcriptase synthesizes DNA in the 5′ → 3′ direction. The enzyme requires a specific transfer RNA molecule that serves as primer and provides a free 3′-OH for initiation. Reverse transcriptase catalyzes the synthesis of double-stranded DNA from single-stranded RNA because it catalyzes three kinds of reactions. These reactions correspond to the first three steps of the mechanism depicted in Figure 18.10:

1. It functions as an RNA-dependent DNA polymerase, capable of synthesizing a DNA strand complementary to an RNA strand (first strand synthesis).
2. It possesses a nuclease activity, *ribonuclease H* (RNase H), that specifically catalyzes RNA hydrolysis in RNA–DNA hybrids.
3. It functions as a DNA-dependent DNA polymerase, capable of synthesizing a DNA strand complementary to another DNA strand (second strand synthesis).

Taken together, these reactions convert the single-stranded viral RNA genome to a double-stranded viral DNA genome. DNA synthesis and RNA hydrolysis are coordinated so that RNA is degraded after it has served as a template. The enzyme possesses two molecular domains; one catalyzes DNA synthesis (using either RNA or DNA as a template), and the other catalyzes RNA hydrolysis.

The double-stranded viral DNA formed becomes incorporated into the double-stranded DNA of infected eukaryotic cells (step 4). The host cell's genome is thereby altered, and the cell is converted from a normal to a malignant cell (step 5). After cell division, viral DNA is transcribed to produce viral RNA, and translation of the viral RNA produces viral proteins. Viral RNA and proteins combine to form new viral particles that bind to the cell surface and are later released from it.

18.4.2. Human Immunodeficiency Virus (HIV)

The **human immunodeficiency virus (HIV),** the causative agent of AIDS, represents a retrovirus of great national concern. The HIV virus consists of an RNA–protein core surrounded by a lipid bilayer membrane (Figure 18.11). The membrane contains two glycoproteins, one spanning the membrane (*gp41*), and the other located on its external surface (*gp120*). The viral core consists of a large number of two kinds of proteins, two copies of the RNA genome, and several molecules of reverse transcriptase.

AIDS, or *acquired immunodeficiency syndrome,* constitutes a new disease that emerged in the early 1980s. It was discovered when afflicted individuals died of rare infections because their immune systems were crippled. AIDS is almost invariably fatal because victims can no longer defend themselves against infectious diseases or cancers. Luc Montagnier identified the cause of the disease in 1983, and its name derives from the fact that nearly all cases are acquired via an exchange of body fluids such as blood and semen.

HIV causes lysis of *helper T cells,* which function in both cell-mediated immunity and antibody formation. Upon infection by HIV, the viral membrane fuses with the membrane of the helper T cell, and the viral core is released into the cell's cytosol. The viral RNA becomes uncoated and then copied by reverse transcriptase into DNA sequences that become incorporated into the host-cell DNA.

Development of an effective vaccine against HIV is difficult because the viral reverse transcriptase is very error-prone. The enzyme lacks the proofreading function used by DNA polymerase to correct base-pairing mistakes made during DNA replication. As a result, HIV mutates rapidly. In fact, the HIV mutation rate is about 65 times that of influenza virus. The rapid mutation rate allows the virus to evade the protective effects conferred by any vaccine.

Treatment of AIDS with drugs has likewise only met with limited success. The most widely used drug is AZT (Figure 7.8), a nucleoside that can be incorporated into DNA in place of thymidine. Unfortunately, AZT has very high toxicity for human DNA polymerase. Additionally, HIV's rapid mutation leads to quick appearance of mutated forms of reverse transcriptase that are resistant to AZT.

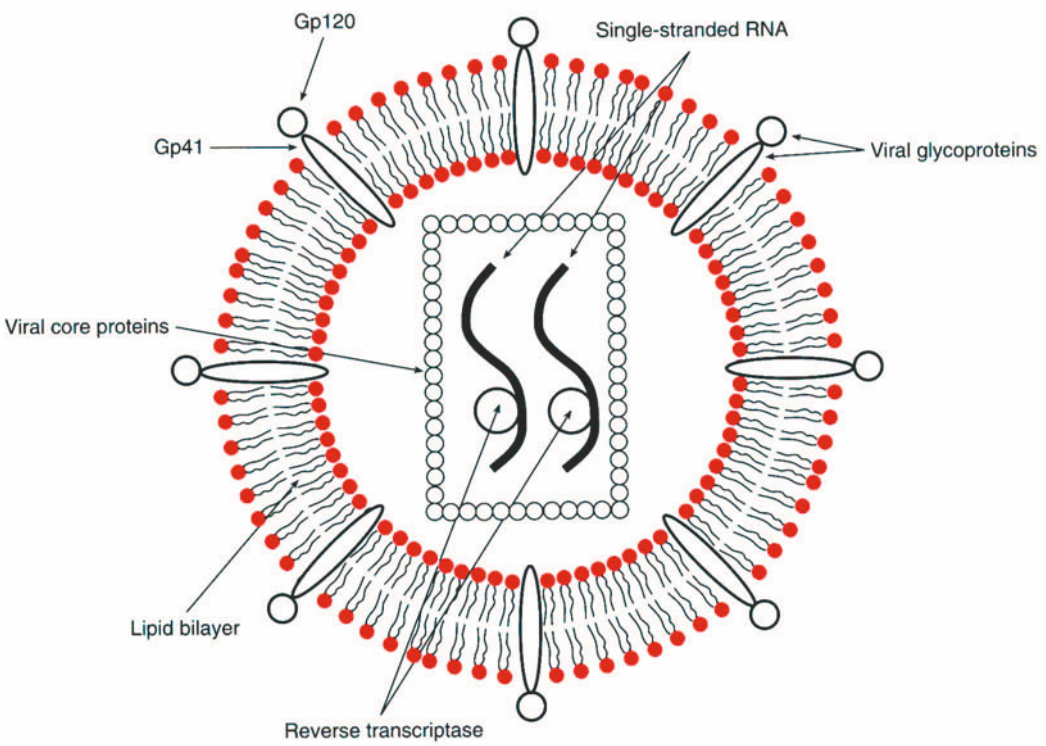


Figure 18.11. Schematic drawing of the human immunodeficiency virus (HIV).

18.4.3. Oncogenes and Proto-Oncogenes

Oncogenes represent genes of retroviruses capable of bringing about malignant transformations of mammalian cells. **Proto-oncogenes** are normal cellular genes that are analogous to viral oncogenes. Some investigators have hypothesized that viral oncogenes may have formed originally from cellular proto-oncogenes. In this scheme, proto-oncogenes mutated to oncogenes during viral infections and were then transferred to the virus particles. Cellular proto-oncogenes can be converted to oncogenes and then cause transformation of normal cells to malignant ones. Oncogene conversion can occur through mutations, such as those induced by certain carcinogens, or by genetic recombination with the genome of a retrovirus. According to current thinking, all cancers result from genetic alterations.

Products of oncogenes function in the nuclear transcription of DNA or in the cytoplasmic regulation of cell growth and differentiation. Oncogene products differ from their normal counterparts in being either modified or expressed in an abnormal fashion. For example, an oncogene may lead to production of a normal cellular protein that has an abnormally high activity or is present in abnormally large amounts. The participation of oncogenes in transcription and cell growth leads to the proliferation of malignant cells.

18.5. GENE REGULATION IN PROKARYOTES

Transcriptional control, regulation of the rate at which genes are transcribed into mRNAs, constitutes the main control mechanism of protein synthesis in both prokaryotes and eukaryotes. *Translational control,* regulation of the rate at which polypeptide chains are translated from mRNAs, constitutes a secondary regulatory mechanism in both types of organisms. In prokaryotes, a major transcriptional control mechanism involves enzyme induction and repression.

18.5.1. Enzyme Induction and Repression

Many enzymes of bacterial cells are present at fixed concentrations, regardless of the metabolic state of the cell; the amount of enzyme per cell stays constant, though the activity of the enzyme may vary. Enzymes of this type, termed **constitutive enzymes,** include those of major catabolic pathways, such as the enzymes of glycolysis. By contrast, other enzymes exist whose amount per cell may vary. These enzymes, named **inducible enzymes,** are normally present in only trace amounts, but their intracellular concentration increases rapidly when specific substances, **inducers,** are added to the medium. An inducer is generally either the substrate of the enzyme or a compound structurally related to the substrate. Adding the inducer results in **enzyme induction:** an increase in transcription of the gene coding for the enzyme, followed by an increase in translation of the corresponding mRNA.

As an example, consider the utilization of carbohydrates by *E. coli.* Normally, this bacterium does not use lactose as long as it has an adequate supply of glucose. Under these conditions, β-*galactosidase,* an inducible enzyme that catalyzes the hydrolysis of lactose to glucose and galactose, is not needed and occurs in only trace amounts of about *five copies* per cell. However, if we place *E. coli* cells in a medium devoid of glucose and containing lactose as the only source of carbon and energy, the cells adapt and begin to use lactose by synthesizing large amounts of β-galactosidase. Within a few minutes, the intracellular concentration of β-galactosidase increases about 1000-fold so that cells contain some *5000 copies* of the enzyme per cell. β-Galactosidase concentration stays at this high level as long as lactose is present in the medium. When lactose has been depleted, the enzyme level returns to its initial low value.

The actual physiological inducer of β-galactosidase is *allolactose,* an isomer of lactose in which galactose is linked β(1→6) to glucose (Figure 18.12A). Allolactose forms from lactose by a side reaction of β-galactosidase.

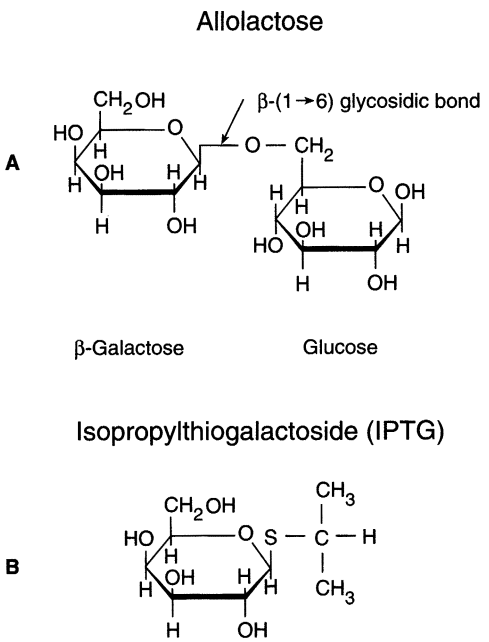


Figure 18.12. The physiological (A) and a gratuitous (B) inducer of β-galactosidase.

Thus, β-galactosidase has two enzymatic activities—it catalyzes both the hydrolysis and the isomerization of lactose.

Some substances, termed **gratuitous inducers,** can serve as inducers even though they are not natural substrates of the enzyme. *Isopropylthiogalactoside (IPTG)* serves as an important gratuitous inducer of β-galactosidase (Figure 18.12B). IPTG has some structural similarity to allolactose and functions as a potent inducer of β-galactosidase, but it is not degraded by the enzyme. Most studies of the β-galactosidase system use IPTG as inducer.

Induction of β-galactosidase by means of either lactose or IPTG is accompanied by induction of two other proteins, β-galactoside permease, a transport protein, and *thiogalactoside transacetylase,* an enzyme. We refer to the induction of a group of related proteins by a single inducer as **coordinate induction.** The permease functions in the transport of lactose and other galactosides from the medium into the cell. The transacetylase catalyzes the transfer of an acetyl group from acetyl CoA to β-galactosides. The enzyme is required for intracellular utilization of lactose, but its precise role is unknown. Some have suggested that it functions in detoxification and excretion of nonmetabolized analogs of β-galactosides.

Enzyme repression constitutes a seemingly opposite effect to that of enzyme induction. In enzyme repression, the concentration of an enzyme (a **repressible enzyme**) decreases upon addition of a specific substance, termed a **co-repressor,** to the medium. A co-repressor is generally either the end product of the enzymatic reaction(s) or a compound structurally related to the end product. To illustrate enzyme repression, consider the bacterial metabolism of amino acids. *E. coli* cells possess the enzymes required for synthesizing all 20 amino acids when they have appropriate carbon and energy sources and ammonium (NH_4^+) as sole nitrogen source. When one amino acid, say, tryptophan, is added to the growth medium, the group of enzymes required for tryptophan biosynthesis becomes repressed and ceases to be produced. Enzyme systems required for biosynthesis of the remaining 19 amino acids continue to be produced. If an entire group of enzymes is repressed by addition of a single end product (as in the case of tryptophan), we refer to the process as **coordinate repression.**

Enzyme induction and repression constitute economical processes, since they provide for synthesis of specific enzymes when the enzymes are needed but stop the synthesis when the enzymes are no longer required. By using these two processes, bacteria can adapt quickly and efficiently to changes in their external environment. As outlined in the following section, enzyme induction and repression represent two manifestations of the same genetic control mechanism.

18.5.2. The Operon Hypothesis

In 1961, François Jacob and Jacques Monod proposed a unifying hypothesis to explain the transcriptional regulation of enzyme induction and repression. Their proposal, called the **operon hypothesis,** has been extensively corroborated by direct biochemical experimentation. According to the operon hypothesis, transcriptional regulation involves three types of DNA segments (Figure 18.13):

1. A **structural gene** that serves as a template and undergoes transcription to produce a messenger RNA molecule. The mRNA is subsequently translated to produce a specific enzyme or other protein.
2. An **operator** that occurs adjacent to the structural gene and controls its transcription.
3. A **regulatory gene** that undergoes transcription to produce a messenger RNA molecule. The mRNA is subsequently translated to produce a **repressor** protein capable of interacting with the operator.

The functional unit of transcription, called an **operon,** consists of **control elements** and one or more structural genes. Control elements comprise a promoter, the site to which RNA polymerase binds, and an operator. A regulatory gene is *not* part of an operon and may be located near the operon or far removed from it.

Enzyme induction and repression are explained by the operon hypothesis as follows. In enzyme induction, *the repressor binds to the operator* (Figure 18.14). The operator/repressor complex stops transcription by blocking RNA polymerase from binding to the promoter. The operator is "turned off," the structural genes controlled by it are not transcribed, and their protein gene products do not form. When an inducer is added, it combines with the free repressor to form an **inducer/repressor complex** that *does not bind to the operator.* As a result, the operator is "turned on," and RNA polymerase can bind to the promoter. The structural genes controlled by the operator are transcribed into mRNA, and translation of the mRNA yields the specific protein gene products.

In enzyme repression, *the repressor does not bind to the operator* and does not prevent RNA polymerase from binding to the promoter. The operator is "turned on," and the structural genes controlled by the operator are transcribed and yield their protein products. When a co-repressor is added, it combines with the repressor to form a **co-repressor/repressor complex** that *does bind to the operator.* Because of this binding, the operator is "turned

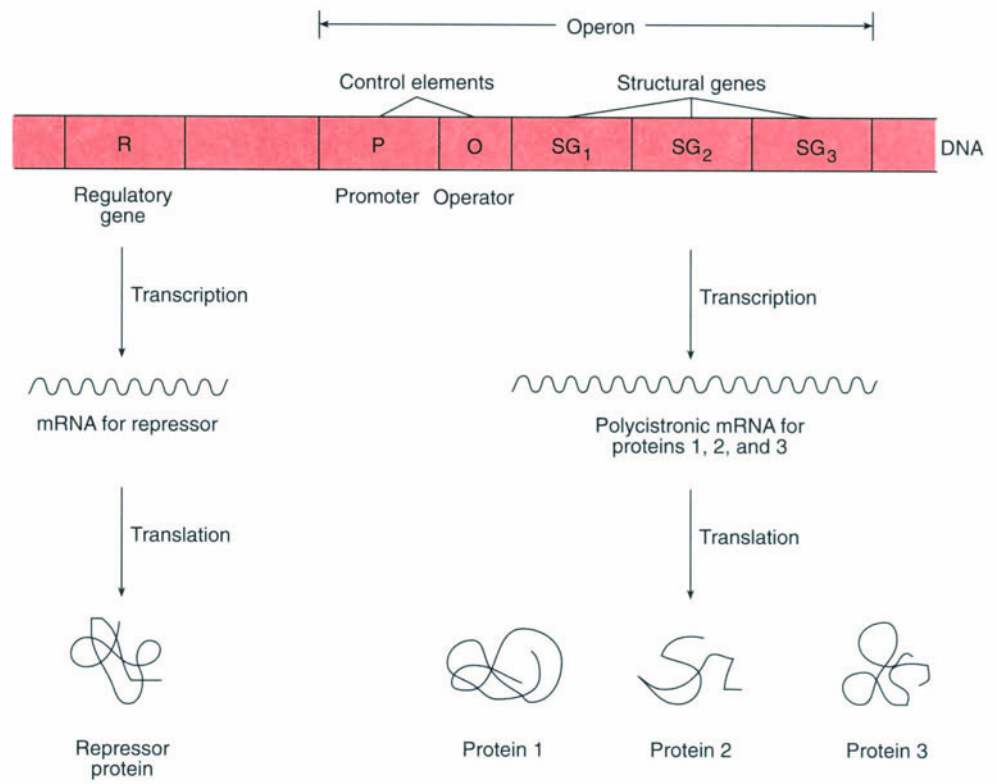


Figure 18.13. Structure of an operon. A regulatory gene, which may be near or far from the promoter, is not part of an operon. In this figure and in Figures 18.14–18.16, structural genes and control elements are not drawn to scale.

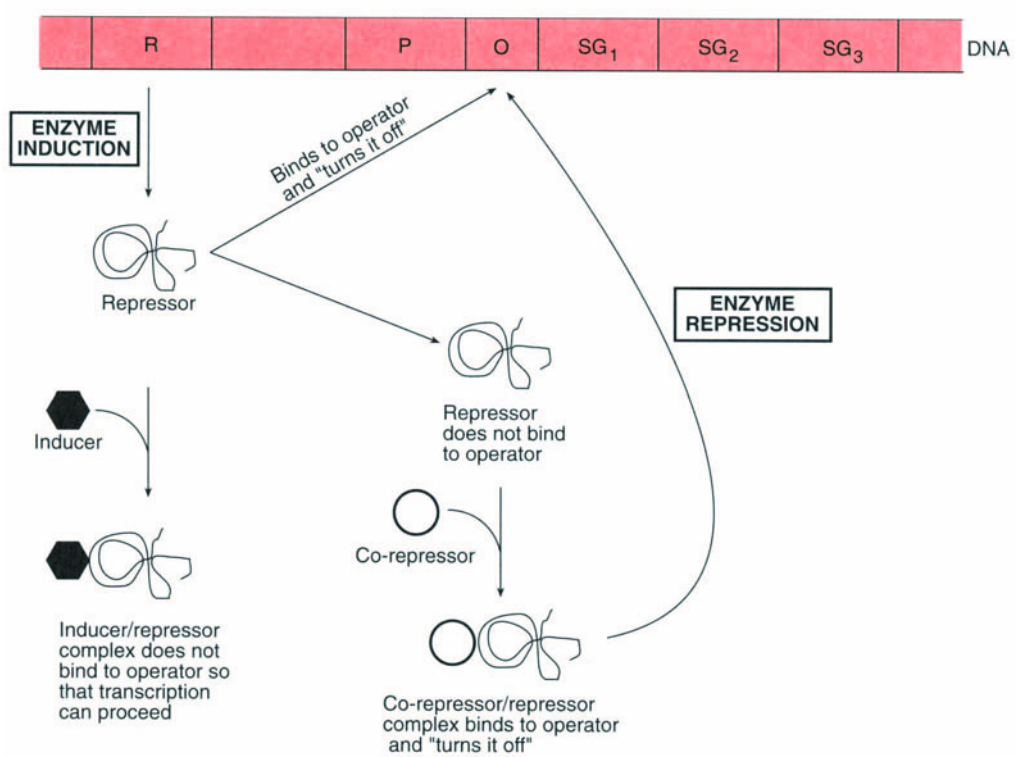


Figure 18.14. Enzyme induction and repression. "Turning off" the operator stops transcription of structural genes.

off," and RNA polymerase can no longer bind to the promoter. The structural genes controlled by the operator are not transcribed and do not yield their protein products.

Control of induction and repression thus involves two general principles:

1. Inducible enzymes are controlled by operons that are normally "turned off" unless induced by an inducer. The *lac* operon that controls the synthesis of β-galactosidase is normally "turned off" but can be "turned on" by adding lactose.
2. Repressible enzymes are controlled by operons that are normally "turned on" unless repressed by a co-repressor. The *trp* operon that controls tryptophan biosynthesis is normally "turned on" but can be "turned off" by adding tryptophan.

18.5.3. The *lac* Operon

The *lac* and *trp* operons of *E. coli* are among the best understood operons. The ***lac* operon** consists of control elements and three structural genes, one each for β-galactosidase, β-galactoside permease, and thiogalactoside transacetylase. The regulatory gene is located next to the promoter, and its product is called the ***lac* repressor** (Figure 18.15). The repressor, isolated in 1966 by Benno Müller-Hill and Walter Gilbert, is a protein (MW = 150,000) that has an extremely high affinity for binding to the operator; one-half of the maximum binding (a term analogous to the K_m of enzymes) is attained at a concentration of $10^{-13}M$. The *lac* repressor is synthesized as inactive monomers (MW = 39,000), which subsequently associate to form an active tetramer. There are usually about 10 tetramers per cell.

The three structural genes of the *lac* operon are transcribed into a single messenger RNA molecule and hence are *coordinately expressed:* all three genes are either transcribed together or not transcribed at all. We call a messenger RNA transcript of two or more genes a **polycistronic mRNA.** Such mRNA serves as a template for translating two or more polypeptide chains specified by adjacent genes in DNA. The polycistronic mRNA of the *lac* operon has a short half-life (about three minutes), permitting rapid alteration in operon expression. Since there is some overlap between the nucleotide sequences of the operator and the promoter, binding of the *lac* repressor to

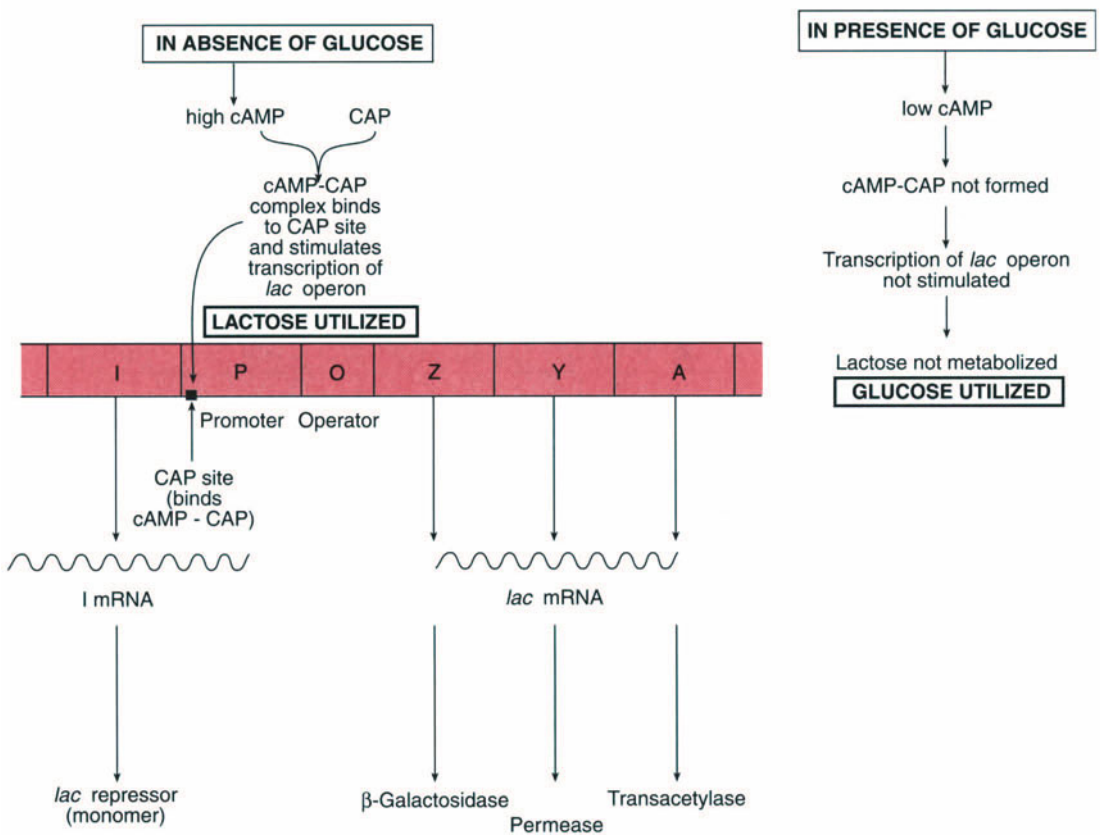


Figure 18.15. The *lac* operon and its control via catabolite repression.

the operator physically prevents RNA polymerase from attaching to the promoter.

The promoter contains recognition sites for two different proteins. One site binds RNA polymerase; the other site binds *cAMP–CAP,* where *cAMP* is *cyclic AMP* and *CAP* is *catabolite activator protein.* Binding of cAMP–CAP stimulates attachment of RNA polymerase to the promoter and increases the *lac* operon's rate of transcription. The cAMP–CAP complex functions in regulating carbohydrate utilization.

If *E. coli* cells are provided with *both* glucose and lactose, glucose interferes with induction of the *lac* operon by lactose, and glucose is used up first. Because this interference can be brought about by either glucose or a catabolite of glucose (such as glucose 6-phosphate or fructose), we term the effect **catabolite repression.**

Catabolite repression results from changes in the concentration of cAMP. Glucose and its catabolites lower the intracellular concentration of cAMP by an as yet unknown mechanism. In the presence of glucose or its catabolites, the cAMP–CAP complex does not form to any great extent, and binding of RNA polymerase is not stimulated. Under these conditions, the *lac* operon is not expressed (it has a low-efficiency promoter), the enzymes for lactose utilization are not produced, and glucose is metabolized instead. In the absence of glucose, the operon is expressed, and lactose is metabolized.

Catabolite repression is not limited to the *lac* operon; glucose and its catabolites interfere with induction of several other operons as well. Some insight into the binding of CAP to DNA has been obtained from determination of its three-dimensional structure. CAP is a dimer of identical subunits (210 amino acids each) and forms a supersecondary structure (Section 3.3). Each subunit contains two α-helical segments joined by a turn. This structural motif, called **helix–turn–helix,** appears to be common to a number of DNA-binding proteins (see below). When CAP binds cAMP, it undergoes a conformational change that increases its affinity for binding to DNA. It appears that the helical segments of the two CAP subunits fit into the major groove of two successive turns of the double helix of B-DNA.

18.5.4. The *trp* Operon

The ***trp* operon** (Figure 18.16) contains five structural genes that code for three enzymes; two of the enzymes are composed of two different subunits. The ***trp* repressor** is a dimer of 107-residue subunits. Normally, there are about 20 copies of *trp* repressor per cell. In contrast to the *lac* repressor, the *trp* repressor is the product of a regulatory gene located far from the operon. The repressor binds to the operator *only* when complexed with tryptophan, which functions as co-repressor. Binding of the co-repressor/repressor complex to the operator blocks the attachment of RNA polymerase to the promoter and results in about a 70-fold decrease in the rate of transcription. An additional decrease in transcription results from a *secondary promoter* and *attenuation.*

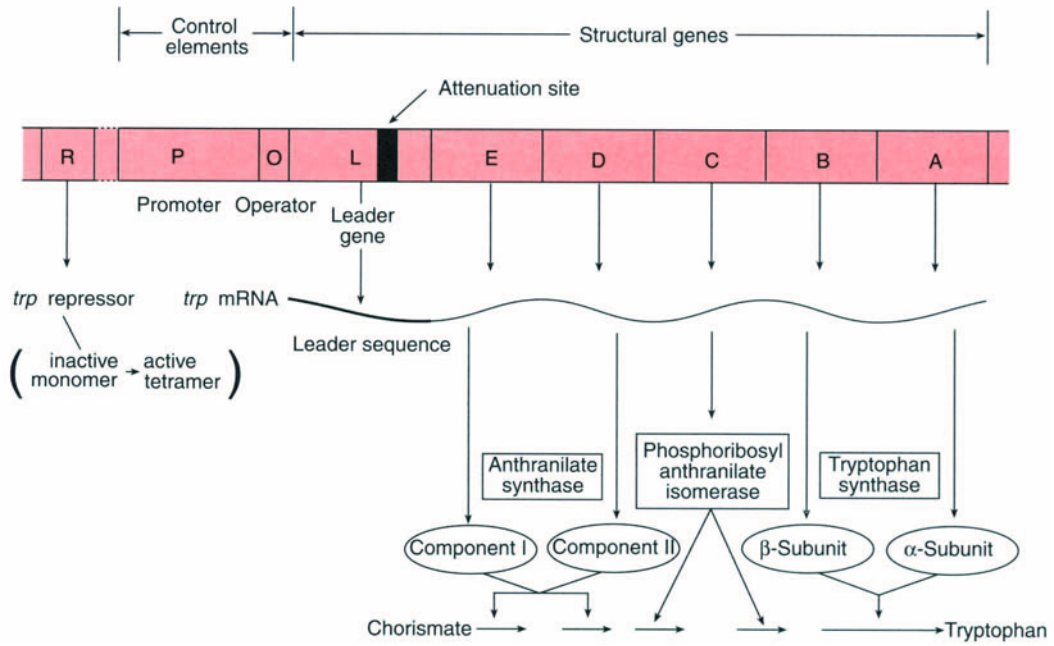


Figure 18.16. Structure of the *trp* operon.

18.5.4A. Secondary Promoter. A secondary promoter is located in the D gene and is not regulated by the repressor. Binding of RNA polymerase to this promoter leads to synthesis of an mRNA transcript of only genes C, B, and A. Hence, the *trp* operon can lead to production of two different polycistronic mRNAs. We do not know the significance of this dual synthesis and the precise role of the secondary promoter in regulation of the operon.

18.5.4B. Attenuation. Attenuation constitutes a mechanism for premature termination of transcription. Charles Yanofsky discovered the phenomenon in 1981. Attenuation serves as a regulatory device for a number of bacterial operons controlling the biosynthesis of amino acids. The *trp* attenuator site is located between the promoter and the start of the first structural gene. It provides a second chance for RNA polymerase to abort transcription if the enzyme was not stopped by the co-repressor/repressor complex and if the cell does not require tryptophan synthesis.

Attenuation is based on the coupling of transcription and translation in prokaryotes and on the occurrence of several consecutive codons for the amino acid end product of the pathway in the leader sequence of mRNA. The **leader sequence** comprises a segment of nucleotides at the 5′-end of mRNA that precedes the AUG *initiation codon.* The segment base-pairs with 16S rRNA and thereby positions the initiation codon at the ribosome for initiation of translation. In the *trp* operon, the template for the leader sequence is located between the operator and the first gene, and the attenuator is located within this template. The leader sequence contains two consecutive codons for tryptophan (Figure 18.17) and can assume different secondary structures (hairpins), depending on the position of the ribosome that follows behind the RNA polymerase. Different secondary structures of the mRNA serve as signals for RNA polymerase.

When the tryptophan concentration is high, cells contain adequate amounts of tryptophanyl-tRNA^{Trp}, and the ribosome proceeds with translation through the section of the leader sequence that contains the consecutive tryptophan codons. The secondary mRNA structure formed serves as a termination signal for RNA polymerase so that transcription stops after synthesis of the short nucleotide transcript of the leader sequence.

When the amount of tryptophan is low, cells contain only small amounts of tryptophanyl-tRNA^{Trp} so that the

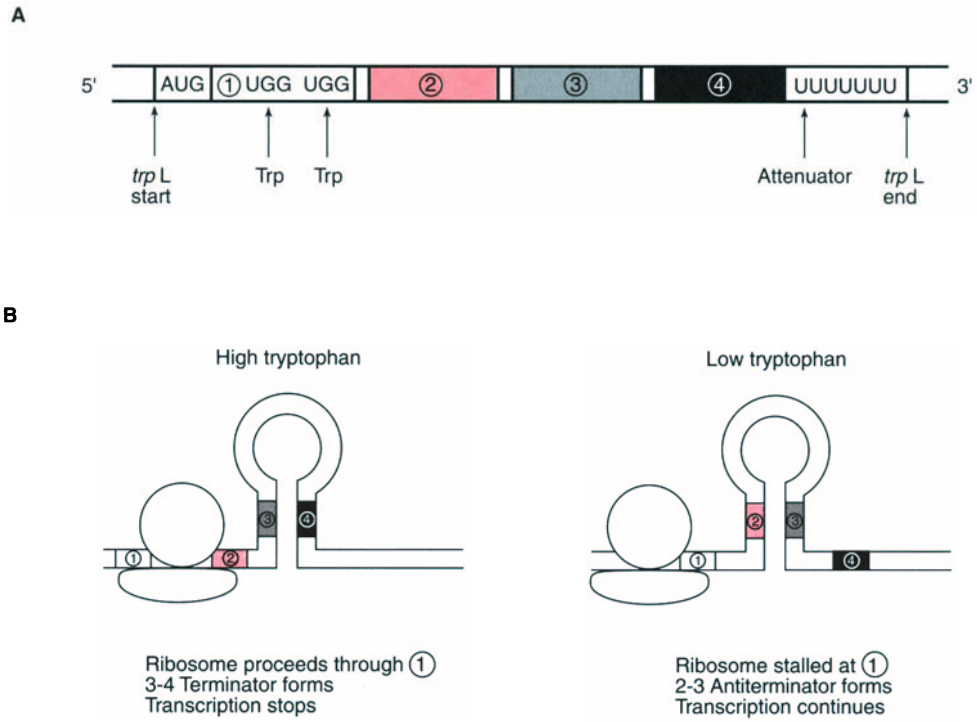


Figure 18.17. Attenuation. (A) Schematic structure of the leader sequence (*trp*L) of *trp* operon mRNA. Complementary base sequences in sections 1–4 allow formation of different hairpin structures: an "antiterminator" forms with sections 2 and 3, and a "terminator" forms with sections 3 and 4. (B) Transcription is terminated at high tryptophan levels but continues at low tryptophan levels. Lengths of the sections and spaces between them are not drawn to scale.

ribosome cannot proceed with translation through the section of the leader sequence that contains the two consecutive tryptophan codons. As a result, the ribosome becomes *stalled.* The mRNA structure formed under these conditions does not constitute a termination signal for RNA polymerase, which continues downstream into the E gene for synthesis of the EDCBA gene products.

In the presence of tryptophan, the attenuator (the control element of attenuation) acts as a rho-independent termination site (see Section 18.2), but in the absence of tryptophan the entire polycistronic mRNA (transcript of five genes) is produced. Researchers have estimated that the attenuator provides about a 10-fold decrease in the rate of transcription. Hence, action of the attenuator, coupled with that of the co-repressor/repressor, results in an approximate 700-fold decrease in the rate of transcription. This decrease is of comparable magnitude to the 1000-fold increase in transcription resulting from induction of the *lac* operon. It appears from these values that transcription can be increased (*lac* operon) and decreased (*trp* operon) to a similar extent.

18.6. GENE REGULATION IN EUKARYOTES

18.6.1. Untranscribed DNA

Regulation of gene expression in eukaryotes is more complex and less well understood than that in prokaryotes. One difference between the two types of organisms involves the fraction of total DNA that is transcribed into mRNA for translation into proteins. Prokaryotes contain a relatively small amount of DNA (see Table 7.3), and most of it codes for proteins. Eukaryotes contain a much larger amount of DNA, but only a small fraction (possibly as little as 2%) codes for proteins. We estimate that the 4 million base pairs of *E. coli* DNA code for some 3000 proteins, and the 2.9 billion base pairs of the human genome code for 30,000–40,000 proteins. On this basis, a human cell has about 700 times as much DNA as an *E. coli* cell but produces only about 10 times as many different proteins. What is the function of the DNA not transcribed into mRNA?

One function of this "other" DNA is to serve as introns between exons of discontinuous genes. Since many eukaryotic genes are discontinuous, the amount of *intron DNA* is appreciable. Expressing discontinuous genes requires excision of introns and splicing of exons in precursor RNAs. Such exon splicing may constitute a regulatory device that permits production of various RNAs by joining different exons of the same precursor RNA.

A special case of complex splicing reactions comprises those that give rise to **antibody diversity,** the enormous number of antibodies an organism can form. Some estimates say that a human can produce more than 10 million different antibodies that bind specific antigens. Since this number greatly exceeds the total number of immunoglobulin genes in the human genome, it is impossible for each antibody to be derived from a unique gene. Instead, the large number of antibodies results from genetic recombination, involving the joining of exons, corresponding to different portions of the antibody molecule.

A second function of "other" DNA is to provide templates for transcription into ribosomal, transfer, and small RNAs. The occurrence of large amounts of DNA not transcribed into mRNA may explain why the amount of DNA per eukaryotic cell does not always correlate with the complexity of the organism (refer to Table 7.3). For example, the lungfish has 35 times as much DNA per cell as a human, despite the fact that it represents a simpler organism.

18.6.2. Repetitive DNA

Eukaryotic DNA differs from prokaryotic DNA not only in its large content of untranscribed DNA but also in its content of many repeated base sequences. Such **repetitive DNA** comprises a heterogeneous group of sequences that vary from a few copies to several million copies per cell.

The most highly repeated sequences consist of about a dozen base pairs. Such DNAs, called **satellite DNAs,** are rich in adenine and thymine. Because DNA density increases with its content of guanine and cytosine, satellite DNAs have lower densities than bulk DNA. Because of this, they can be detected during density gradient centrifugation as a *shoulder* or a *satellite peak,* distinct from the main DNA peak. Satellite DNA typically constitutes about 10–20% of an organism's total DNA. A clue to its function has come from studies on mouse satellite DNA. Researchers found most of it to be located in the centromeres, the attachment sites for the mitotic spindle. Hence, they postulated that satellite DNA plays a role in cell division, where it might help align chromosomes and/or provide binding sites for proteins that attach the spindle fibers.

In addition to satellite DNA, there exist many other classes of repeated DNA sequences. We commonly divide them into two large groups: *short interspersed repeats* typically have a length of 100–500 base pairs, whereas *long interspersed repeats* are several thousand base pairs in length. Some of these repeated DNA sequences occur in introns, and some constitute functional genes, such as those coding for rRNA, tRNA, and histones. In many

cases, the occurrence of repeated DNA sequences seems to play a useful metabolic role by allowing extensive production of specific transcripts.

18.6.3. Transcriptional Controls

Eukaryotic gene regulation, much as that of prokaryotes, occurs primarily at the level of transcription through the action of promoters located upstream of a gene. However, eukaryotic genes are not organized into operons, and there are no operators to regulate promoters. Instead, changes in transcription rates are produced by means of specific base sequences in the DNA (*enhancers* and *silencers*) and specific proteins (*transcription factors*) that bind to these sequences.

Enhancers and **silencers** represent base sequences in eukaryotic DNA that facilitate and inhibit transcription, respectively. These base sequences are located on the same strand as the transcribed gene but can be either upstream or downstream and can be either near or far from the affected gene or its promoter. Enhancers and silencers function by binding specific proteins; we call those binding to enhancers *activators,* and those binding to silencers *repressors.* Binding of these proteins accelerates or decreases transcription as the case may be.

Activators and repressors are examples of **transcription factors**—proteins that bind to DNA and affect the rate of transcription. We know of at least two other types of transcription factors. *Basal factors* function in the initiation of transcription by positioning RNA polymerase at the start of the coding region. *Coactivators* are believed to function as "adapter" molecules that integrate signals from activators—and perhaps repressors—and transmit the integrated signals to the basal factors.

Transcription factors frequently exhibit one of two structural motifs, resulting in overall molecular shapes that fit into segments of double-stranded DNA (Figure 18.18). One motif, called *helix–turn–helix,* was mentioned above in connection with the catabolite activator protein. A second motif, called **zinc finger,** binds to the major groove of DNA and interacts with about five successive base pairs. One variant of this structure consists of a polypeptide segment containing two cysteine and two histidine residues that coordinate a divalent zinc ion.

Since eukaryotic genes are not organized into operons, mRNA is usually **monocistronic,** coding for a single polypeptide chain. Because of this, coordinating the regulation of related genes represents a more difficult problem than in prokaryotes. We do find, however, that eukaryotic genes are frequently organized into a *gene family,* such as that coding for hemoglobin variants during development.

Some hormones exert their effects on transcription. These hormones often are hydrophobic compounds (for instance, steroid hormones) that diffuse across the nuclear membrane, bind to particular receptors in the cell nucleus, and stimulate transcription of specific genes.

18.6.4. Other Regulatory Aspects

Expression of eukaryotic DNA is more complex than that of prokaryotic DNA for a number of reasons. First, packaging eukaryotic DNA in the cell nucleus involves several levels of organization. Nuclear DNA associates specifically with large numbers of protein molecules (histones). The resulting nucleosomes constitute the structural elements of chromatin from which eukaryotic chromosomes are constructed. By contrast, prokaryotic DNA occurs in the nucleoid region of the cytoplasm, where its associa-

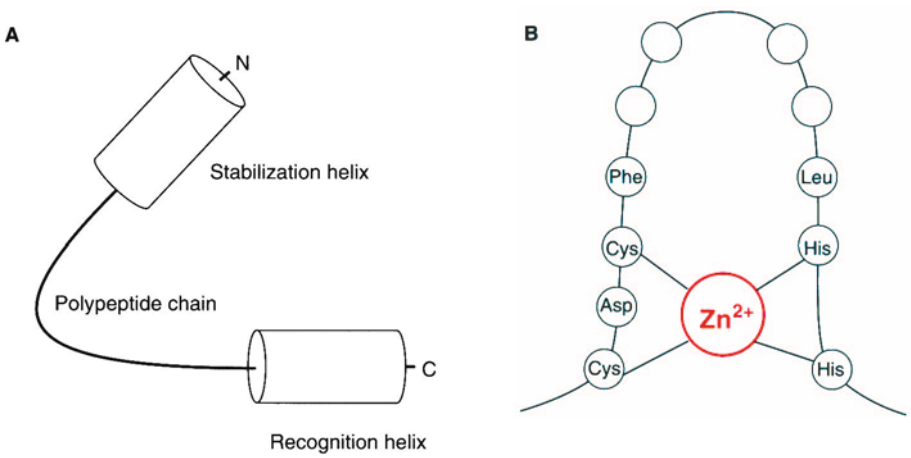


Figure 18.18. Schematic structural elements of some DNA-binding proteins. (A) Helix–turn–helix motif; (B) zinc finger.

tion with proteins and its structural organization are less extensive.

Additionally, eukaryotic genes may sometimes be subject to **gene amplification** whereby specific genes are selectively replicated during development. As an example, the number of copies of the genes coding for rRNA in the toad *Xenopus laevis* increases more than 400-fold during oogenesis. Gene amplification generally leads to a high level of expression of the amplified genes.

Prokaryotes have one major source of genetic information—the DNA of the nucleoid. Some additional information occurs in the form of plasmids. Eukaryotes, however, have two or three major sources of genetic information—the DNA of the nucleus, mitochondria, and chloroplasts. Moreover, the genetic code of mitochondria differs somewhat from the "universal" code of nuclear DNA.

Lastly, eukaryotic cells of multicellular organisms vary in the expression of their genetic material. The entire genetic information for production of a complete eukaryotic organism is contained within the fertilized egg. When the egg develops into a multicellular organism, the process of *cell differentiation* takes place. Cells that are identical at the start develop differently; some become liver cells, others muscle cells, and so on. Cell differentiation represents an extremely complex process.

SUMMARY

RNA polymerase catalyzes transcription of the template strand of DNA into a complementary and antiparallel RNA strand using ribonucleoside triphosphates as substrates. Pyrophosphate is cleaved off from the substrate, and the ribonucleoside monophosphate becomes incorporated. The polymerization is driven by hydrolysis of PP_i. Initiation of transcription in prokaryotes occurs when the sigma subunit binds to the core enzyme of RNA polymerase, forming a holoenzyme that attaches to the promoter region of DNA. Subsequently, RNA polymerase binds ATP (GTP) in a self-priming step. ATP (GTP) provides the 3′-OH for polymerization of the second nucleotide. DNA unwinds locally as transcription proceeds. Termination of prokaryotic transcription involves a specific protein (rho) and a rho-independent mechanism based on the occurrence of specific sites in DNA that lead to formation of hairpin-type structures in the RNA transcript.

Ribosomal and transfer RNAs of both prokaryotes and eukaryotes undergo extensive processing. Primary transcripts of these RNAs consist of large precursor molecules that are processed to form smaller, mature species. Prokaryotic mRNA is not processed and is translated as soon as it forms. Processing of eukaryotic mRNA requires excision of introns and splicing of exons. Further processing involves formation of a methylated cap at the 5′-end and a poly(A) tail at the 3′-end of the mRNA. Some RNAs (ribozymes) have catalytic activities and may be involved in their own processing.

Retroviruses are RNA-containing viruses that carry reverse transcriptase, an enzyme catalyzing DNA synthesis using either an RNA or a DNA template. Upon infection of eukaryotic cells, reverse transcriptase catalyzes synthesis of a double-stranded viral DNA that contains oncogenes and becomes incorporated into host DNA, thereby transforming normal cells to malignant cells.

In prokaryotic cells, constitutive enzymes are present at fixed concentrations. Inducible enzymes are present at low concentrations but form rapidly in response to an added inducer. Synthesis of some enzymes is repressed by adding a co-repressor. Enzyme induction and repression can be explained by the operon hypothesis, according to which the functional unit of transcription is an operon composed of control elements (promoter and operator) and structural genes that code for specific proteins.

In enzyme induction (as in the *lac* operon), the repressor binds to the operator, turns it off, and prevents transcription of the structural genes. An inducer combines with the repressor, prevents its attachment to the operator, and thereby turns the operator on. In enzyme repression (as in the *trp* operon), the co-repressor/repressor complex binds to the operator and turns it off. Gene regulation in eukaryotes is more complex than in prokaryotes.

SELECTED READINGS

Adhyer, S. (ed.), RNA polymerases and associated factors, Part A, *Methods in Enzymology,* Vol. 273, Academic Press, San Diego (1996).

Andrake, M. D., and Skalka, A. M., Retroviral integrase, putting the pieces together, *J. Biol. Chem.* 271:19633–19636 (1996).

Belfort, M., and Perlman, P. S., Mechanisms of intron mobility, *J. Biol. Chem.* 270:30237–30240 (1995).

Brennan, R. G., and Matthews, B. W., The helix–turn–helix DNA binding motif, *J. Biol. Chem.* 264:1903–1906 (1989).

Cooper, G. M., *Oncogenes,* Jones & Bartlett, Boston (1990).

Freed, E. O., and Martin, M. A., The role of human immunodeficiency virus type I envelope glycoproteins in virus infection, *J. Biol. Chem.* 270:23883–23886 (1995).

Friedberg, E. C., Relationships between DNA repair and transcription, *Annu. Rev. Biochem.* 65:15–42 (1996).

Katz, R. A., and Skalka, A. M., The retroviral enzymes, *Annu. Rev. Biochem.* 63:133–174 (1994).

Krämer, A., The structure and function of proteins involved in mammalian pre-mRNA splicing, *Annu. Rev. Biochem.* 65:367–409 (1996).

Laine, R. O., Hutson, R. G., and Kilberg, M. S., Eukaryotic gene expression: Metabolite control by amino acids, *Prog. Nucleic Acid Res. Mol. Biol.* 53:219–248 (1996).

McKnight, S. L., and Yamamoto, K. R. (eds.), *Transcriptional Regulation,* Cold Spring Harbor Laboratory Press, Plainview, New York (1993).

Rhodes, D., and Klug, A., Zinc fingers, *Sci. Am.* 268:56–65 (1993).

Richardson, J. P., Structural organization of transcription termination factor Rho, *J. Biol. Chem.* 271:1251–1254 (1996).

Rosenberg, P. S., Scope of the AIDS epidemic in the United States, *Science* 270:1372–1375 (1995).

Sachs, A., and Wahle, E., Poly(A) tail metabolism and function in eukaryotes, *J. Biol. Chem.* 268:22955–22958 (1993).

Sharp, P. A., Split genes and RNA splicing, *Cell* 77:805–815 (1994).

Yanofsky, C., Transcription attenuation, *J. Biol. Chem.* 263:609–612 (1988).

REVIEW QUESTIONS

A. Define each of the following terms:

Consensus sequence	Spliceosome
Repetitive DNA	Retrovirus
Attenuation	Discontinuous (split) genes
Promoter	Gratuitous inducer
Polycistronic mRNA	Catabolite repression
Transcription factor	Exon splicing

B. Differentiate between the two terms in each of the following pairs:

Sigma subunit/sigma cycle	*lac* operon/*lac* repressor
Enhancer/silencer	Primary transcript/mature RNA
Oncogene/proto-oncogene	Capping/polyadenylation
Operon/operator	Repressor/co-repressor
Constitutive enzyme/inducible enzyme	DNA-dependent RNA polymerase/reverse transcriptase
Structural gene/regulatory gene	Rho-dependent termination/rho-independent termination

C. (1) Describe the prokaryotic mechanisms involved in the following: (a) Termination of transcription; (b) attenuation; (c) catabolite repression.
(2) What are the postulates of the operon hypothesis? How are enzyme induction and repression explained by this hypothesis?
(3) What are the structures and main features of the *lac* and *trp* operons?
(4) Compare and contrast gene regulation in prokaryotes and eukaryotes.
(5) What steps are involved in initiation of transcription in prokaryotes?
(6) How does an RNA-containing virus bring about changes in the DNA of a eukaryotic cell?
(7) Describe the mechanism of exon splicing.

PROBLEMS

18.1. What complexes can be produced with the following and what is the effect of forming each complex? (a) A repressor in enzyme induction and in enzyme repression; (b) cAMP in the operation of the *lac* operon.

18.2. How many base pairs are there in the following hypothetical DNA sites? (a) From −13 to +27; (b) from +1 to +59; (c) from −42 to +18.

18.3.* How many nucleotides are in a double-stranded bacterial DNA if we know that 75% of the DNA codes for a total of 2100 different proteins that have an average molecular weight of 30,000? Assume that the molecular weight of an amino acid residue is 100.

18.4. What is the mass, in grams, of the cellular DNA in the previous problem, if the molecular weight of a nucleotide residue is taken as 300?

18.5.* What is the total number of nucleotides of the mRNAs that code for all of the proteins in Problem 18.3?

18.6. What are the base sequence and base composition of the double-stranded DNA segment that contains the structural gene for the heptapeptide Met-Met-Trp-Trp-Met-Trp-Trp?

18.7.* You receive a sample of purified 16S ribosomal RNA from *E. coli*. What experiment could you perform on this material that should provide support for the notion that the 16S rRNA is transcribed from one of the DNA strands and not from both?

18.8. Would it be correct to refer to an inducer of the *lac* operon as either an *antirepressor* or a *de-repressor?* Could the same terms be used for the co-repressor of the *trp* operon as well?

18.9. Since all three structural genes of the *lac* operon are transcribed as a unit, can you conclude that there exist equal numbers of copies of the corresponding three gene products (β-galactosidase, β-galactoside permease, and thiogalactoside transacetylase) in the cell? Explain your answer.

18.10. Why can transcription in eukaryotes *not* be regulated by an attenuation mechanism?

18.11.* Consider two identical aliquots of a growing culture of *E. coli*. To one aliquot you add a small amount of tryptophan, and to the other you add an identical amount (in micromoles) of histidine. The leader sequence of the *his* operon of *E. coli* contains seven consecutive codons for histidine. Which culture might you expect to show a greater degree of attenuation and why?

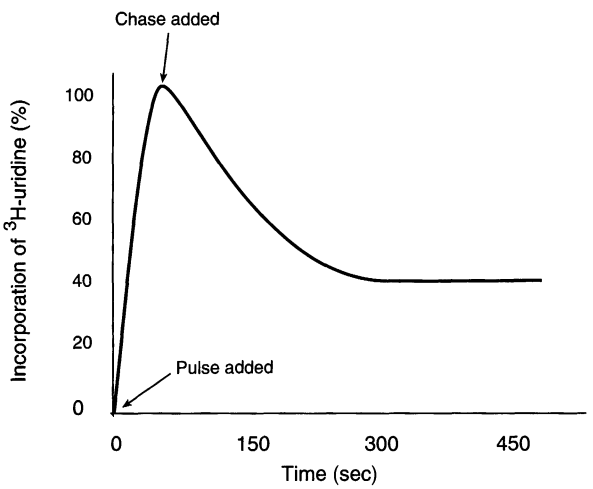
18.12. The sense strand of a stuctural gene has the base sequence 5′-AGCCATTCGAGGTGTTCACGTAAA-3′. What is the base sequence of the mRNA?

18.13. What is the amino acid sequence of the peptide coded for by the structural gene of the previous problem?

18.14. You fragment *E. coli* DNA, heat-denature it, and then hybridize it with bulk mRNA from the same organism. You find that a maximum of 50% of the DNA forms hybrids under optimum conditions. How do you explain this result?

18.15. The maximum *in vivo* rate of transcription is about 45 nucleotides per second in *E. coli*. On this basis, calculate the length of time required for transcription of the gene for: (a) the inactive monomer of the *lac* repressor (360 amino acids); (b) a protein composed of 800 amino acid residues.

18.16.* A *pulse–chase experiment* is one in which we expose a system briefly to a radioactive substance (the pulse), followed by a large dose of the same but unlabeled substance (the chase), which stops incorporation of the radioactive label. A student exposes a growing culture of *E. coli* to a pulse of [^{3}H]uridine, which is incorporated into RNA. After 60 seconds, the student adds a large amount of unlabeled uridine together with rifampicin. Rifampicin inhibits initiation of RNA synthesis but and has no effect on chain elongation. This pulse–chase experiment yields the plot shown below. Keeping in mind that mRNA has a shorter half-life than other RNAs, calculate the approximate fraction of incorporated label taken up by mRNA. What is the approximate half-life of this labeled mRNA?



18.17. Cordycepin (3′-deoxyadenosine) is an antibiotic. Its nucleotide analog, cordycepin 5′-triphosphate, is incorporated into RNA by RNA polymerase. Once incorporated into the RNA, cordycepin 5′-triphosphate inhibits any further transcription. Why is this so? What can you conclude about the direction of growth of the newly synthesized RNA strand?

18.18. What would be the expected effect of a mutation in DNA that led to production of a sigma subunit capable of binding by itself (in addition to binding as the holoenzyme) to the promoter site during the initiation stage of transcription?

18.19.* What must be the reason that the core enzyme, rather than the holoenzyme, continues with transcription after the initiation stage? In view of your answer, what effect would you predict for a mutation that led to the sigma subunit *not* dissociating from the core enzyme?

18.20. What is the likely effect of a mutation that results in a deletion of the regulatory gene from: (a) the *lac* operon; (b) the *trp* operon?

18.21.* A researcher suggests that (a) the poly(A) tail of eukaryotic mRNA is shortened by one A residue each time the mRNA is used for transcription of a protein and (b) as soon as the tail reaches a minimal length of five residues, the shortened tail serves as a signal to initiate degradation of the mRNA. How could an analysis of bulk mRNA provide support for, or argue against, this hypothesis?

18.22. An oncogene of the single-stranded RNA of a retrovirus has the following base composition (mol %): A, 15; U, 25; G, 25; C, 35. What is the base composition of the double-stranded DNA segment, corresponding to this oncogene, that can be transferred from the virus into the host eukaryotic cell?

18.23. What would you expect to be the effect on lactose utilization of a mutation in the *lac* operon that resulted in formation of (a) inactive β-galactoside permease or (b) β-galactoside permease of increased activity?

18.24. What would you predict to be the effect on the biosynthesis of tryptophan of a mutation in the *trp* operon that resulted in formation of an inactive product from (a) gene E or (b) gene A?

18.25. The base composition of a sense strand of DNA is (mol %): A, 25; T, 19; G, 28; C, 28. The template strand of the DNA is transcribed completely. What is the base composition of the newly made RNA strand?

18.26. What accounts for the fact that prokaryotic mRNA constitutes only about 5% of cellular RNA but represents about 50% of the cellular RNA synthesized at any given time?

18.27. The holoenzyme of RNA polymerase forms a tight complex with *E. coli* DNA. A researcher subjects this complex to digestion with deoxyribonuclease and finds that a segment of the bacterial double-stranded DNA (approximately −20 to +20) remains intact. How can this be?

Translation—The Synthesis of Protein

19

In protein biosynthesis, or *translation,* a sequence of mRNA codons directs the polymerization of amino acids into a polypeptide chain. The process is appropriately called translation, since it involves a transfer of genetic information by means of two different "languages" using nucleotides and amino acids.

Most of our knowledge of protein synthesis has been obtained over the last 40 years, following three key discoveries made in the 1950s: (1) identification of ribosomes as the sites of protein synthesis; (2) characterization of transfer RNA as the adapter that transports amino acids to the ribosomes; and (3) demonstration of amino acid activation whereby an amino acid is converted first to an aminoacyl adenylate and subsequently to an aminoacyl-tRNA.

By all measures, protein synthesis constitutes the most complex biosynthetic mechanism known, and its elucidation represents a major triumph of modern biochemistry. Protein synthesis requires the interplay of several hundred different molecules, about 200 in prokaryotes and 300 in eukaryotes (Table 19.1). Despite this enormous complexity, a polypeptide chain grows rapidly. In *E. coli,* protein synthesis accomplishes the polymerization of about 20 amino acids per second. Thus, about 20 codons, or 60 nucleotides, of mRNA must "pass through" a ribosome per second. In rabbit reticulocytes, hemoglobin synthesis proceeds at a rate of about one amino acid per second.

19.1. PRINCIPLES OF PROTEIN BIOSYNTHESIS

19.1.1. Overview of Translation

Synthesis of a protein, like that of any other polymer, requires *initiation, elongation,* and *termination* of the chain. Preceding chain initiation, each amino acid becomes *activated* in a two-step process. After the amino acid is linked to AMP, forming an **aminoacyl adenylate,** it is transferred to a specific tRNA, forming an **aminoacyl-tRNA.** The activated amino acid is transported to the ribosomes for incorporation into protein. Once synthesized, the polypeptide chains undergo modifications through *posttranslational processing.* Thus, we can divide translation into five stages: (1) amino acid activation; (2) initiation; (3) elongation; (4) termination; and (5) posttranslational processing.

Our discussion of protein synthesis will focus on

Table 19.1. Number of Molecules Required for Translation

Type	Prokaryotes	Eukaryotes
Ribosomal proteins	55	82
Ribosomal RNA	3	4
Amino acid-activating enzymes (minimum number)	20	20
Cytosolic protein factors	9	12
Transfer RNAs (probable approximate number)	60	100
Posttranslational processing enzymes (probable approximate number)	50	100
Total	197	318

prokaryotes, specifically on the translational system of *E. coli* (Table 19.2). Protein synthesis in eukaryotes is similar to that in prokaryotes but more complex and not as well understood.

19.1.2. Ribosomes—Sites of Protein Synthesis

19.1.2A. Alignment of Ribosome and mRNA. Recall that the ribosome consists of two unequal subunits, designated by their sedimentation coefficients: 30S and 50S particles in prokaryotes; 40S and 60S particles in eukaryotes. Both subunits have a roughly globular structure, and each contains one or more RNA molecules and a large number of protein molecules (see Table 16.2).

The 30S subunit has a cleft, or tracking groove, that holds an mRNA segment having a length of about 25 codons or 75 nucleotides. When the two subunits associate, the cleft is bounded by the 50S subunit (Figure 19.1).

As mRNA translation proceeds, the ribosome and mRNA must continuously realign. A new mRNA segment, containing codons that have not yet been translated, must replace the previously translated segment in the cleft. According to the favored mechanism for this realignment, the mRNA is "pulled through" the ribosome much as a thread is pulled through the eye of a needle.

19.1.2B. Binding Sites on the Ribosome. In addition to binding mRNA, ribosomes bind tRNA. Each ribosome has two tRNA binding sites (Figure 19.2), an **A site** (**aminoacyl** or **acceptor site**) and a **P site** (**peptidyl** or **donor site**). The A site binds the incoming aminoacyl-tRNA. The P site binds the growing polypeptide chain when the next amino acid arrives at the ribosome. Subsequent transfer of the peptidyl group from the P to the A site results in peptide bond formation with the incoming amino acid.

Thus, each tRNA molecule can bind both its specif-

Table 19.2. Components Required for Protein Synthesis in *E. coli*

Stage	Components
1. Amino acid activation	Amino acids tRNAs Aminoacyl-tRNA synthetases ATP, Mg^{2+}
2. Initiation	$fMet\text{-}tRNA^{fMet}$ Initiation codon (AUG) of mRNA Shine–Dalgarno sequence of mRNA 30S ribosomal subunit 50S ribosomal subunit Initiation factors (IF-1, IF-2, IF-3) GTP, Mg^{2+}
3. Elongation	Aminoacyl-tRNAs Amino acid codons of mRNA 70S ribosome Elongation factors (EF-Tu, EF-Ts, EF-G) GTP, Mg^{2+}
4. Termination	Termination codons (UAA, UAG, UGA) of mRNA 70S ribosome Release factors (RF-1, RF-2, RF-3) GTP, Mg^{2+}
5. Posttranslational processing	Various proteins, enzymes, and other biomolecules

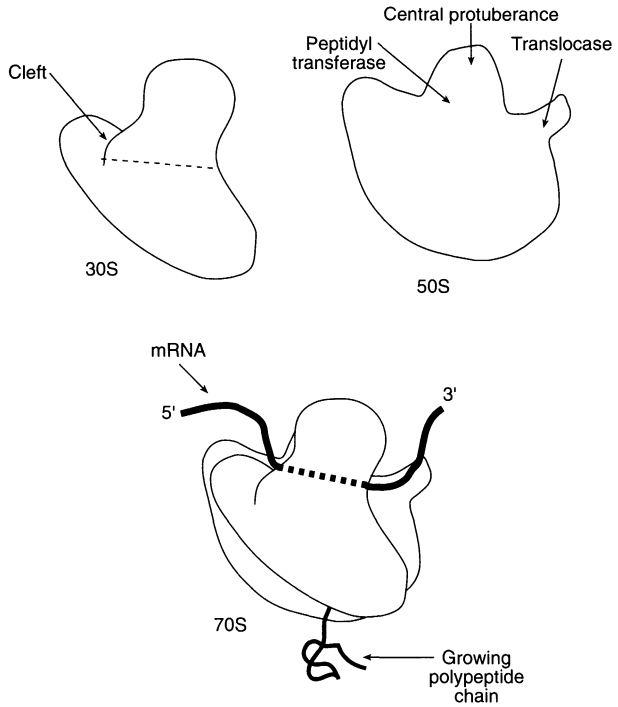


Figure 19.1. Structural features of prokaryotic ribosomes. [Adapted, with permission, from J. A. Lake, *Annu. Rev. Biochem.* 54:507–530 (1985). Copyright © 1985 by Annual Reviews, Inc.]

ic amino acid and the growing polypeptide chain of which the amino acid forms the C-terminus (Figure 19.3). A transfer RNA with an attached polypeptide chain constitutes a **peptidyl-tRNA.** As we will see, the P site binds essentially only peptidyl-tRNA; it binds one aminoacyl-tRNA, that corresponding to the initial, N-terminal amino acid. The A site can have either an aminoacyl-tRNA or a peptidyl-tRNA bound to it.

Binding of tRNA to ribosomes is *codon-directed at all times.* This means that aminoacyl-tRNA becomes linked to mRNA by hydrogen bonding of base pairs between the anticodon in tRNA and the codon in mRNA (Figure 19.4). The same base pairing exists when the tRNA carries the growing polypeptide chain of which the specific amino acid represents the C-terminus. As always, bonding between any two nucleotide strands is complementary and antiparallel.

19.1.2C. Ribosome Cycle. Initiation of protein synthesis involves the smaller ribosomal subunit, which becomes linked to the larger one to form a complete ribosomal particle. As the mRNA is "pulled through," the complete ribosome effectively "moves along" the mRNA. Once synthesis of the polypeptide chain has been terminated, the ribosome dissociates from the mRNA and separates into its two component subunits. We call this set of reactions the **ribosome cycle** (Figure 19.5).

Occurrence of the ribosome cycle was demonstrated by R. O. R. Kaempfer, M. Meselson, and H. J. Raskas in 1968. These investigators grew *E. coli* cells in a medium containing heavy isotopes (^{15}N, ^{13}C, ^{2}H) and then transferred the cells to a medium containing the corresponding light isotopes (^{14}N, ^{12}C, ^{1}H). They isolated and analyzed ribosomes by density gradient centrifugation and found three classes—"heavy," "light," and "hybrid" ribosomes. The first two types had both subunits labeled

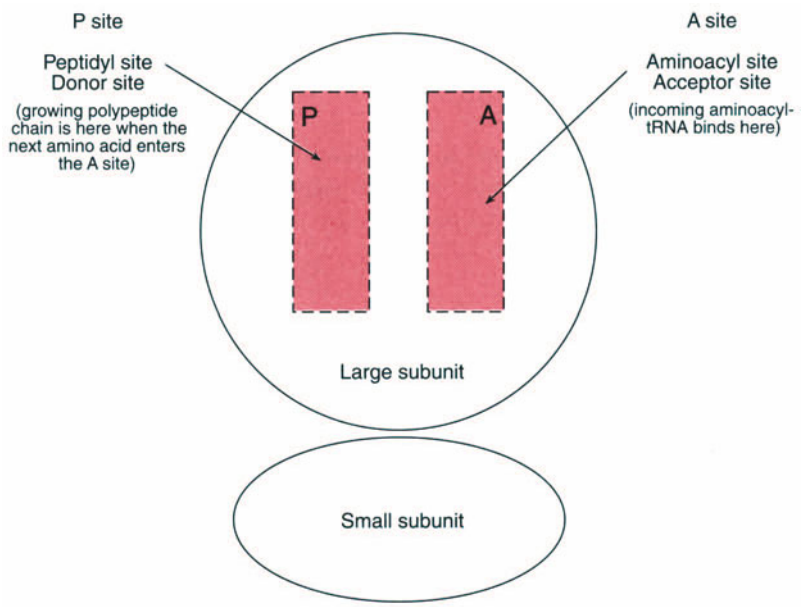


Figure 19.2. Binding sites on the ribosome.

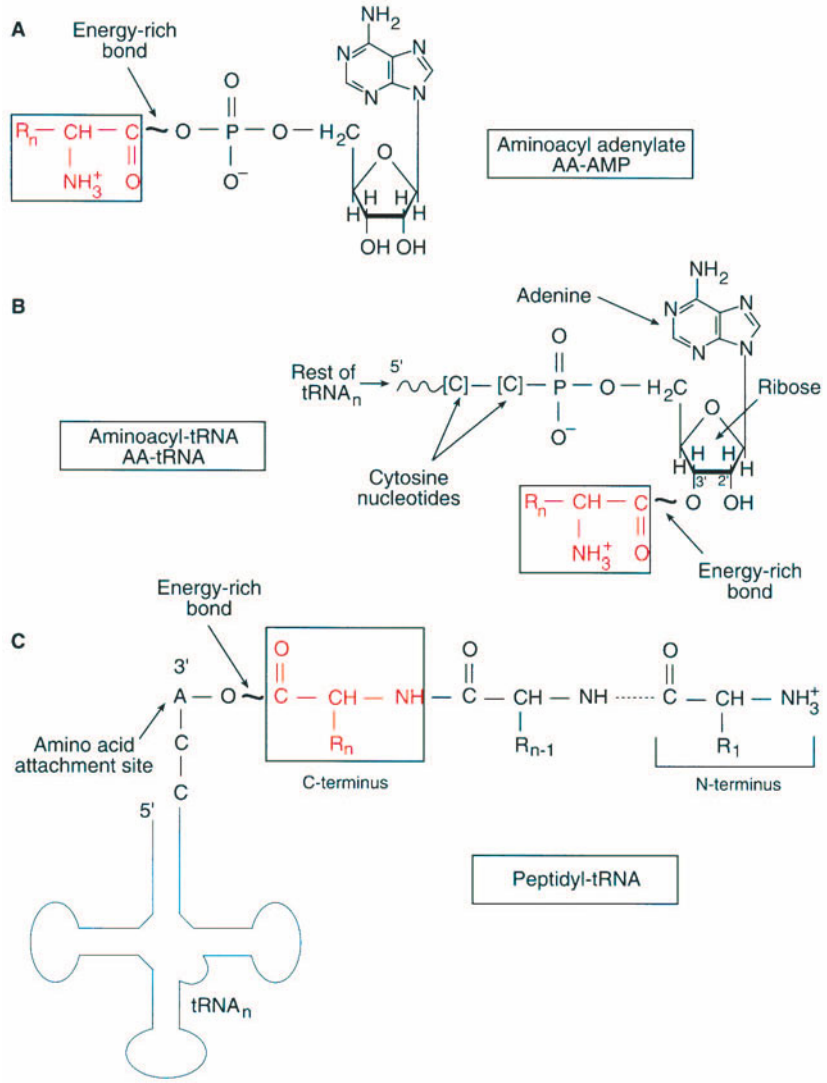


Figure 19.3. Structures of aminoacyl adenylate (A), aminoacyl-tRNA (B), and peptidyl-tRNA (C).

with either heavy or light isotopes; hybrid ribosomes consisted of one heavy and one light subunit. The occurrence of hybrid ribosomes indicated that both heavy and light ribosomes dissociated and then recombined with different subunits.

After the first ribosome becomes engaged in polypeptide chain synthesis and moves down some distance on the mRNA, a second ribosome attaches and begins synthesizing a second, identical polypeptide chain. This process repeats so that each mRNA molecule has many ribosomes attached to it. Each ribosome carries an identical polypeptide chain but in different stages of completion (Figure 19.6). We call this aggregate of mRNA and ribosomes a *polysome* or *polyribosome*. A polysome has a maximum density of about one ribosome per 80 mRNA nucleotides. In some ways, polysomes resemble transcriptional aggregates in which multiple RNA polymerase molecules attach to a strand of DNA (see Figure 18.4).

19.1.3. Directionality in Translation

19.1.3A. Synthesis of Polypeptide Chains.

The polypeptide chain is synthesized from the N- to the C-terminus: synthesis begins with the N-terminal amino acid and ends with the C-terminal amino acid. In 1961, Howard Dintzis determined the directionality of polypeptide chain growth for hemoglobin of *reticulocytes*—immature red blood cells that actively synthesize hemoglobin. He exposed cells to a *pulse* of [^{3}H]leucine for varying lengths of time. After isolating the hemoglobin, he sepa-

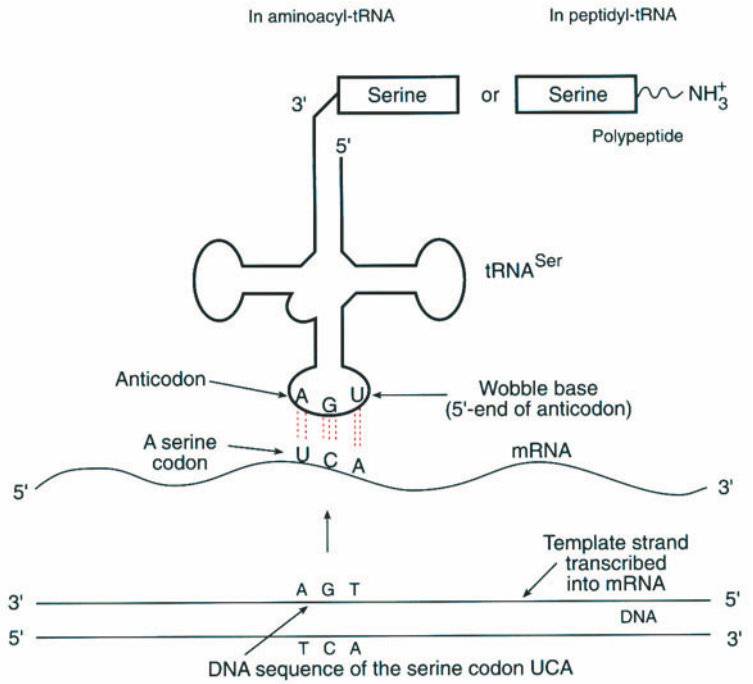


Figure 19.4. Codon–anticodon binding. To show the complementary and antiparallel binding, the tRNA has been drawn with its 3′-end on the left.

rated the two chains and analyzed completed α-chains for distribution of [^{3}H]leucine (Figure 19.7). The results indicated that the polypeptide chain grew from the N- to the C-terminus.

19.1.3B. Translation of Messenger RNA.

Messenger RNA is translated in the 5′ → 3′ direction: the 5′-end is translated first, the 3′-end last. The direction of "reading" mRNA was determined by the use of synthetic mRNAs in a Nirenberg-type, cell-free amino acid-incorporating system (see Section 16.3). For example, use of the polyribonucleotide

$$5'\ \mathrm{AAA}-(\mathrm{AAA})_n-\mathrm{AAC}\ 3'$$

as synthetic mRNA led to formation of the polypeptide

$$\mathrm{H_2N}-\mathrm{Lys}-(\mathrm{Lys})_n-\mathrm{Asn}-\mathrm{COOH}$$

proving that mRNA is "read" in the 5′ → 3′ direction.

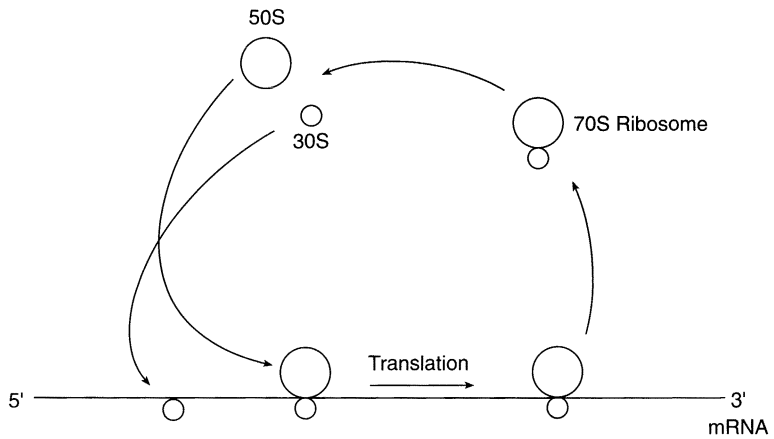


Figure 19.5. The ribosome cycle. The two subunits associate at the start of translation and dissociate at its termination.

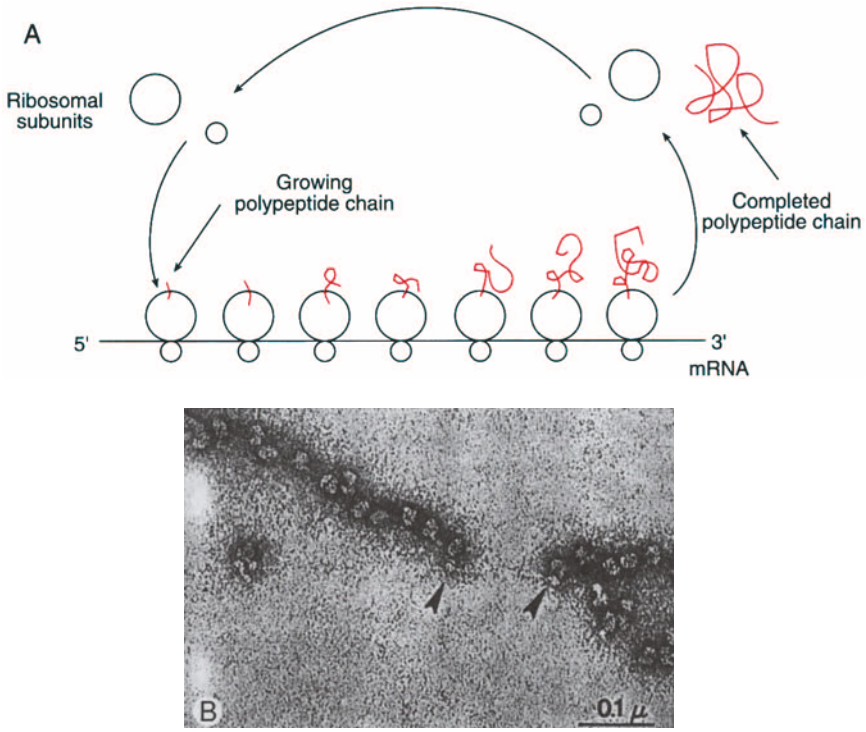


Figure 19.6. Polysomes. (A) Polysome formation as part of the ribosome cycle. (B) Electron micrograph of two polysomes from *E. coli*. Arrows point to attached molecules of RNA polymerase (see also Figure 19.8). [Reprinted, with permission, from B. A. Hamkalo, and O. L. Miller, *Annu. Rev. Biochem.* 42:379–396 (1973). Copyright © 1973 by Annual Reviews, Inc.]

19.1.3C. Simultaneous Transcription and Translation. Recall that synthesis of mRNA also proceeds in the 5′ → 3′ direction. Because both mRNA synthesis and its reading proceed in the same direction, *transcription and translation can proceed simultaneously* (Figure 19.8). As soon as mRNA transcription from DNA begins, ribosomes can atttach to the mRNA and translate it into a polypeptide chain.

Simultaneous transcription and translation is possible (and does occur) only in prokaryotes, where the two

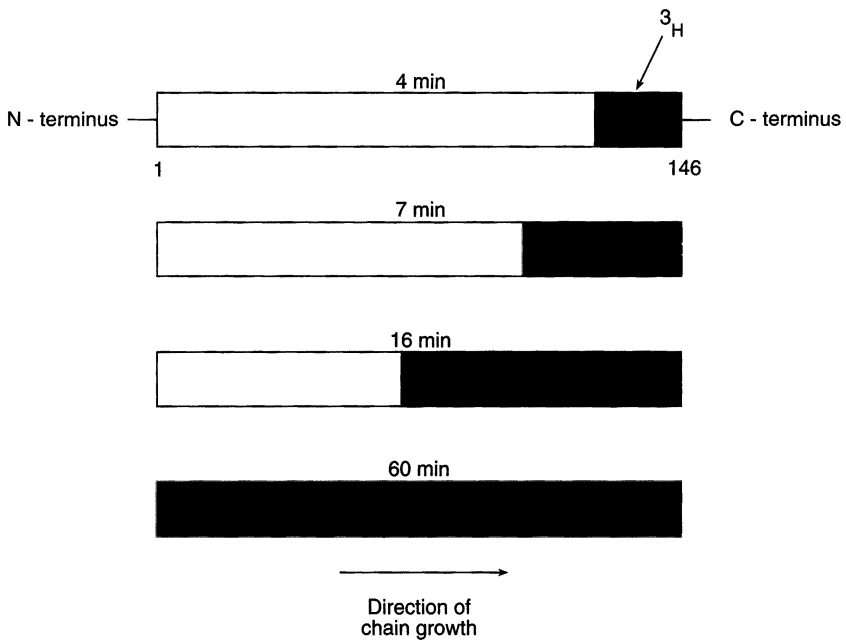


Figure 19.7. Distribution of radioactive label in the α-chains of hemoglobin synthesized by reticulocytes. With a short exposure to the radioactive label, only a small segment at the C-terminus became labeled, indicating that the most recently added amino acids were those at the C-terminus. Thus, the polypeptide chain grew from the N- to the C-terminus.

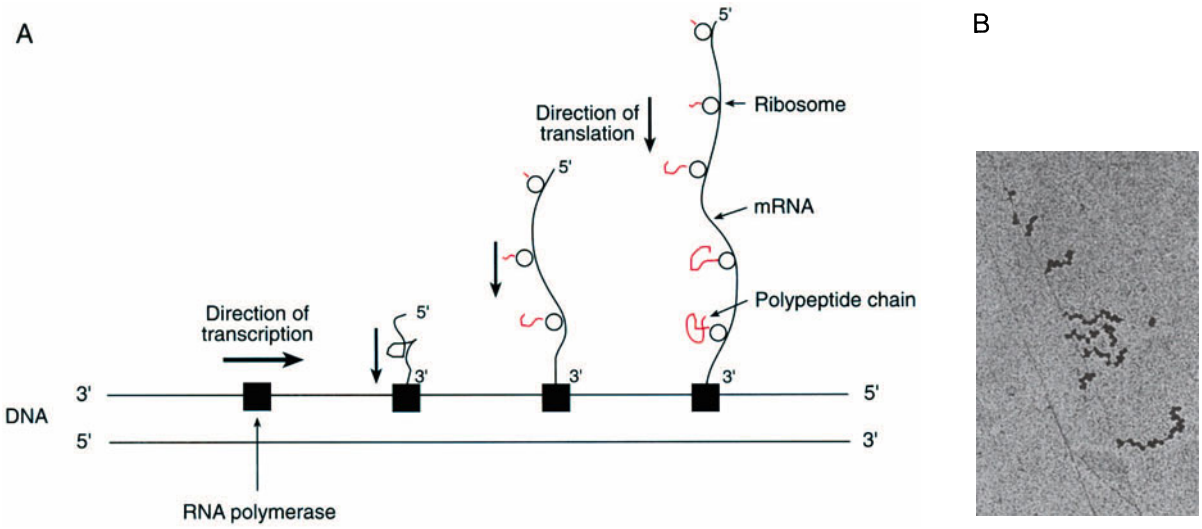


Figure 19.8. Coupled transcription and translation in prokaryotes. As mRNA is synthesized, ribosomes attach to it and begin its translation. (A) Schematic drawing. (B) Electron micrograph from *E. coli.* [Reprinted, with permission, from O. L. Miller, B. A. Hamkalo, and C. A. Thomas, Jr., *Science* 169:392–395 (1970). Copyright © 1970 American Association for the Advancement of Science.]

processes occur in the same cytoplasmic compartment. In eukaryotes, simultaneous transcription and translation is impossible because the processes occur in two different compartments (nucleus and cytoplasm), separated by a membrane.

Prokaryotic ribosomes occur mostly free in the cytoplasm, but a few are bound to the cell membrane. In eukaryotes, ribosomes occur both as free ribosomes and as ribosomes tightly bound to the membrane of the endoplasmic reticulum. **Free ribosomes** synthesize proteins that function in the cytoplasm or that are targeted for the nucleus, mitochondria, and other cellular organelles. **Membrane-bound ribosomes** synthesize proteins destined to be secreted by the cell or targeted for the cell membrane and lysosomes.

19.2. STAGE 1: AMINO ACID ACTIVATION

19.2.1. Formation of Aminoacyl Adenylate

Because the biochemical free energy change for synthesis of peptide bonds is of the order of $\Delta G^{\circ\prime} = +13$ kJ mol^{-1}, peptide bond formation cannot be accomplished by simple polymerization of amino acids in aqueous solution. Instead, amino acids undergo conversion to a more reactive form prior to polymerization. This process, called **amino acid activation,** takes place in the cytosol, not on the ribosomes, and is catalyzed by **aminoacyl-tRNA synthetases.** At least one specific synthetase exists for each amino acid (glycyl-tRNA synthetase, histidyl-tRNA synthetase, etc.), resulting in a minimum of 20 different synthetases per cell. Aminoacyl-tRNA synthetases differ in the number and size of their subunits as well as in their amino acid sequences. Synthetases isolated from different organisms and specific for the same amino acid exhibit considerable *sequence homology.* However, synthetases from one organism but specific for different amino acids show little sequence homology.

Amino acid activation occurs in two steps (Figure 19.9), both catalyzed by aminoacyl-tRNA synthetase, which requires magnesium ions as a cofactor. In the first step, an amino acid (AA) becomes covalently linked to an AMP residue derived from ATP, forming an *aminoacyl adenylate (AA-AMP).* Pyrophosphatase catalyzes hydrolysis of the *pyrophosphate* (PP_i), cleaved out during the reaction, to inorganic phosphate:

$$\mathrm{AA + ATP} \xrightleftharpoons{\text{aminoacyl-tRNA synthetase}} \mathrm{AA\text{-}AMP + PP_i}$$

$$\mathrm{PP_i} \xrightarrow{\text{pyrophosphatase}} \mathrm{2P_i}$$

Aminoacyl adenylates constitute energy-rich compounds, belonging to the class of acid anhydrides (see Section 9.2). Their mixed anhydride structure consists of a C=O group linked to a P=O group via an oxygen atom. A. Meister and P. Berg first isolated aminoacyl adenylates in 1958.

19.2.2. Formation of Aminoacyl-tRNA

In the second step of amino acid activation, aminoacyl-tRNA synthetase catalyzes a transfer of the aminoacyl

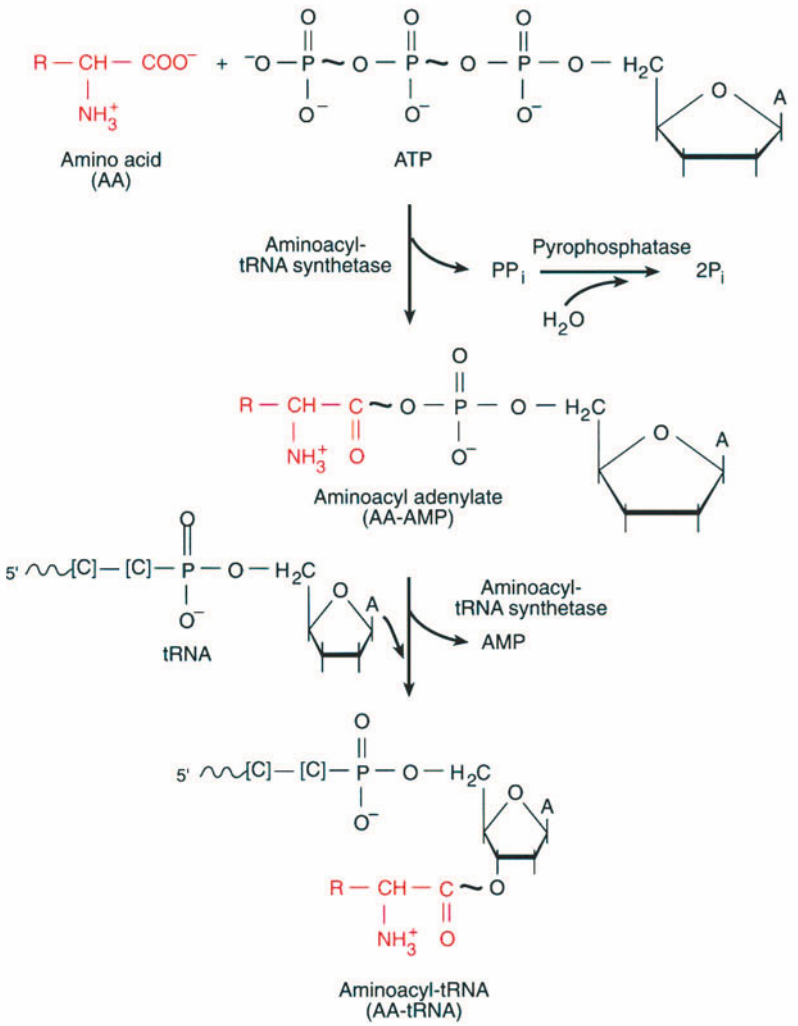


Figure 19.9. Amino acid activation.

moiety from AMP to the tRNA specific for that amino acid (its *cognate tRNA*). We refer to the linking of an amino acid to tRNA as *charging*. Charging of tRNA yields *aminoacyl-tRNA* and AMP:

$$\text{AA-AMP} + \text{tRNA} \xrightleftharpoons{\text{aminoacyl-tRNA synthetase}} \text{AA-tRNA} + \text{AMP}$$

In aminoacyl-tRNA, an amino acid has become linked covalently to the 2′- or 3′-OH of the ribose of the terminal adenine nucleotide of tRNA (Figure 19.9). The 3′-end of tRNA constitutes the acceptor site for the amino acid. In solution, the acyl group migrates back and forth between the 2′ and 3′ positions, resulting in rapid equilibration between 2′ and 3′ compounds.

Aminoacyl-tRNAs, like aminoacyl adenylates, represent energy-rich compounds. The presence of a vicinal ribose hydroxyl group and a protonated α-amino group make the acyl linkage a special ester bond ($\Delta G^{\circ\prime} = -29.3$ kJ mol^{-1}).

Biochemists have conducted many studies in an attempt to define the *synthetase recognition site,* that portion of the tRNA molecule that becomes bound to aminoacyl-tRNA synthetase. The nature of this enzyme–substrate interaction appears to vary from system to system. Some small synthetases seem to "recognize" only the amino acid acceptor region of tRNA; other synthetases recognize varying portions of the inner (concave) side of tRNA's L-shaped structure (see Figure 16.6).

19.2.3. Energetics of Amino Acid Activation

Combining the two reactions catalyzed by aminoacyl-tRNA synthetase yields two energetically coupled reac-

tions in which aminoacyl adenylate (more precisely, enzyme-bound aminoacyl adenylate) represents the common intermediate. Indicating the number of energy-rich bonds in a compound by a number in parentheses, we can write these reactions as follows:

$$\text{AA} + \overset{(2)}{\text{ATP}} \rightleftharpoons \overset{(1)}{\text{AA-AMP}} + \overset{(1)}{\text{PP}_i}$$

$$\overset{(1)}{\text{AA-AMP}} + \text{tRNA} \rightleftharpoons \overset{(1)}{\text{AA-tRNA}} + \text{AMP}$$

Overall reaction:

$$\overset{(2)}{\text{AA} + \text{ATP}} + \text{tRNA} \rightleftharpoons \overset{(1)}{\text{AA-tRNA}} + \text{AMP} + \overset{(1)}{\text{PP}_i} \xrightarrow{} 2\text{P}_i$$

You can see that the number of energy-rich bonds is the same for both the reactants and the products in each reaction. Therefore, each reaction proceeds without appreciable energy expenditure and is readily reversible. What drives amino acid activation, as it does RNA and DNA polymerization, is the exergonic hydrolysis of PP_i catalyzed by pyrophosphatase. Because of PP_i hydrolysis, amino acid activation proceeds spontaneously and essentially irreversibly. The requirement of PP_i hydrolysis means that two energy-rich bonds are ultimately expended for the activation of each amino acid: ATP $\rightarrow$ AMP; $PP_i \rightarrow 2P_i$.

For a polypeptide containing 100 amino acid residues, amino acid activation alone consumes energy equivalent to the hydrolysis of 200 ATP. As we will see, other steps in translation also require an input of energy.

Amino acid activation achieves two things. First, amino acids become linked to specific tRNAs that transport them to ribosomes for incorporation into protein. Second, amino acids become converted to reactive energy-rich compounds whose energy helps drive the endergonic synthesis of peptide bonds.

19.2.4. Fidelity of Aminoacyl-tRNA Synthetase

While aminoacyl-tRNA synthetase catalyzes both steps of amino acid activation, its specificity for the two steps differs greatly (Table 19.3). The enzyme has low specificity for forming AA-AMP but high specificity for forming AA-tRNA. These differences in specificity make good metabolic sense. Because AA-AMP constitutes only an intermediate in amino acid activation, it matters little which enzyme converts which amino acid to AA-AMP. The main thing is that amino acids become activated and readied for the second step, in which they become linked to tRNA.

This second step, however, must be carried out with utmost fidelity, since it is the tRNA that recognizes the amino acid codon on mRNA. That recognition consists of base pairing between the tRNA anticodon and the mRNA codon and determines which amino acid becomes incorporated into the polypeptide chain. Attaching a wrong amino acid to tRNA results in *misincorporation* of that amino acid. If isoleucyl-tRNA were to bind valine, for example, valine would be misincorporated instead of isoleucine, resulting in formation of a mutated protein. Fidelity of protein synthesis rests largely on the specificity of aminoacyl-tRNA synthetase in the second step of amino acid activation.

The high specificity results from the presence of a *hydrolytic site* on the enzyme that differs from the *synthetic site,* which functions in AA-AMP and AA-tRNA formation. A wrong amino acid attached to AMP is hydrolyzed off at the hydrolytic site and prevented from linking to tRNA. The capacity to catalyze hydrolysis of misactivated amino acids constitutes a *proofreading (editing) function,* much like that of the $3' \rightarrow 5'$ exonuclease activity associated with DNA polymerase (see Section 17.2). Proofreading represents a costly process, since production of an incorrect AA-AMP requires ATP hydrolysis.

We can illustrate the effectiveness of proofreading with the example of isoleucyl-tRNA synthetase, which occasionally catalyzes mischarging of $tRNA^{Ile}$ with valine instead of isoleucine. The two amino acids have similar structures and differ only by one methylene group. In the absence of proofreading, one valine is mischarged for about every 4000 isoleucines (you cannot deduce the extent of mischarging from the data of Table 19.3). With proofreading, mischarging of valine decreases to about one for every 50,000 isoleucines. Since valine is smaller than isoleucine, Val-AMP can fit into the hydrolytic site of the enzyme but Ile-AMP cannot. Such "sizing" may enable the synthetase to screen out mischarged amino acids *smaller* than the correct amino acid.

Table 19.3. Specificity of *E. coli* Ile-tRNA Synthetase[a]

	Enzyme activity[b] (μmoles formed (mg enzyme)$^{-1}$ h^{-1})	
Amino acid	AA-AMP	AA-tRNA
Ile	768	3.3
Leu	31	<0.07
Val	416	<0.03
Met	41	0.07

[a]Adapted, with permission, from P. Berg *et al., J. Biol. Chem.* 236:1726–1734 (1961).
[b]The assay mixture included unfractionated *E. coli* tRNA, ATP, Mg^{2+}, isoleucyl-tRNA synthetase, and the ^{14}C-labeled amino acid indicated.

19.3. STAGE 2: INITIATION

19.3.1. Role of Methionine

The **initiation** stage of protein synthesis produces the N-terminus of the polypeptide chain. In both prokaryotes and eukaryotes, methionine serves as the N-terminal amino acid and is coded by the mRNA codon *AUG,* called the **initiation codon** or **start codon** (5′-AUG-3′).

Analysis of amino acid composition and sequence data provided an early clue to the role of methionine in protein synthesis. In one such study, researchers identified the N-terminal amino acids in total unfractionated proteins of *E. coli.* They found that methionine was the N-terminal amino acid for almost half of these proteins despite the fact that methionine constituted only a small fraction of the total amino acids. This puzzling finding led to the discovery that all *E. coli* proteins start with methionine as their N-terminus, but many subsequently undergo posttranslational processing that involves removal of the N-terminal methionine.

19.3.2. Initiator tRNA

Two *isoacceptor transfer RNAs* recognize the AUG codon. One represents ordinary tRNA ($tRNA^{Met}$) that binds methionine and transports it to ribosomes for incorporation into protein. The other tRNA also binds methionine but has the added property that the *methionine, after it has become bound to the tRNA, can be formylated* (Figure 19.10). **Formylation** is catalyzed by a **transformylase** that requires N^{10}-formyltetrahydrofolate, a folate coenzyme, as cofactor (see Figure 14.18). The enzyme catalyzes only formylation of methionine attached to this specific tRNA; it does not catalyze formylation of free methionine. We designate the tRNA involved, called **initiator tRNA,** as $tRNA^{fMet}$ or $tRNA_f^{Met}$ and the reaction product, *N*-formylmethionyl-tRNA, as fMet-$tRNA^{fMet}$ or fMet-$tRNA_f^{Met}$

Both tRNAs have the same anticodon and recognize AUG codons in mRNA (Table 19.4). Ordinary tRNA recognizes AUG codons that occur *within a coding region* of mRNA and places methionine at internal locations in the growing polypeptide chain. Initiator tRNA, on the other hand, recognizes only an AUG initiation codon that occurs at the *start of a coding region* in mRNA and places methionine at the N-terminal end of the growing polypeptide chain.

Another important difference between the two tRNAs is that Met-$tRNA^{Met}$, like all other aminoacyl-tRNAs, binds to the A site of the ribosome. However, fMet-$tRNA^{fMet}$ binds only to the P site. By so doing, fMet-$tRNA^{fMet}$ provides the proper configuration on the ribosome for the incoming second amino acid, which, in the form of aminoacyl-tRNA, can bind only to the vacant A site

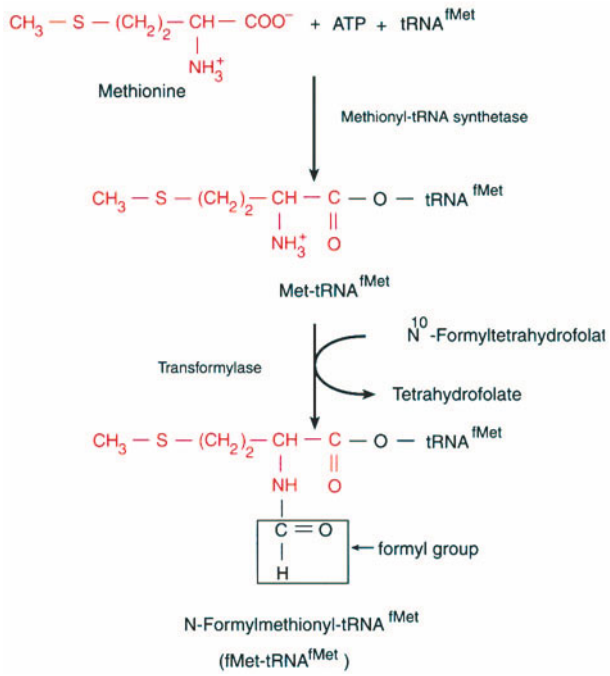


Figure 19.10. Formation of *N*-formylmethionyl-$tRNA^{fMet}$.

(shown later in Figure 19.14A). By occupying the P site, fMet-$tRNA^{fMet}$ guarantees that polypeptide chain synthesis begins with *N*-formylmethionine. Moreover, in *N*-formylmethionine the amino group has been chemically blocked. Hence, the second amino acid can only attach to the carboxyl group of methionine, thereby providing for polypeptide chain growth from the N- to the C-terminus.

Since any given polypeptide chain is likely to contain a number of methionine residues, the question arises as to how the system can select the initiation codon and avoid starting the protein with one of the internal methionine codons in mRNA. The answer is that a purine-rich segment at the 5′-end of mRNA is complementary to a pyrimidine-rich segment at the 3′-end of 16S rRNA. We call this mRNA segment the **Shine–Dalgarno sequence** in honor of the individuals who, in 1974, first pointed out

Table 19.4. Properties of Ordinary and Initiator Methionyl-tRNA

	Ordinary tRNA ($tRNA^{Met}$)	Initiator tRNA ($tRNA^{fMet}$)
Has anticodon:	5′-CAU-3′	5′-CAU-3′
Recognizes codon:	5′-AUG-3′	5′-AUG-3′
Places methionine:	Internally	At N-terminus
Bound methionine can be formylated	No	Yes
Aminoacyl-tRNA binds to:	A site	P site

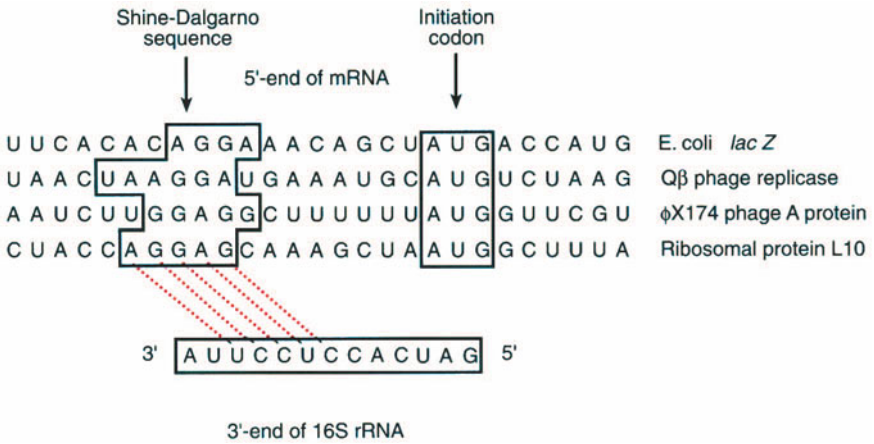


Figure 19.11. Shine–Dalgarno sequences in mRNAs that bind to 16S rRNA of *E. coli.*

its significance (Figure 19.11). The Shine–Dalgarno sequence is centered about 10 nucleotides upstream from the AUG initiation codon and forms part of the nontranslated *leader sequence* at the 5′-end of mRNA.

During initiation, mRNA binds to the 30S ribosomal subunit. In this process, the Shine–Dalgarno sequence in mRNA base-pairs with the complementary region in 16S rRNA of the 30S subunit. This places the AUG initiation codon of mRNA at the P site when the 50S subunit links up with the smaller one to form the 70S initiation complex (Figure 19.12).

Prokaryotic mRNAs frequently contain a number of coding segments, each beginning with an initiation codon that is preceded by a Shine–Dalgarno sequence. Such *polycistronic* mRNAs encode several polypeptide chains. By contrast, most eukaryotic mRNAs are *monocistronic* and code for single polypeptide chains.

19.3.3. Mechanism of Initiation

Initiation in prokaryotes, depicted schematically in Figure 19.13, requires three proteins termed **initiation factors** and denoted IF-1, IF-2, and IF-3 (Table 19.5). The term *factor* for accessory proteins in translation reflects investigators' initial uncertainty as to the protein nature of these substances. Factors involved in translation are not part of the ribosome proper but become associated with it transiently and increase both the efficiency and the fidelity of

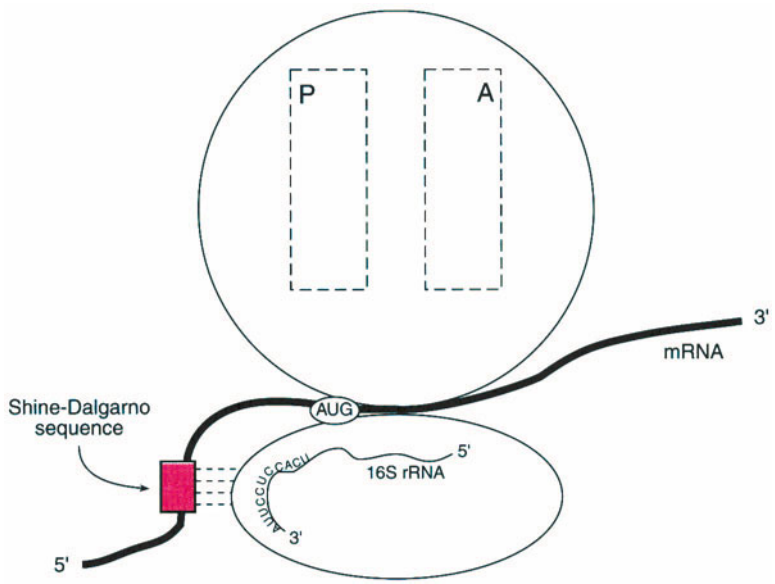


Figure 19.12. Positioning of the initiation codon by binding of the Shine–Dalgarno sequence to rRNA.

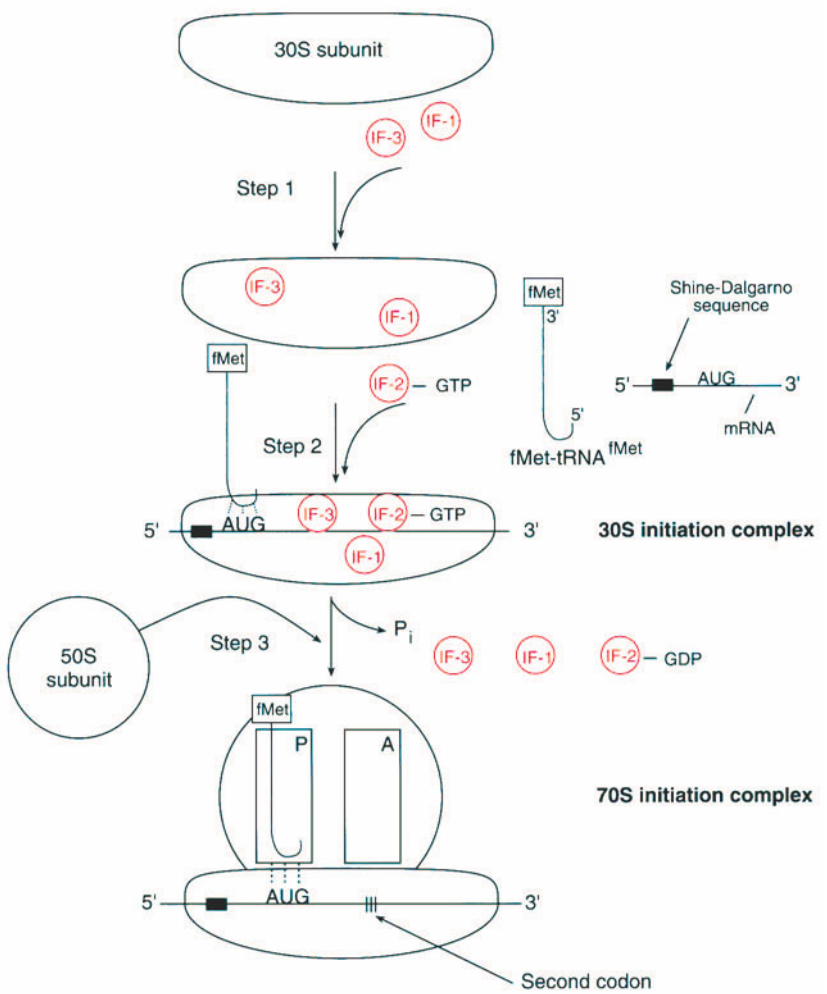


Figure 19.13. The initiation stage of protein synthesis. IF-1, IF-2, and IF-3 are initiation factors. Transfer RNA is represented by a simple (rather than L-shaped) structure to stress the antiparallel bonding between it and mRNA.

Table 19.5. Cytosolic Factors of the *E. coli* Protein Synthesis System

Factor	Molecular weight	Function
Initiation factors		
IF-1	9,000	Supports the binding of IF-3
IF-2	97,000	Binds fMet-tRNA^{fMet} and GTP
IF-3	22,000	Binds to 30S subunit and aids mRNA binding
Elongation factors		
EF-Tu	43,000	Binds aminoacyl-tRNA and GTP
EF-Ts	74,000	Displaces GDP from EF-Tu•GDP complex
EF-G	77,000	Functions in translocation by binding GTP
Release factors		
RF-1	36,000	Binds to UAA and UAG termination codons
RF-2	38,000	Binds to UAA and UGA termination codons
RF-3	46,000	Binds GTP and facilitates binding of RF-1 and RF-2

translation. The initiation process in prokaryotes consists of three steps and includes formation of a **30S initiation complex** and a **70S initiation complex:**

1. *Binding of initiation factors.* Initiation factors IF-1 and IF-3 bind to the 30S ribosomal subunit and prevent its association with the 50S subunit.
2. *Formation of the 30S initiation complex.* Initiation factor IF-2, bound to GTP, binds to the 30S subunit. The complex binds initiator tRNA that carries *N*-formylmethionine (fMet-tRNA^{fMet}). The complex also interacts with mRNA by base-pairing between the 16S rRNA of the subunit and the Shine–Dalgarno sequence of the mRNA.
3. *Formation of the 70S initiation complex.* A 50S subunit associates with the 30S initiation complex. Factors IF-1 and IF-3 dissociate. GTP bound to IF-2 undergoes hydrolysis to GDP, and the GDP/IF-2 complex is released. The initiation codon is base-paired to the anticodon of fMet-tRNA^{fMet}, which is positioned in the P site of the ribosome.

19.3.4. Initiation in Eukaryotes

The overall process of initiation in eukaryotes resembles that in prokaryotes: AUG serves as the initiation codon, two initiation complexes (40S and 80S) form, and there exist two tRNAs, one for placing methionine internally and one for placing it at the N-terminus of the polypeptide chain. However, there are also important differences between the two systems. Eukaryotic initiator tRNA (tRNA$_i^{Met}$) differs from that of prokaryotes in that the *methionine does not undergo formylation after having been attached to the tRNA.* Eukaryotic cells do not possess the enzyme transformylase. Thus, synthesis of all eukaryotic proteins starts with methionine rather than with *N*-formylmethionine. As in prokaryotes, some proteins lose their N-terminal methionine during posttranslational processing.

Eukaryotes require a larger number of initiation factors (designated eIF-1, eIF-2, and so on) than prokaryotes, and many factors consist of subunits. One of the initiation factors, called **cap-binding protein,** along with several other protein factors, interacts specifically with the methylated cap at the 5′-end of eukaryotic mRNA. The mRNA/cap-binding protein complex binds to the small ribosomal subunit, followed by binding of a ternary complex composed of eIF-2, GTP, and Met-tRNA^{Met} to yield a **40S initiation complex.** The ribosomal subunit with the bound ternary complex moves along the mRNA, away from the methylated cap, until it encounters an initiation codon. Because this movement is coupled to hydrolysis of ATP, eukaryotic systems require hydrolysis of both ATP and GTP for initiation. In the final stage of initiation, a 60S ribosomal subunit binds to the complex, the initiation factors dissociate, and an **80S initiation complex** forms. At this point, Met-tRNA^{Met} has been positioned so that it interacts with the initiation codon at the P site of the ribosome.

Eukaryotic mRNAs do not have Shine–Dalgarno sequences, and the initiation codon frequently represents the first AUG codon of the message. Since most eukaryotic mRNAs are monocistronic, they have only a single initiation codon per mRNA. Interestingly, initiation of protein synthesis in mitochondria and chloroplasts of eukaryotic cells proceeds with *N*-formylmethionine rather than methionine. This finding supports the theory that mitochondria and chloroplasts have evolved from bacteria that took on a symbiotic relationship with early eukaryotic cells (see Section 1.1).

19.4. STAGE 3: ELONGATION

The bulk of protein synthesis occurs during the **elongation** stage and consists of successive additions of amino acid residues to the growing polypeptide chain. The first amino acid added represents the protein's second amino acid, the one following N-terminal methionine. The last amino acid added constitutes the C-terminus of the polypeptide chain.

To discuss the elongation process in general terms (Figure 19.14), we will assume that n amino acids (AA) have previously been polymerized and that amino acid $n + 1$ must be incorporated next. In that case, AA$_n$ constitutes the C-terminus of the growing polypeptide chain; it serves as the link between the polypeptide chain and its own cognate transfer RNA, tRNA$_n$. Because tRNA$_n$ carries a growing polypeptide chain, it represents a *peptidyl-tRNA,* bound at the P site of the ribosome. The tRNA$_n$ anticodon is base-paired to the appropriate mRNA codon$_n$. The A site is vacant and ready to accept the next incoming amino acid (AA$_{n+1}$) attached to its cognate tRNA, tRNA$_{n+1}$ (Figure 19.14A).

Incorporation of each amino acid into the polypeptide chain proceeds by means of an elongation cycle, composed of three steps, which we shall discuss in order: aminoacyl-tRNA binding, peptide bond formation, and translocation.

19.4.1. Aminoacyl-tRNA Binding

Elongation starts with binding of the incoming amino acid (AA$_{n+1}$), linked to tRNA$_{n+1}$, at the A site of the ribosome. Binding of the aminoacyl-tRNA is codon-directed: The anticodon of tRNA$_{n+1}$ base-pairs with codon $n + 1$ in mRNA. At this point, both ribosome binding sites are oc-

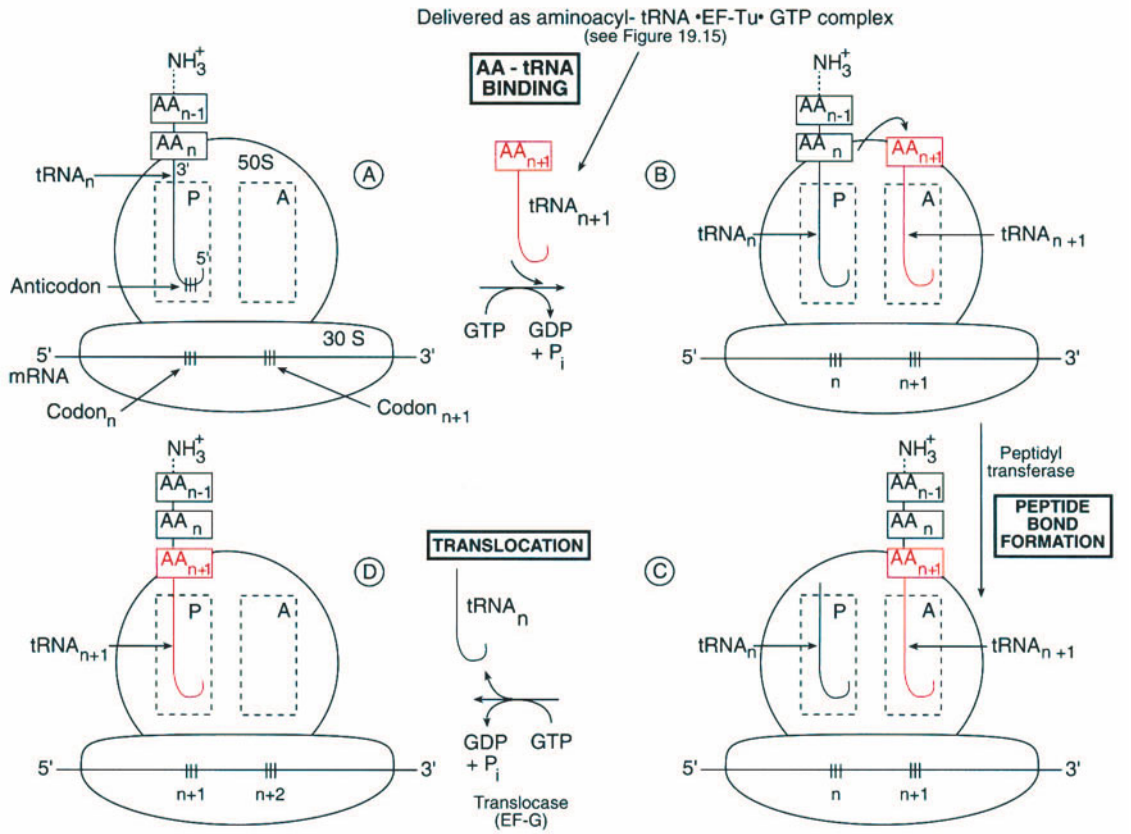


Figure 19.14. The elongation stage of protein synthesis. Following the elongation cycle for incorporation of amino acid$_{n+1}$ (shown), the system is ready for incorporation of amino acid$_{n+2}$. The first elongation cycle occurs when fMet-tRNA^{fMet} binds at the P site and amino acid$_2$ enters the A site.

cupied, the P site with peptidyl-tRNA, and the A site with aminoacyl-tRNA (Figure 19.14B).

Aminoacyl-tRNA binding requires **GTP** hydrolysis. GTP represents part of a complex regulatory system that includes **elongation factor T (EF-T).** This factor consists of two subunits, **EF-Ts** and **EF-Tu,** denoting a "temperature-stable" and a "temperature-unstable" protein, respectively. EF-Tu is the most abundant protein of *E. coli* and constitutes about 5% of the total cellular protein. Approximately 100,000 copies of EF-Tu occur per cell. Binding of GTP to the complete EF-T results in dissociation of EF-Ts and linking of EF-Tu to GTP. The binary complex of EF-Tu and GTP binds the aminoacyl-tRNA to form a ternary complex (aminoacyl-tRNA·EF-Tu·GTP), which binds to the A site of the ribosome. The aminoacyl-tRNA remains in the A site, GTP is hydrolyzed to GDP and P_i, and the EF-Tu·GDP complex dissociates from the ribosome. EF-Ts replaces GDP in the EF-Tu·GDP complex, thereby regenerating the complete EF-T factor (Figure 19.15).

EF-Tu plays an important role in translation because it does not react with either formylated or unformylated methionyl-tRNA^{fMet}. Hence, initiator tRNA cannot bind at the A site and binds at the P site instead. Additionally, involvement of EF-Tu helps to minimize misincorporation of amino acids. After the ternary complex binds to the ribosome, some time must elapse to allow for GTP hydrolysis and dissociation of Ef-Tu·GDP. Thus, the new peptide bond cannot form immediately upon arrival of an aminoacyl-tRNA at the A site. The built-in delay allows for correction of errors made in aminoacyl-tRNA binding. If anticodon–codon matching is incorrect, the wrong aminoacyl-tRNA dissociates rapidly from the ribosome. Proofreading occurs both before and after GTP hydrolysis to ensure that only the correct aminoacyl-tRNA remains bound. As a result, the observed error frequency in translation is only of the order of about one misincorporation per 10^4 amino acid residues.

19.4.2. Peptide Bond Formation

The second step of the elongation cycle accomplishes **peptide bond formation.** The bond forms between an incoming amino acid (AA_{n+1}) and the C-terminus of the growing polypeptide chain by shifting the entire growing chain onto the incoming amino acid (Figure 19.16). This

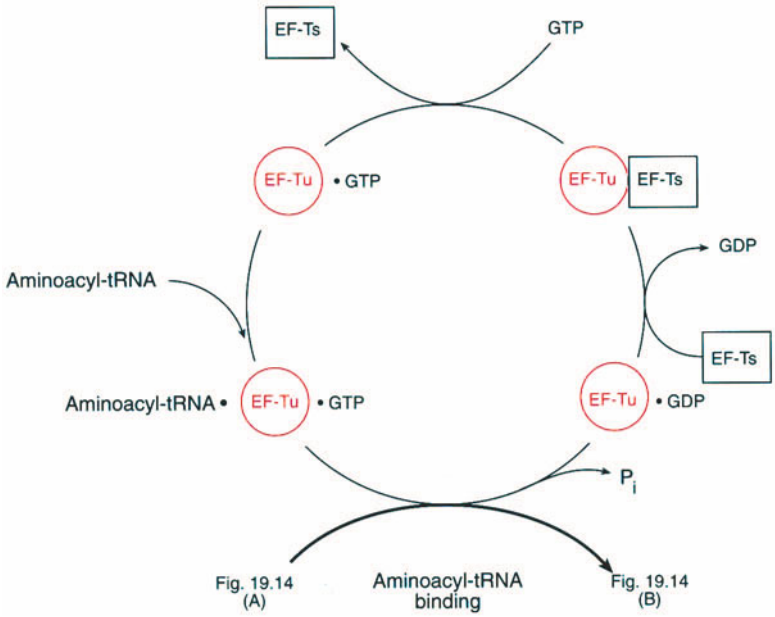


Figure 19.15. The cyclic set of reactions involving elongation factor T (EF-Tu•EF-Ts). The aminoacyl-tRNA•EF-Tu•GTP complex positions the aminoacyl-tRNA at the A site of the ribosome. GTP is hydrolyzed, and the EF-Tu•GDP complex diffuses away.

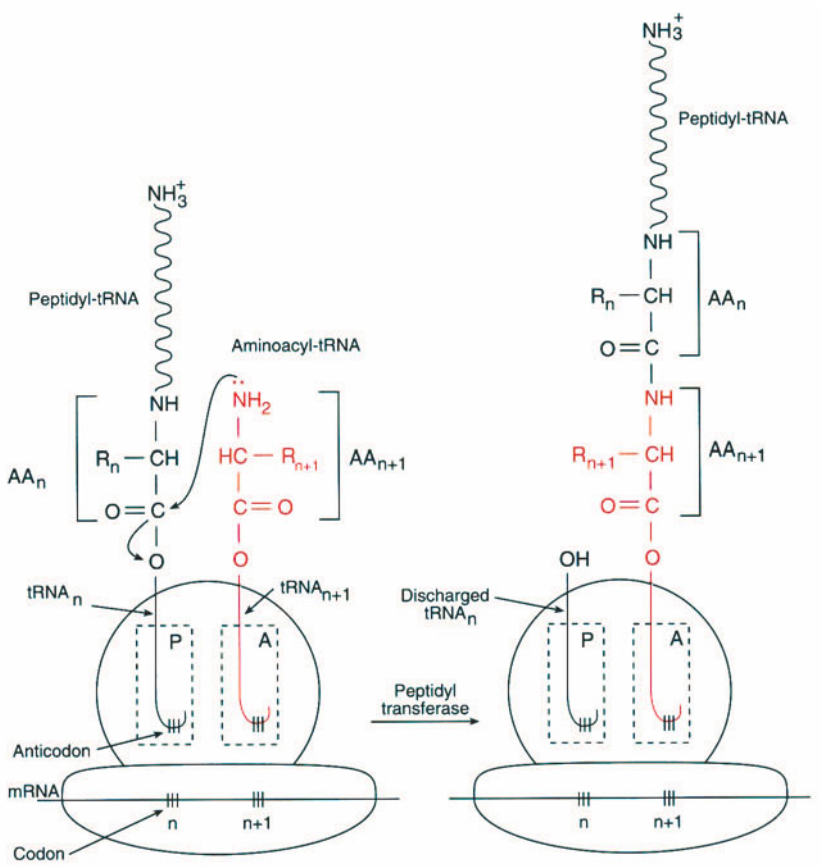


Figure 19.16. Action of peptidyl transferase in peptide bond formation. The growing polypeptide chain is transferred from a peptidyl-tRNA in the P site to an aminoacyl-tRNA in the A site.

transpeptidation, catalyzed by **peptidyl transferase,** involves a nucleophilic attack of the amino group of an incoming amino acid on the carbonyl carbon of peptidyl-tRNA. The reaction proceeds without input of energy. Cleavage of the energy-rich ester bond in peptidyl-tRNA provides sufficient energy to drive peptide bond formation. Action of peptidyl transferase results in formation of a new peptidyl-tRNA, *bound at the A site.* The growing polypeptide chain is attached to $tRNA_{n+1}$ via the appropriate amino acid (AA_{n+1}). The previous, **discharged tRNA** ($tRNA_n$) remains attached at the P site (Figure 19.14C).

Peptidyl transferase is a component of the ribosome and not an external factor. Activity of the *E. coli* enzyme has been shown to be associated with the central protuberance of the 50S subunit (see Figure 19.1) and to result from juxtaposition of a number of 50S ribosomal proteins as well as 23S ribosomal RNA.

19.4.3. Translocation

Translocation constitutes the third step of the elongation cycle and consists of a relative alignment of ribosome and mRNA. It begins with ejection of the discharged tRNA ($tRNA_n$) from the P site. After the P site has been vacated, the peptidyl-tRNA shifts from the A to the P site. During this shift, the tRNA remains H-bonded to mRNA; anticodon–codon base pairs are maintained. Hence, as $tRNA_{n+1}$ shifts from the A site to the P site, the mRNA is "pulled" through the cleft in the ribosome over a distance equivalent to the length of one codon. At the end of the translocation step, codon $n + 1$ is at the P site while codon $n + 2$ is at the A site. Translocation requires the action of **elongation factor EF-G (translocase)** and hydrolysis of GTP. Translocation accomplishes the following:

- *Ejecting the discharged tRNA* of the previous aminoacyl-tRNA from the P site
- *Shifting the peptidyl-tRNA* from the A site to the P site
- *Pulling the mRNA* through the ribosome cleft by the length of one codon

Following the action of translocase, the translational apparatus has regained its initial configuration except that the P site holds the new peptidyl-tRNA and the A site is juxtaposed to the next codon ($n + 2$), ready for accepting the next incoming amino acid (AA_{n+2}) (Figure 19.14D).

19.5. STAGE 4: TERMINATION

Termination begins when one of the three *termination* or *stop codons (UAA, UAG, UGA)* enters the A site on the ribosome (Figure 19.17). Because these codons are not recognized by ordinary tRNAs, they cannot function in ordinary codon–anticodon binding. They are, however, recognized by one of two **release factors:** *RF-1* binds to codons UAA and UAG, and *RF-2* binds to codons UAA and UGA. The third release factor, *RF-3,* binds GTP and facilitates the binding of RF-1 and RF-2 to the A site. Attachment of a release factor to the A site prevents binding of aminoacyl-tRNA to this site.

Binding of a release factor to the A site also alters the specificity of peptidyl transferase. Consequently, in the second stage of the termination process, the enzyme catalyzes transfer of the polypeptide chain to water rather than to an incoming amino acid. In this reaction, peptidyl transferase acts as a hydrolase; it causes hydrolysis of the completed polypeptide chain from tRNA, resulting in dis-

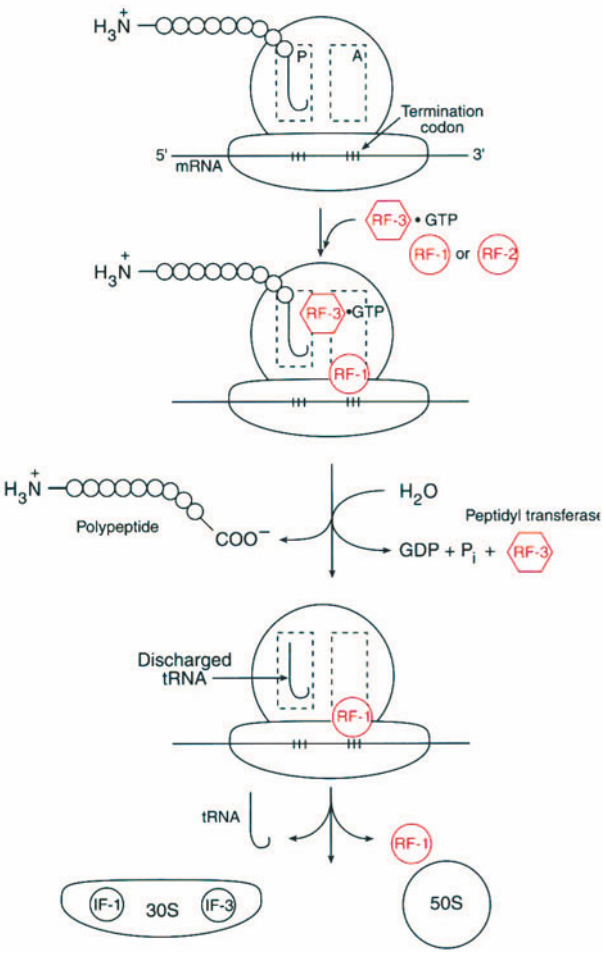


Figure 19.17. The termination stage of protein synthesis. RF-1, RF-2, and RF-3 are release factors. Upon dissociation of the ribosome, initiation factors IF-1 and IF-3 bind to the 30S subunit to prevent its association with a 50S subunit in the absence of mRNA.

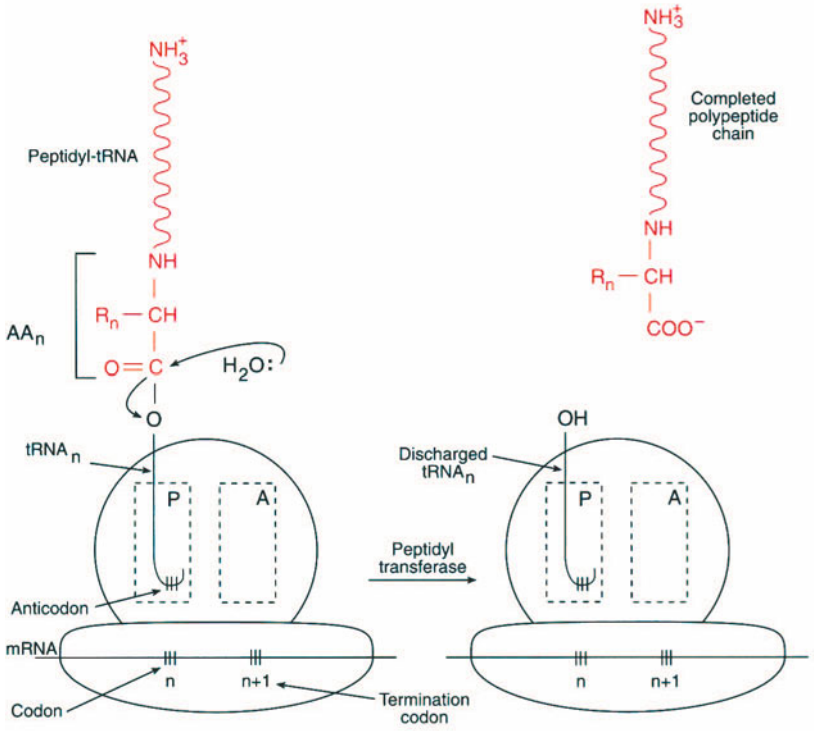


Figure 19.18. Action of peptidyl transferase in termination of translation. The enzyme functions as a hydrolase, and the completed polypeptide chain is transferred from the peptidyl-tRNA in the P site to a molecule of water.

sociation of the polypeptide chain from the ribosome (Figure 19.18).

In the final stage of termination, release of the completed polypeptide chain is followed by release of discharged tRNA, mRNA, and release factors and dissociation of the ribosome into its two component subunits. Concomitantly, GTP undergoes hydrolysis to GDP and P_i. Binding of IF-1 and IF-3 to the 30S subunit prevents its association with a 50S subunit and formation of a dead-end 70S ribosome, devoid of bound mRNA.

19.6. ENERGETICS AND CONTROL

19.6.1. Energy Considerations

Initiation, elongation, and termination require hydrolysis of GTP rather than ATP. It appears unlikely that hydrolysis of GTP provides energy to drive protein synthesis because (a) translation does occur, albeit at a very slow rate, in the absence of GTP hydrolysis, and (b) no energy-rich intermediates have been identified in which GTP or GDP forms covalent bonds with specific substances. Hence, researchers think that GTP functions as an allosteric effector, binding to specific ribosomal components and causing them to undergo conformational changes. According to this view, the conformational changes lead to activation of ribosomal components and to catalysis of GTP hydrolysis. Conversion of GTP to GDP and P_i results in relaxation of the ribosomal components and a return to their initial conformation.

GTP-binding proteins of translation may belong to the group of *G proteins,* regulatory proteins that serve as cellular signals and transfer factors (see Section 10.5). A G protein becomes activated upon binding GTP. Bound GTP undergoes slow hydrolysis, which, when completed, results in a termination of the G protein's regulatory effect. At that point, the G protein can again be reactivated by binding GTP.

Table 19.6 summarizes the energy balance sheet of translation in terms of energy-rich bonds. Amino acid activation requires an expenditure of two energy-rich bonds per amino acid (ATP $\rightarrow$ AMP and $PP_i \rightarrow 2P_i$). Initiation and termination of the polypeptide chain require hydrolysis of one GTP each, and every elongation cycle requires hydrolysis of two GTPs. You can calculate the energy expenditure for synthesis of a hypothetical protein, consisting of n amino acid residues, from these data. The elongation cycle is used $n - 1$ times, since all but the N-terminal amino acid must pass through it. Thus, the to-

Table 19.6. Energy Requirements of Translation

Step	Number of energy-rich bonds: Per amino acid	Per protein
Amino acid activation	2	
Initiation of polypeptide chain		1
Elongation		
Aminoacyl-tRNA binding	1	
Peptide bond formation		
Translocation	1	
Termination of polypeptide chain		1
Total	4	2

tal number of energy-rich bonds hydrolyzed can be tabulated as follows:

Step	Number of energy-rich bonds
Amino acid activation	$2n$
Initiation	1
Elongation	$2(n-1)$
Termination	1
Total	$4n$

Assuming that a protein consists of 300 amino acids ($n = 300$) and that the intracellular hydrolysis of an energy-rich bond has a free energy change of $\Delta G' = -30.5$ kJ mol^{-1}, the energy expenditure is (4)(300)(30.5) = 36,600 kJ mol^{-1} per protein or 36,600/299 = 122.4 kJ mol^{-1} per peptide bond. Since the biochemical standard free energy change ($\Delta G^{\circ\prime}$) for peptide bond formation has a value of only 8–13 kJ mol^{-1}, you can see that cells pay a very high price for protein synthesis. The process is so expensive because it must proceed essentially in an irreversible manner and with a high degree of fidelity. Because the cell manufactures proteins *of defined amino acid sequences,* many steps require proofreading, which takes energy. In most organisms, protein synthesis consumes more energy than any other biosynthetic process. Researchers estimate that, in *E. coli,* up to 90% of the biosynthetic energy is used for protein synthesis.

Construction of a unique amino acid sequence represents a tremendous synthetic achievement for the cell. A polypeptide chain of 300 amino acid residues and made from 20 different amino acids can be put together in 20^{300}, or approximately 10^{390}, different ways. In protein biosynthesis, the cell accomplishes the synthesis of one specific sequence out of this astronomical number of possible sequences. Looked at in these terms, the energy price paid by the cell for protein synthesis does not seem exorbitant at all.

19.6.2. Regulation of Protein Synthesis

In both prokaryotes and eukaryotes, control at the level of transcription is the primary determinant of the amount of protein synthesized. In eukaryotes, some control at the level of translation seems to be more widespread, and regulatory effects appear to be exercised primarily at the stage of polypeptide chain initiation. The two best-understood cases involve the regulation of hemoglobin biosynthesis in reticulocytes and the effects of interferons on protein synthesis in virus-infected cells.

Reticulocytes contain a kinase, called **heme-controlled inhibitor,** that provides a link between the level of heme and the level of globin, for which heme serves as a prosthetic group. The kinase is inactive as long as the heme concentration is adequate. When the heme concentration falls to a certain value, the kinase undergoes activation and catalyzes a phosphorylation of initiation factor eIF-2. The phosphorylated factor cannot function properly in initiation and thereby blocks translation. As the level of heme decreases, the synthesis of globin decreases as well.

Interferons are small glycoproteins (MW = 26,000–38,000) whose secretion is induced in vertebrate cells by infection with viruses. Interferons *interfere* with viral replication and act in one of two ways. They may induce a nuclease that specifically degrades viral mRNA, or they may induce a kinase that inactivates initiation factor eIF-2 by catalyzing its phosphorylation. The resultant block in translation prevents production of viral proteins and intact viral particles and inhibits viral replication.

19.6.3. Effects of Antibiotics

An *antibiotic* is a substance produced by one microorganism (bacterium or fungus) and toxic to other microorganisms (Table 19.7 and Figure 19.19). Many of the known antibiotics function by blocking translation. Because of the differences between translational systems in prokaryotes and eukaryotes, most antibiotics effective against prokaryotes are harmless to eukaryotes. This property accounts for the widespread use of antibiotics. A translation-apparatus-directed antibiotic produced by a given microorganism does not inhibit that microorganism's own translational machinery. The reason is that the microorganism carries plasmids, which contain genes that confer resistance to the antibiotic.

Diphtheria toxin, while not an antibiotic, also acts by inhibiting translation. Diphtheria toxin is produced by *Corynebacterium diphtheriae,* a bacterium that harbors a

Table 19.7. Mode of Action of Several Antibiotics

Antibiotic	Effect on protein synthesis
Tetracycline	Inhibits binding of aminoacyl-tRNA to 30S ribosomal subunit. Acts in *prokaryotes*.
Streptomycin	Inhibits chain initiation by binding to 30S ribosomal subunit and causes misreading by interfering with base-pairing between codon and anticodon. Acts in *prokaryotes*.
Chloramphenicol (Chloromycetin)	Inhibits peptidyl transferase of 50S ribosomal subunit and thus blocks chain elongation. Acts in *prokaryotes*.
Erythromycin	Binds to the 50S ribosomal subunit and inhibits translocation, thus blocking chain elongation. Acts in *prokaryotes*.
Fusidic acid	Prevents dissociation of EF-G•GDP from 50S ribosomal subunit and thus blocks chain elongation. Acts in *prokaryotes*.
Cycloheximide	Inhibits peptidyl transferase of 60S ribosomal subunit and thus blocks chain elongation. Acts in *eukaryotes*.
Puromycin	Causes premature chain termination because of its structural similarity to the 3′-end of aminoacyl-tRNA. Puromycin enters the A site and transfers to the growing chain, thereby blocking further extension and causing premature release of the incompleted chain as peptidyl puromycin. Acts in both *prokaryotes and eukaryotes*.

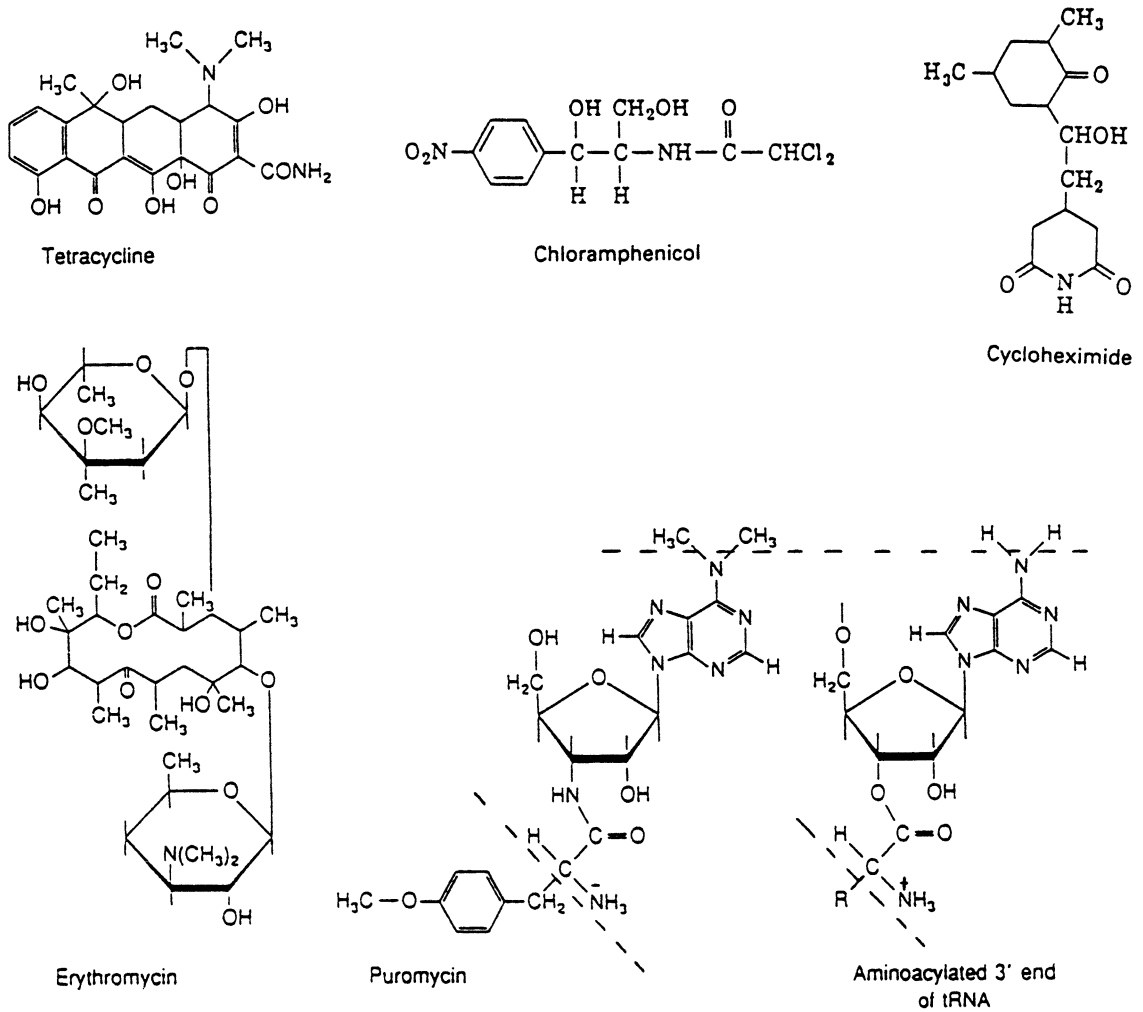


Figure 19.19. Structures of some antibiotics. Dashed lines show the structural similarity of puromycin to the 3′-end of aminoacylated tRNA.

bacteriophage, *corynephage* β. The toxin consists of an acidic, single-chain, phage-encoded protein that specifically inhibits elongation factor eEF-2. Diphtheria toxin acts as a catalyst for a reaction in which ADP-ribose is linked to eEF-2. Minute amounts of the toxin block protein synthesis completely. Consequently, pure diphtheria toxin represents one of the deadliest substances known.

19.7. STAGE 5: POSTTRANSLATIONAL MODIFICATIONS

19.7.1. Folding and Processing

Proteins fold spontaneously to assume their unique three-dimensional structures. The polypeptide chain begins to fold while still attached to the ribosome during translation. Support for this concept comes from experiments in which incompletely synthesized proteins were shown to possess enzymatic activity or antigen-binding capacity.

For some proteins, the amino acid sequence constitutes the sole determinant of the folding pattern. For others, especially some larger proteins, the primary structure alone is insufficient for correct folding, and additional factors are required to achieve the proper three-dimensional configuration. Some factors are enzymes that catalyze a reshuffling of sulfhydryl groups and disulfide bonds. Other factors, called **chaperones,** consist of proteins that function in the folding of other polypeptide chains. Some chaperones prevent incorrect folding, or reverse incorrect folding, by binding to hydrophobic regions of the polypeptide chain, followed by release of the chaperone from the protein once folding has been completed. Often, binding and release occur repeatedly. Chaperones also function to prevent improper intermolecular associations of polypeptide chains. Some chaperones appear to function by promoting specific stages of the folding process. Many chaperones are ATPases (enzymes that catalyze ATP hydrolysis) and apparently apply the free energy released to their role in polypeptide chain folding.

In addition to taking on specific three-dimensional configurations, proteins also undergo many types of *post-translational processing.* Prokaryotic proteins may require alterations of their N-terminal ends. Typical changes include removal of the formyl group of *N*-formylmethionine by a *deformylase,* removal of *N*-formylmethionine, or removal of *N*-formylmethionine plus one or more amino acids that follow methionine in the chain. The last two modifications require the action of one or more *peptidases.* Eukaryotic proteins, made with N-terminal methionine, may require similar processing.

Other posttranslational modifications of proteins include phosphorylation to produce phosphoproteins, attachment of metals and other prosthetic groups, formation of disulfide bonds, linking of proteins to carbohydrates or lipids, and hydroxylation to produce hydroxyamino acids like hydroxylysine and hydroxyproline.

19.7.2. Protein Targeting

After a protein has been synthesized on the ribosome, it must be transported to some intracellular or extracellular locus where it serves its particular function. Dispatching proteins to given locations, or **targeting,** is accomplished by endowing proteins with specific markers, expressed in terms of amino acid sequences and/or structural characteristics.

A major targeting strategy deals with proteins destined to be secreted across a membrane—the cell membrane in prokaryotes and the endoplasmic reticulum in eukaryotes. In both instances, membrane-bound ribosomes complete the synthesis of such proteins even though synthesis of *all* proteins begins on free cytoplasmic ribosomes. For proteins destined to be secreted across a membrane, the initial polypeptide segment synthesized in the cytoplasm provides a signal for the ribosome to move toward and become bound to the membrane, where the chain is subsequently completed. In 1970, David Sabatini and Günter Blobel proposed a **signal hypothesis** to describe protein synthesis by membrane-bound ribosomes.

According to this hypothesis, the signal for membrane-bound synthesis consists of a sequence of amino acids near the N-terminus (methionine) of the polypeptide chain, called the **signal peptide** and typically comprising about 15–30 amino acid residues. Signal peptides have no unique amino acid sequences but always contain a highly hydrophobic core of 10–15 amino acids. In eukaryotes, the signal peptide is recognized by a **signal recognition particle (SRP)** (Figure 19.20) that consists of six small proteins and one 7S RNA molecule (total MW = 325,000). SRP binds tightly to ribosomes containing a growing polypeptide chain that carries a signal peptide. SRP binding stops translation temporarily and prevents the ribosome-bound protein from being released into the cytosol. The SRP–ribosome complex diffuses toward the membrane of the endoplasmic reticulum, where SRP binds to a specific receptor (**SRP receptor**), located in the membrane.

Binding of the SRP–ribosome complex to the endoplasmic reticulum leads to resumption of protein synthesis on the ribosome, which is now membrane-bound. The signal peptide becomes inserted into the lipid bilayer of the membrane, a process aided by the hydrophobic nature of the peptide. This further anchors the ribosome to the membrane. As translation proceeds, an ATP-dependent

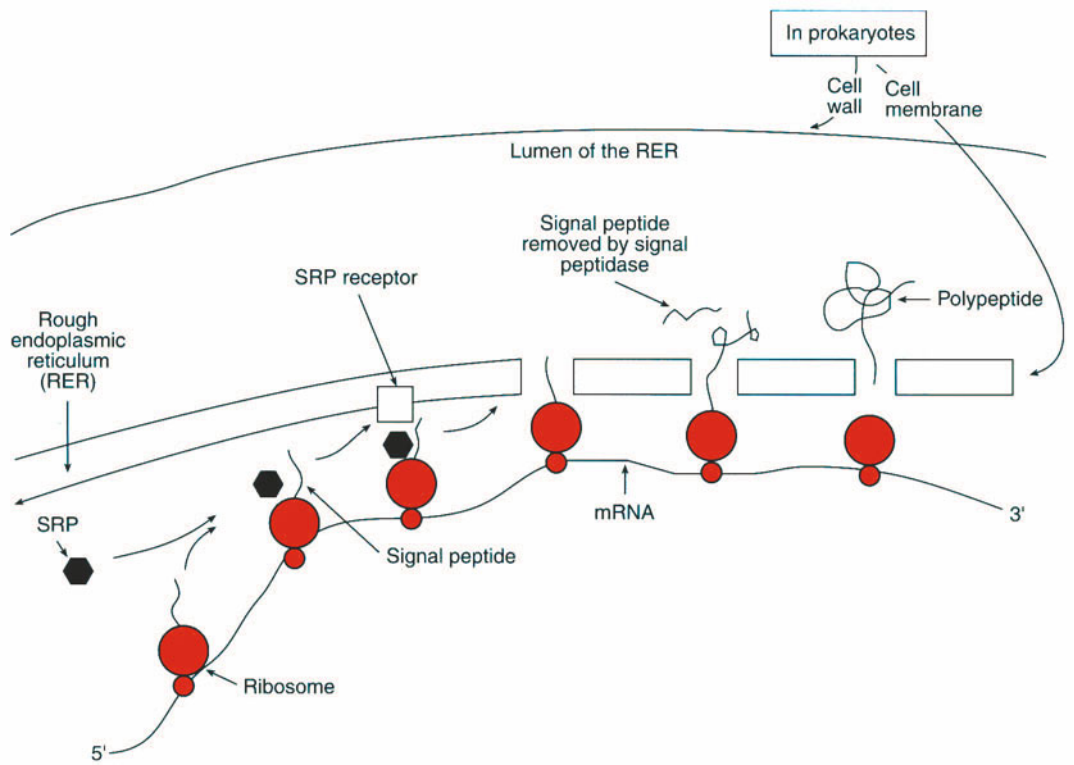


Figure 19.20. Eukaryotic protein synthesis by membrane-bound ribosomes according to the signal hypothesis. SRP, Signal recognition particle. In prokaryotes, proteins destined for export from the cell are synthesized via a similar mechanism, indicated on the right.

process causes extrusion of the chain, led by the signal peptide, through the membrane. Translation and extrusion of the protein occur simultaneously.

Translation is terminated by deposition of the protein in the lumen (enclosed space) of the endoplasmic reticulum and dissociation of the ribosome from the membrane. **Signal peptidase**—an integral membrane protein—catalyzes cleavage of the signal peptide from the protein. Membrane proteins have signal sequences that resist the action of signal peptidase so that the signal sequences of such proteins are not cleaved off, and they remain anchored in the membrane. All other proteins pass completely through the membrane and into the *Golgi apparatus,* a major sorting center of the cell. From there, proteins are dispatched to various locations within the cell and are *exported* from the cell. Many proteins are first altered by *glycosylation* (Section 5.5), using either O-linked or N-linked oligosaccharides. *O-linked oligosaccharides* are synthesized in the Golgi apparatus; *N-linked oligosaccharides* are synthesized in the endoplasmic reticulum but altered in many ways in the Golgi apparatus. Carbohydrate portions of glycosylated proteins provide additional recognition signals for targeting.

In prokaryotes, synthesis of proteins by membrane-bound ribosomes proceeds via a mechanism similar to that shown in Figure 19.20. Ribosomes move to the cell membrane upon recognition of a signal peptide, and proteins are extruded through this membrane.

19.7.3. Protein Degradation

Proteins have a finite life span. They are subject to environmental damage resulting in oxidation or denaturation. Proteins may also be altered by accidental biochemical reactions, such as an improper enzymatic hydrolysis or phosphorylation. Altered proteins undergo degradation, and their replacement occurs via synthesis of new molecules. Hence, proteins are continuously synthesized and degraded, resulting in a *turnover* that we can measure by a *half-life*—the time required for one-half of the mass of a protein to be metabolized, degraded, or excreted.

Half-lives vary greatly. Ornithine decarboxylase, phosphoenolpyruvate carboxylase, and lactate dehydrogenase have half-lives of 12 min, 5 h, and 6 days, respectively. Major regulatory proteins tend to have short half-lives. Protein degradation helps to control the concentration of critical enzymes and protects an organism against accumulation of abnormal and potentially harmful

proteins that may be produced by mutations, errors in transcription or translation, oxidation, or denaturation.

Proteins may be degraded by particulate systems derived from lysosomes that fuse with the cell surface or with subcellular structures and subsequently release their proteolytic enzymes. Proteins may also be degraded by means of soluble, ATP-dependent systems that occur in both prokaryotes and eukaryotes. The best-known eukaryotic system is that involving **ubiquitin,** a small protein of *ubiquitous* occurrence in eukaryotic cells. Ubiquitin consists of 76 amino acid residues (MW = 8500) and constitutes one of the most highly conserved proteins known. Ubiquitins isolated from many diverse organisms are either identical or differ only by a few amino acids.

Ubiquitin combines covalently with proteins destined for degradation (Figure 19.21). In this process, ubiquitin first undergoes activation by the formation of a thioester linkage between its carboxyl group (C-terminus)

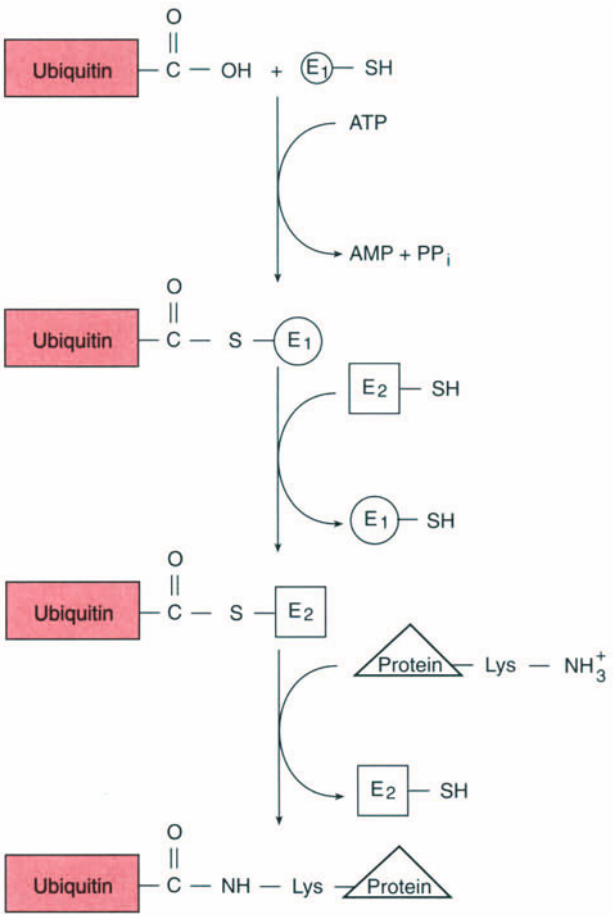


Figure 19.21. Marking of cytosolic proteins for degradation by covalent attachment of ubiquitin. E_1 and E_2 are enzymes functioning in the transfer of ubiquitin to the ϵ-NH_2 group of a lysine residue in the target protein.

Table 19.8. Dependence of the Half-Life of Cytosolic Proteins on the Nature of Their N-Terminal Amino Acid[a]

N-Terminal amino acid	Protein half-life (*S. cerevisiae*)
Stabilizing	
Ala, Gly, Met, Ser, Thr, Val	>20 h
Destabilizing	
Ile, Glu	~30 min
Tyr, Gln	~10 min
Highly destabilizing	
Leu, Phe, Asp, Lys	~3 min
Arg	~2 min

[a]Adapted, with permission, from A. Bachmair *et al., Science* 234:179–186 (1986). Copyright © 1986 American Association for the Advancement of Science.

and the sulfhydryl group of an enzyme, E_1. ATP hydrolysis drives this step. Next, ubiquitin transfers to a second enzyme, E_2, forming another thioester bond. In the third step, ubiquitin transfers from E_2 to the ϵ-amino group of a lysine residue in the "condemned" protein. We refer to the covalent linking of ubiquitin to a protein as *ubiquitination.* Ubiquitin-marked proteins are degraded by specific proteases that recognize the ubiquitin marker. Ubiquitin is removed by hydrolysis, and the "condemned" protein is degraded in an ATP-dependent process.

We do not yet fully understand how proteins are encoded for shorter or longer half-lives and how they are selected for degradation. Researchers believe both of these aspects of protein metabolism to be determined by the amino acid sequences and/or other structural characteristics of protein molecules. Two such molecular properties have already been identified—nature of the N-terminus and occurrence of PEST sequences.

Composition of the N-terminus correlates with the half-life of proteins (Table 19.8). Certain N-terminal amino acids stabilize proteins, while others destabilize them. The relationship between a protein's in vivo half-life and the identity of its N-terminal residue has been termed the N-end rule. Similar but distinct versions of the N-end rule operate in all organisms examined, from bacteria to mammals. In eukaryotes, the N-end rule pathway constitutes part of the ubiquitin system.

Occurrence of *PEST sequences* in proteins also correlates with their half-lives. PEST sequences are amino acid sequences rich in proline (P), glutamic acid (E), serine (S), and threonine (T). It appears that the greater the number of PEST sequences is in a protein, the shorter the half-life. While composition of the N-terminus and occurrence of PEST sequences play important roles, it is likely that additional factors are involved in determining protein half-lives and marking them for degradation.

SUMMARY

Amino acids are converted by means of ATP and aminoacyl-tRNA synthetase first to aminoacyl adenylates and subsequently to amino acyl-tRNAs. Hydrolysis of PP_i drives amino acid activation. Aminoacyl-tRNA synthetases have low specificity for forming aminoacyl adenylates but high specificity for forming aminoacyl-tRNAs.

All prokaryotic proteins start with *N*-formylmethionine as their N-terminus. *N*-Formylmethionyl-tRNA binds to the P site on the ribosome, thereby positioning the N-terminal amino acid for peptide bond formation. The 5′-end of mRNA contains a purine-rich sequence (Shine–Dalgarno sequence) that base-pairs with a pyrimidine-rich sequence at the 3′-end of 16S rRNA during initiation, thereby positioning the initiation codon at the P site. Eukaryotic proteins start with methionine.

Elongation consists of three reactions, repeated for each amino acid: aminoacyl-tRNA binding, peptide bond formation, and translocation. In the first reaction, an incoming aminoacyl-tRNA binds at the A site on the ribosome. Binding requires hydrolysis of GTP and a complex regulatory system involving two soluble factors, EF-Tu and EF-Ts. In the second reaction, the growing polypeptide chain transfers from the P site to the incoming amino acid at the A site. Peptidyl transferase, a component of the 50S ribosomal subunit, catalyzes peptide bond formation. In the third reaction, peptidyl tRNA shifts from the A to the P site, and mRNA is "pulled" through a cleft in the ribosome by the length of one codon. Translocation is catalyzed by translocase and requires GTP hydrolysis.

Termination begins when a termination codon enters the A site on the ribosome. Binding of a release factor alters the specificity of peptidyl transferase so that it transfers the growing polypeptide chain to water, thereby terminating translation. Termination requires GTP hydrolysis.

In both prokaryotes and eukaryotes, the amount of protein synthesized is determined primarily by control at the level of transcription. In eukaryotes, some regulation of translation appears to be more widespread and seems to occur primarily at the stage of polypeptide chain initiation. Posttranslational processing includes final folding of the polypeptide chain, changes at the N-terminus, and other enzymatic modifications. Proteins secreted from the cell are synthesized by membrane-bound ribosomes and carry a signal peptide that aids in their transfer through the endoplasmic reticulum in eukaryotes and the cell membrane in prokaryotes.

Proteins have half-lives of varying lengths and are degraded by both particulate and soluble systems. The best-known soluble system in eukaryotes is that in which "condemned" proteins are marked with ubiquitin. Scientists believe that composition of the N-terminus and the occurrence of PEST sequences in a protein are, in part, responsible for determining the length of the protein's half-life and its marking for degradation by ubiquitin.

SELECTED READINGS

Arnstein, H. R. V., and Cox, R. A., *Protein Biosynthesis,* IRL Press, Oxford (1992).

Austen, B. M., and Westwood, O. M. R., *Protein Targeting and Secretion,* IRL Press, Oxford (1991).

Dunn, W. A., Jr., Autophagy and related mechanisms of lysosome-mediated protein degradation, *Trends Cell Biol.* 4:139–143 (1994).

Gesteland, R. F., and Atkins, J. F., Recoding: Dynamic programming of translation, *Annu. Rev. Biochem.* 65:741–768 (1996).

Graves, D. J., Martin, B. L., and Wang, J. H., *Co- and Posttranslational Modification of Proteins: Chemical Principles and Biological Effects,* Oxford University Press, New York (1994).

Hartl, F. U., Molecular chaperones in cellular protein folding, *Nature (London)* 381:571–580 (1996).

Hill, W. E. (ed.), *The Ribosome: Structure, Function, and Evolution,* American Society for Microbiology, Washington (1990).

Holtzman, E., Intracellular targeting and sorting, *BioScience* 42:608–620 (1992).

Jacobson, A., and Peltz, S. W., Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells, *Annu. Rev. Biochem.* 65:693–739 (1996).

Kawashima, T., Colominas, C. B., Wulff, M., Cusack, S., and Leberman, R., The stucture of the *Escherichia coli* EF-Tu·EF-Ts complex at 2.5Å resolution, *Nature (London)* 379:511–518 (1996).

Lieberman, K. R., and Dahlberg, A. E., Ribosome-catalyzed peptide bond formation, *Prog. Nucleic Acid Res. Mol. Biol.* 50:1–23 (1995).

Rapoport, T. A., Jungnickel, B., and Kutay, U., Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes, *Annu. Rev. Biochem.* 65:271–303 (1996).

Sandoval, I. V., and Bakke, O., Targeting of membrane proteins to endosomes and lysosomes, *Trends Cell Biol.* 4:292–297 (1994).

Scheffner, M., Nuber, U., and Huibregtse, J. M., Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade, *Nature (London)* 373:81–83 (1995).

Tate, W. P., Poole, E. S., and Mannering, S., Hidden infidelities in the translational stop signal, *Prog. Nucleic Acid Res. Mol. Biol.* 52:293–335 (1996).

Varshavsky, A., The N-end rule pathway of protein degradation, *Genes to Cells* 2:13–28 (1997).

REVIEW QUESTIONS

A. Define each of the following terms:

Ribosome cycle	Discharged tRNA
Transpeptidation	Initiation factors
Aminoacyl adenylate	Translocation
Targeting	Release factors
Signal peptide	Shine–Dalgarno sequence
Ubiquitin	PEST sequence
Cap-binding protein	Signal recognition particle

B. Differentiate between the two terms in each of the following pairs:

EF-Tu/EF-Ts	EF-T/EF-G
A site/P site	Interferons/chaperones
Peptidyl-tRNA/ initiator tRNA	Initiation codon/ 30S initiation complex
Signal peptidase/ peptidyl transferase	70S initiation complex/ 80S initiation complex

C. (1) Write out the equations, using structural formulas, for (a) activation of AA to AA-AMP, (b) conversion of AA-AMP to AA-tRNA, and (c) conversion of Met-tRNA^{fMet} to fMet-tRNA^{fMet}.
(2) Summarize the events that constitute (a) chain initiation, (b) chain elongation, and (c) chain termination.
(3) Discuss the role of ATP and GTP in translation.
(4) Why do proteins have a half-life? How is this half-life affected by the protein's structure? How does ubiquitin function in protein degradation?
(5) Describe the experiments mentioned in this chapter that helped shed light on the following: (a) ribosome cycle; (b) direction of synthesizing polypeptide chains; (c) direction of "reading" mRNA.
(6) What mechanisms exist to ensure that protein synthesis proceeds with as great a fidelity as possible? At what step in translation does each mechanism come into play?
(7) Outline the signal hypothesis.

PROBLEMS

19.1. You can desulfurize Cys-tRNA^{Cys} by treating it with finely divided nickel (Raney nickel) to yield Ala-tRNA^{Cys}. When you use this aminoacyl-tRNA in a Nirenberg-type cell-free protein-synthesizing system (see Section 16.3), will alanine be incorporated in response to mRNA codon UGU or GCA?

19.2. How could you use the inhibition of translation by puromycin (Table 19.7) to show that proteins begin to fold before synthesis of the entire chain has been completed?

19.3. A sample of bacterial culture contains 2.0 mg of ^{3}H-labeled ornithine decarboxylase. How much of the tritium-labeled enzyme will be left after 36 min, assuming that the amount of enzyme is governed by its half-life (12 min)?

19.4. Match up the following components with their corresponding functions:

(a) Transformylase	1. Binds fMet-tRNA^{fMet} and GTP
(b) IF-1	2. Binds codons UAA and UGA
(c) EF-Tu	3. Formylates Met in Met-tRNA^{fMet}
(d) RF-1	4. Binds codons UAA and UAG
(e) Peptidyl transferase	5. Binds aminoacyl-tRNA and GTP
(f) IF-2	6. Synthesizes aminoacyl adenylate

(g) EF-Ts	7. Realigns ribosomes and mRNA
(h) RF-2	8. Supports binding of IF-3
(i) Aminoacyl-tRNA synthetase	9. Displaces GDP from EF-TU • GDP
(j) IF-3	10. Binds GTP; aids binding of release factors
(k) RF-3	11. Binds to 30S subunit; aids binding of mRNA
(l) Translocase	12. Synthesizes peptide bonds

19.5 Complete oxidation of glucose under aerobic conditions yields 36 molecules of ATP per molecule of glucose (Section 12.6). Using this value, calculate the number of amino acids that could be activated to aminoacyl-tRNA per molecule of glucose catabolized. Include the necessary hydrolysis of PP_i.

19.6. Using the information in the previous problem, calculate the number of amino acids that could be polymerized in translation per molecule of glucose catabolized. Assume equivalence of GTP and ATP.

19.7. Which of the following are likely to act as competitive inhibitors in a ribosome binding assay (see Section 16.3) of Ala-$tRNA^{Ala}$? (a) Ala-$tRNA^{Cys}$; (b) Cys-$tRNA^{Ala}$; (c) Cys-$tRNA^{Cys}$.

19.8. What is the cost in terms of the number of energy-rich bonds for peptide bond formation by peptidyl transferase (Figure 19.14, going from B to C)? What is the approximate $\Delta G'$ for this reaction?

19.9. What do the antibiotics chloramphenicol and fusidic acid have in common?

19.10.* In a newly isolated bacterium, amino acids are polymerized at the rate of 30 residues per second. What is the rate at which mRNA is pulled through the ribosome?

19.11. In terms of the number of energy-rich bonds, what is the energy expenditure for synthesizing a polypeptide chain of 150 amino acids?

19.12. Assume that the polypeptide chain in the previous problem constitutes one of six identical subunits of an oligomeric protein. What is the energy expenditure, in terms of the number of energy-rich bonds, for synthesizing the entire oligomer?

19.13.* A student incubates a cell-free amino acid-incorporating system once in the presence and once in the absence of puromycin. How would the molecules of a specific protein formed in the two experiments compare with respect to (a) types of N-terminal amino acids, (b) types of C-terminal amino acids, and (c) average chain length?

19.14. Rabbit muscle contains 50 mg of compound A per 100 g of tissue. Isolating compound A takes one hour. How much muscle tissue must you start with if you want to isolate 400 mg of compound A and if, during the isolation, compound A has an effective half-life of 20 min?

19.15.* Assume that translation has an error frequency leading to the misincorporation of one out of 10,000 amino acid residues. On that basis, calculate the probability of incorporating (a) 100 and (b) 1000 amino acid residues in perfect order.

19.16. What peptide would be synthesized from the following mRNA molecule, assuming that translation begins with the first initiation codon?

5′-AUGCCAGACUUCAUGGUCCAACUGUCG
AGCGAGAUG-3′

19.17. A specific cytosine in the mRNA of the previous problem undergoes mutation by deamination to uracil. What peptide would be formed if the mutated cytosine represents residue (a) 4, (b) 12, or (c) 19?

19.18. From data given in the introduction to this chapter, compute the time it would take to synthesize a protein of 500 amino acid residues in *E. coli* and in eukaryotes.

19.19. What is common to the mechanisms of action of heme-controlled inhibitor and interferons?

19.20. In which of the following preparations do you expect the addition of fMet-$tRNA^{fMet}$ to lead to stimulation of protein synthesis? (a) Cytoplasmic ribosomes from *E. coli;* (b) cytoplasmic ribosomes from reticulocytes; (c) mitochondria from reticulocytes. Why?

19.21. What is the anticodon in a cognate tRNA if the amino acid is coded by the sequence 5′-GAC-3′ in (a) the template strand or (b) the coding strand of DNA?

19.22. What is the minimum number of binding sites that must exist on a molecule of aminoacyl-tRNA synthetase? What substance does each site bind?

19.23. You determine the number of PEST sequences for one group of proteins and the types of N-terminal amino acids for a second group of proteins and obtain the following results:

	Protein A	Protein B	Protein C
(a) Number of Pest sequences	1	7	3

	Protein X	Protein Y	Protein Z
(a) N-terminal amino acid	Ala	Phe	Ile

Based on this information, arrange the proteins in each group in the expected order of stability, from the most stable to the least stable (longest half-life to shortest half-life).

19.24. You incubate glycyl-tRNA synthetase with ^{32}P-labeled pyrophosphate (both phosphorus

atoms are labeled) and a number of other components as listed below. Which of these *in vitro* incubation mixtures will result in formation of ^{32}P-labeled ATP?

(a). $tRNA^{Gly}$, ATP, $^{32}PP_i$
(b). $tRNA^{Ile}$, ATP, $^{32}PP_i$
(c). ATP, $^{32}PP_i$
(d). Glycine, ATP, $^{32}PP_i$
(e). Isoleucine, ATP, $^{32}PP_i$

19.25.* Four arginine codons are CGU, CGC, CGA, and CGG. What are the structures of all possible anticodons for these codons? Which anticodon–codon pairings will be held together by strong hydrogen bonding and which are likely to result in slower rates of protein synthesis? Why?

19.26. What peptide is specified by the following base sequence in the template strand of DNA? Assume that translation begins with the first initiation codon.

3′-CCATGACAGTACGACTTGAACGCAATG
GGAATTACT-5′

19.27.* An investigator wishes to modify a bacterium so that a cytoplasmic protein is secreted into the extracellular medium instead. What modification of the bacterial DNA might bring about the desired result?

19.28. Write out all possible mRNA sequences that could code for the following tripeptide:

methionine-alanine-histidine

19.29.* Based on your answer to the previous problem, compute the total number of mRNA sequences that could code for the pentapeptide

methionine-alanine-histidine-leucine-valine

19.30. You wish to purify an aminoacyl-tRNA synthetase and determine whether it has a hydrolytic site at which it carries out proofreading. Would you be better off studying phenylalanyl-tRNA synthetase or glycyl-tRNA synthetase? Explain your answer.

Appendix A
Acid–Base Calculations

A.1. IONIC STRENGTH

Ionic strength (I) is a measure of ion concentration, defined by the equation

$$I = \tfrac{1}{2} \Sigma\, c_i z_i^2$$

where c_i is the concentration of ion i, z_i is the charge of ion i, and Σ is "the sum of."

Problem: What is the ionic strength of a 0.1M solution of $(NH_4)_2SO_4$?

Solution: Each "molecule" of $(NH_4)_2SO_4$ yields two NH_4^+ ions and one SO_4^{2-} ion. Hence,

$$[NH_4^+] = c_1 = 0.2M; \qquad z_1 = +1$$
$$[SO_4^{2-}] = c_2 = 0.1M; \qquad z_2 = -2$$
$$I = \tfrac{1}{2}\Sigma\{(0.2 \times 1^2) + [0.1 \times (-2)^2]\} = 0.3$$

A.2. pH

Problem 1: What is the pH of 0.010M HCl, assuming that the HCl is 100% ionized?

Solution:

$$[H^+] = 0.010M = 1.0 \times 10^{-2}M$$
$$pH = -\log[H^+] = -\log(1.0 \times 10^{-2}) = 2.0$$

Problem 2: What is the pH of 0.020M $Ca(OH)_2$, assuming that it is 100% ionized?

Solution: Since each "molecule" of $Ca(OH)_2$ yields two OH^- ions,

$$[OH^-] = 0.040M$$

and based on Eq. (1.2)

$$[H^+] = (1.0 \times 10^{-14})/0.040 = 2.5 \times 10^{-13}$$

so that

$$pH = -\log[H^+] = -\log(2.5 \times 10^{-13}) = 12.6 = 13$$

A.3. HENDERSON–HASSELBALCH EQUATION

Problem 1: Calculate the pH of a solution made by mixing 15.0 ml of $1.00 \times 10^{-1}M$ K_2HPO_4 and 25.0 ml of $2.00 \times 10^{-2}M$ KH_2PO_4.

Solution: Molar concentrations of the two salts in the final volume (15.0 + 25.0 = 40.0 ml) are:

for K_2HPO_4: $15.0 \times (1.00 \times 10^{-1})/40.0 = 0.0375M$

for KH_2PO_4: $25.0 \times (2.00 \times 10^{-2})/40.0 = 0.0125M$

Substituting these values into the Henderson–Hasselbalch equation and using $pK_a' = 7.21$ yields

$$pH = 7.21 + \log \frac{(0.0375)}{(0.0125)} = 7.69$$

Problem 2: Given that lactic acid (HA = CH_3—CHOH—COOH; A^- = CH_3—CHOH—COO^-) has a pK_a' of 3.86, calculate the percent of lactic acid present in

its dissociated form (A^-) at a pH that is one unit above its pK_a' (i.e., at pH 4.86).

Solution: At pH 4.86, the Henderson–Hasselbalch equation takes the form

$$4.86 = 3.86 + \log \frac{[A^-]}{[HA]}$$

Hence,

$$\log \frac{[A^-]}{[HA]} = 1.00$$

so that

$$\frac{[A^-]}{[HA]} = 10 = \frac{10}{1}$$

For every mole of HA per liter, there exist 10 moles of A^- per liter. The percentage of lactic acid present in the A^- form (*percent dissociation*) is

$$\%A^- = \frac{[A^-]}{[A^-] + [HA]} \times 100 = \frac{10}{(10 + 1)} \times 100 = 91\%$$

In like manner you can calculate that, at *a pH that is one unit below the pK_a' value* (pH = 2.86), 91% of the lactic acid will be present in its undissociated form (HA). You can also show that at *a pH that is two units above (or below) the pK_a' value,* the percentage of lactic acid present in the dissociated (or undissociated) form rises to 99%.

A.4. BUFFERS

Problem 1: How would you prepare 500 ml of a 0.400*M* phosphate buffer at pH 6.50 from solid KH_2PO_4 and K_2HPO_4? The second pK_a' of H_3PO_4 is 7.21.

Solution: A 0.400*M* phosphate buffer contains a total of 0.400 mole of phosphate salts per liter. Hence, letting x equal the number of moles of KH_2PO_4 per liter, it follows that $(0.400 - x)$ is the number of moles of K_2HPO_4 per liter. Substituting into the Henderson–Hasselbalch equation yields

$$6.50 = 7.21 + \log \frac{(0.400 - x)}{x}$$

$$\log \frac{(0.400 - x)}{x} = -0.71$$

$$\frac{(0.400 - x)}{x} = 0.195$$

$$x = 0.335 \text{ mol of } KH_2PO_4\text{/liter}$$

$$(0.400 - x) = 0.065 \text{ mol of } K_2HPO_4\text{/liter}$$

Since you wish to prepare not 1 liter but only 500 ml, you will need one-half of the above number of moles, namely, 0.168 mol of KH_2PO_4 (22.8 g) and 0.033 mol of K_2HPO_4 (5.75 g). To prepare the buffer, you would weigh out these amounts, dissolve them in water, and dilute the solution to a final volume of 500 ml.

Problem 2: How would you prepare 500 ml of a 0.100*M* acetate buffer at pH 5.00 from acetic acid and NaOH? The pK_a' of acetic acid is 4.76.

Solution: Substituting into the Henderson–Hasselbalch equation yields

$$5.00 = 4.76 + \log \frac{[A^-]}{[HA]}$$

Hence,

$$\log \frac{[A^-]}{[HA]} = 0.24$$

$$\frac{[A^-]}{[HA]} = 1.74$$

Since you wish to prepare a 0.100*M* buffer, the total, combined concentration of A^- and HA must be 0.100*M*, that is,

$$[A^-] + [HA] = 0.100$$

Each of the last two equations has two unknowns, $[A^-]$ and [HA]. You can solve two equations with two unknowns simultaneously. From the first equation, you have that

$$[A^-] = 1.74[HA]$$

Substituting for $[A^-]$ into the second equation yields

$$1.74[HA] + [HA] = 2.74[HA] = 0.100$$

so that

$$[HA] = 0.0365M \text{ (0.0365 mol/liter)}$$
$$[A^-] = 1.74[HA] = 0.0635M \text{ (0.0635 mol/liter)}$$

As you wish to prepare only 500 ml, not 1 liter, you will need one-half of the above number of moles, namely, 0.0182 mol of HA (acetic acid) and 0.0318 mol of A^- (acetate). Since both buffer components are derived from acetic acid, you must start with 0.0500 mol of acetic acid and convert some of it to acetate by adding NaOH. From the neutralization reaction involved,

$$CH_3COOH + NaOH \rightleftarrows CH_3COO^-Na^+ + H_2O$$

you see that adding 1.0 mol of NaOH converts 1.0 mol of acetic acid to 1.0 mol of acetate. To produce 0.0318 mol of acetate, you must add 0.0318 mol of NaOH. After adding the NaOH, you will be left with the required amount of acetic acid:

$$0.0500 - 0.0318 = 0.0182 \text{ mol}$$

To sum up, you would prepare this buffer by obtaining 0.0500 mol of acetic acid (2.87 ml of concentrated acetic acid, which is 17.4*M*), adding 0.0318 mol of solid NaOH (1.27 g), and diluting the mixture to 500 ml.

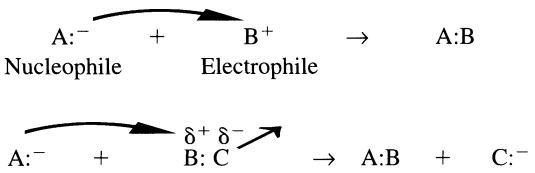
Appendix B
Principles of Organic Chemistry

B.1. FUNCTIONAL GROUPS

Functional groups consist of two or more atoms and possess characteristic structures and chemical reactivities. A given functional group generally behaves the same way in all molecules containing that group. Table B.1 shows some of the functional groups found in biomolecules.

B.2. POLAR REACTIONS

Polar reactions result from the attractive force between positive and negative charges (or partial charges) on molecules. We call the two reactants in a polar reaction **nucleophile** and **electrophile.** A nucleophile consists of an atom or a group of atoms that has an electron-rich site and forms a bond by *donating a pair of electrons.* By contrast, an electrophile consists of an atom or a group of atoms that has an electron-poor site and forms a bond by *accepting a pair of electrons.* We depict the electron-pair movement by means of a curved arrow, using the convention that the electron pair moves from the tail to the head of the arrow:



The second arrow in the second reaction indicates that C leaves, taking the two electrons of the B—C bond with it.

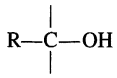
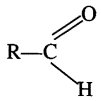
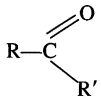
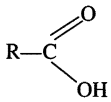
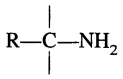
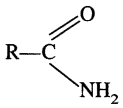
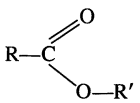
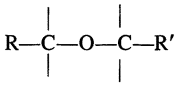
B.3. CHIRALITY

Many objects are asymmetric in their structure; they have a "handedness" like that of the left and right hands. If you visualize your two hands placed on either side of a flat mirror, one hand will appear to be the *mirror image* of the other. But the two hands are not identical. You cannot put one hand on top of the other, with both palms down. *The two hands cannot be superimposed in space.*

In much the same manner, biomolecules can have structural asymmetry, resulting in mirror images that cannot be superimposed in space. We use the term **chirality** ("handedness") to refer to the right- and left-handedness of a molecule. At the molecular level, chirality arises when a compound contains one or more **chiral centers.** A chiral center comprises either a **chiral carbon atom** or some other asymmetric region in the molecule. A chiral carbon atom has four different substituents attached to it (Figure B.1). Because tetravalent carbon is tetrahedral, these groups occupy the corners of a tetrahedron. Because of carbon's tetrahedral nature, the two mirror images of a chiral carbon cannot be superimposed.

Aside from carbon, several other atoms that form compounds having a tetrahedral structure (Si, N, P, S) can exist as chiral centers under proper circumstances. Chiral centers also result from molecular asymmetry that is not due to the presence of chiral atoms. The helical structures of proteins and nucleic acids provide an example. A helix is intrinsically chiral; a left-handed helix constitutes a nonidentical and nonsuperimposable mirror image of a right-handed helix, much as a left-handed screw differs from a right-handed screw.

Table B.1. Some Common Functional Groups in Biomolecules

Compound type	Structure	Functional group
Alcohol		Hydroxyl
Aldehyde		Carbonyl
Ketone		Carbonyl
Acid		Carboxyl
Amine		Amino
Amide		Amide
Ester		Ester
Ether		Ether
Thiol	R—S—H	Sulfhydryl
Disulfide	R—S—S—R′	Disulfide

B.4. OPTICAL ISOMERISM

Chiral compounds are optically active; they exhibit **optical rotation.** When plane-polarized light is passed through solutions of chiral compounds, they rotate the plane of polarization. Most optically active compounds encountered in biochemistry owe their activity to the presence of one or more chiral carbons. We describe optical rotation in terms of a **specific rotation,** [α], usually measured at 25°C, using light of the sodium D line (589.3 nm):

$$[\alpha]_D^{25} = \frac{100\alpha}{dc}$$

where α is the observed rotation in degrees, d is the optical path length through the solution (in decimeters), and c is the concentration of the solution (in grams per 100 ml).

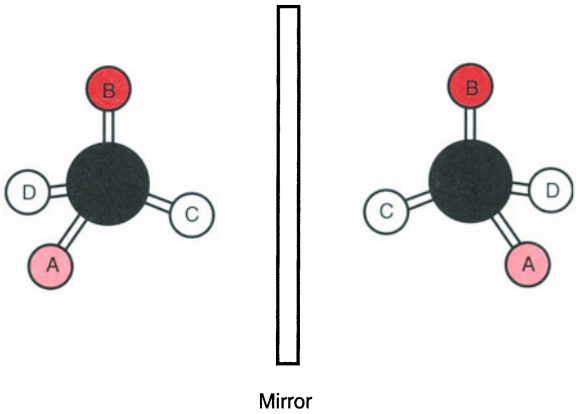
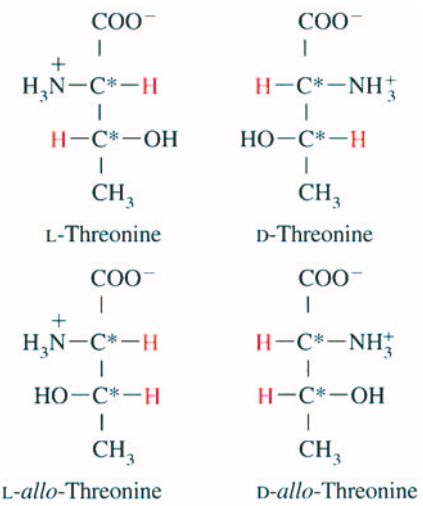


Figure B.1. The two enantiomers (mirror images) of a chiral carbon atom.

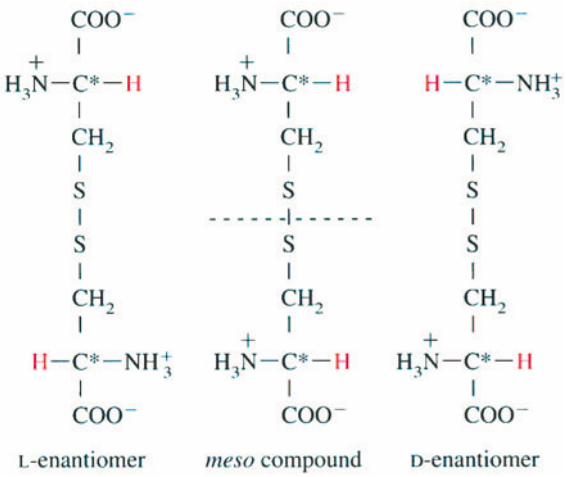
The two mirror images of a compound containing a chiral carbon, called **enantiomers,** differ in their optical rotation; they represent **optical isomers.** If we separate an equimolar mixture of the two enantiomers, we find that one enantiomer rotates the plane of plane-polarized light to the right; it is **dextrorotatory** and designated (+). The other enantiomer rotates the plane the same number of degrees but to the left; it is **levorotatory** and designated (−). An equimolar mixture of the two enantiomers has a net zero optical rotation and is called a **racemic mixture.**

In addition to enantiomers, two other types of optical isomers can occur—diastereomers and *meso* compounds. **Diastereomers** constitute optical isomers of a compound that are *not* mirror images. For example, the amino acid *threonine* has two chiral carbons(*) and a total of four optical isomers:



Of these, L- and D-threonine are enantiomers, and so are L-*allo*- and D-*allo*-threonine. However, L- and L-*allo*-threonine are diastereomers, and so are D- and D-*allo*-threonine.

A ***meso* compound** contains two or more chiral carbon atoms but has no optical activity because the molecule possesses a plane of symmetry. For example, the biochemically important structure formed by linking two molecules of the amino acid *cysteine* via a *disulfide bond* (often referred to as *cystine*) can exist in the form of three isomers:



L-enantiomer *meso* compound D-enantiomer

The *meso* compound has a plane of symmetry that passes through the disulfide bond, at right angles to the axis of this bond. Because of the plane's bisection, one part of the *meso* compound is identical to half of the levorotatory enantiomer, while the other is identical to half of the dextrorotatory enantiomer. Thus, half of the *meso* compound tends to rotate the plane of plane-polarized light to the right, while the other half tends to rotate it the same number of degrees to the left. The two parts produce an internal compensation, making the molecule, as a whole, optically inactive.

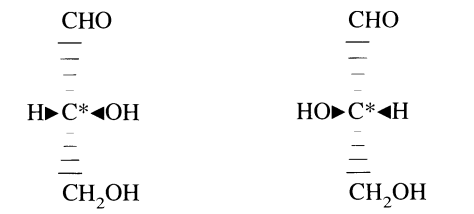
B.5. CONFORMATION AND CONFIGURATION

Conformation refers to the spatial arrangement of atoms in a molecule resulting from their freedom of rotation about single bonds. Production of different conformations involves no change in the covalent structure of the molecule; the change from one conformation to another does not require breaking and remaking of covalent bonds. When we speak of the conformation of a molecule, we are, in reality, dealing with an entire family of structures, differing from one another by one or more slight rotations of atoms about single bonds. A single, specific conformational form, or **conformer,** cannot be isolated.

Configuration refers to a unique and fixed spatial arrangement of atoms in a molecule such that the molecule can be isolated in that particular stereochemical form. The change from one configuration to another requires breaking covalent bonds and re-forming them in a different sense. The covalent structure of the molecule is changed as one configuration is replaced by another. Different configurational forms constitute **stereoisomers.**

B.6. D, L AND *R, S* SYSTEMS

We assign specific configurations to enantiomers by reference to *glyceraldehyde,* $CHO-CHOH-CH_2OH$. This compound was chosen as a standard because chemists were able to isolate its two enantiomers in pure form and to determine the precise configuration of each enantiomer. The dextrorotatory enantiomer was designated D, and the levorotatory enantiomer L. In a two-dimensional projection, in which we write the structure of glyceraldehyde as shown below, *the D-isomer has the H of the chiral carbon on the left.* In the **D, L system,** we designate optically active compounds, such as amino acids and carbohydrates, D or L by matching up their structures with that of glyceraldehyde.

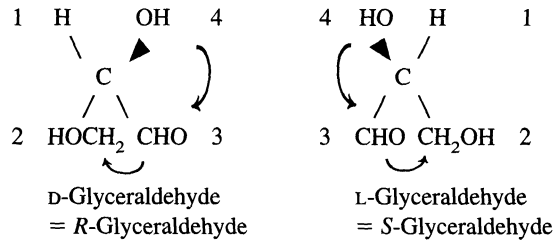


D-Glyceraldehyde L-Glyceraldehyde

For molecules other than glyceraldehyde, there exists no necessary correspondence between *configuration* (D or L) and *optical rotation* (+ or −). Originally, the terms D and L were meant to indicate the direction of rotation, dextrorotatory and levorotatory. Moreover, D-glyceraldehyde and many other D-monosaccharides are indeed dextrorotatory, D(+). However, many other D-compounds are levorotatory, D(−). Likewise, there occur both levorotatory L-compounds, L(−), and dextrorotatory L-compounds, L(+). The magnitude and direction of optical rotation are complicated functions of the electronic structure surrounding the chiral center.

The need for a reference compound, as in the D, L system, is eliminated when we use the ***R*, *S* system.** By means of this system, we can assign absolute configurations to any compound by examining its stereochemical structure.

In the *R*, *S* system, we assign a *priority* to each group attached to a chiral carbon by using a set of rules. We then examine the structural formula with the group of lowest priority facing away from the viewer. If the priority of the remaining three groups *decreases* in a *clockwise* order, we designate the absolute configuration as *R* (from the Latin *rectus,* meaning "right"). If the priority of the groups decreases in a counterclockwise order, we designate the configuration as *S* (from the Latin *sinister,* meaning "left"). Let us illustrate the convention for glyceraldehyde:



D-Glyceraldehyde = *R*-Glyceraldehyde

L-Glyceraldehyde = *S*-Glyceraldehyde

Despite the advantages of the *R*, *S* system, the D, L terminology is still commonly used in biochemistry and is employed in this book.

Appendix C
Tools of Biochemistry

C.1. SPECTROPHOTOMETRY

Many compounds of biological interest absorb light in the ultraviolet, visible, or near-infrared regions of the spectrum. When radiation is absorbed by a substance, the energy of the radiation raises the substance from one energy level to a higher one; the substance undergoes a *transition*. Different wavelengths of radiation have different energies associated with them and cause different types of transitions. We determine absorption of light by passing it through a solution and measuring the incident and transmitted intensities in a *spectrophotometer*. These measurements yield the **absorbance (A),** defined by **Beer's law:**

$$A = \log(I_0/I) = Elc$$

where I_0 is the intensity of the incident light, I is the intensity of the transmitted light, c is the concentration of the absorbing substance, l is the length of the light path through the solution, and E is the extinction coefficient.

The **extinction coefficient** represents the absorbance of a solution when the concentration and the length of the light path are both unity. Because absorbance is a dimensionless quantity, the units of the extinction coefficient depend on l and c and must be such that Elc has no units. If c represents a molar concentration, we call E a *molar extinction coefficient* (ϵ). We can also express absorption of light in terms of the fraction or the percent of incident light transmitted, called **transmittance (T)** or **percent transmittance (%T),** respectively:

$$T = I/I_0 \qquad \%T = (I/I_0) \times 100$$

Typical spectrophotometric assays require construction of a **standard curve**—a plot of absorbance as a function of concentration for several standard solutions of a compound. Drawing the line of best fit through the points allows us to determine the concentration of an unknown from its absorbance and the standard curve.

C.2. CHROMATOGRAPHY

Chromatography comprises a group of methods for separating complex mixtures of molecules based on their *repetitive distribution* between a *mobile* and a *stationary phase.* Distribution of molecules between the two phases is governed by one or more of four basic processes: *adsorption, ion exchange, partitioning,* and *gel filtration.* Movement of the mobile phase results in a differential migration, or *resolution,* of the molecules along the stationary phase.

C.2.1. Adsorption Chromatography

In **adsorption chromatography,** the stationary phase consists of a solid, typically in the form of a *column* (Figure C.1), and the mobile phase is an aqueous or nonaqueous solution. We apply the sample to the top of the column, *develop* the column by passing a suitable liquid through it, and collect the *effluent,* frequently by means of an automatic *fraction collector.* Components are *eluted* when they emerge in the effluent. As the material moves down the column, adsorption and desorption occur repeatedly. Adsorption involves van der Waals forces, hydrogen bonds, and hydrophobic interactions. The rate of movement of a substance through the column depends on

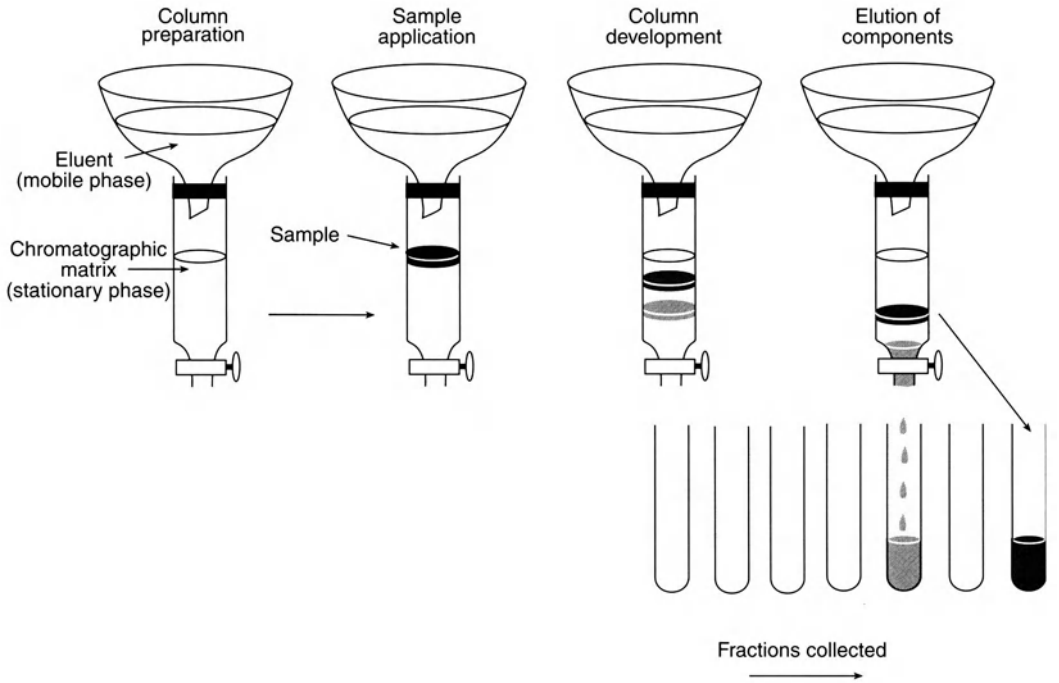


Figure C.1. Column chromatography. Concentration in the effluent is measured as a function of fraction number or elution volume.

the degree of adsorption of the substance to the stationary phase. A substance that adheres more strongly to the column will move more slowly through it.

C.2.2. Ion-Exchange Chromatography

In **ion-exchange chromatography,** the stationary phase (typically, a column) consists of an *ion-exchange resin,* a cross-linked polymer having many charged functional groups. A *cation-exchange resin* has a negative charge and binds cations; an *anion-exchange resin* has a positive charge and binds anions. The mobile phase consists of an aqueous solution. Ions in the sample bind electrostatically to oppositely charged groups of the resin (Figure C.2). Subsequently, we disrupt these ionic bonds by changing the pH and/or ionic strength of the eluting buffer. Binding

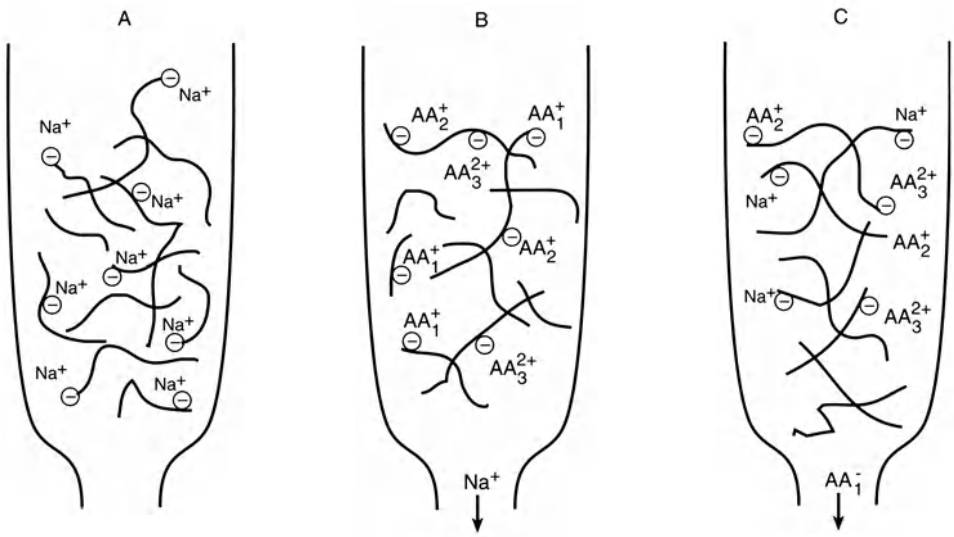


Figure C.2. Ion-exchange chromatography. (A) A cation-exchange resin with bound Na^+ ions. (B) Amino acid cations replace Na^+ on the column. (C) Changing the pH and/or the ionic strength of the eluting buffer alters the charges of the bound amino acid ions and leads to their elution.

and desorption of ions occur repeatedly as the material moves down the column. The extent to which ions are retarded in their movement through the column depends on the sign and magnitude of their charge.

C.2.3. Partition Chromatography

In **partition chromatography,** molecules distribute themselves repeatedly between two immiscible phases according to their solubilities in these phases. In **paper chromatography** (Figure C.3), the stationary phase comprises a polar layer of water molecules bound to the hydroxyl groups of cellulose in the paper. The mobile phase usually consists of a relatively nonpolar organic solvent that moves over the stationary aqueous phase. In such a system, nonpolar molecules move faster along the paper than polar ones. In *descending chromatography,* material moves down the filter paper by gravity; in *ascending chromatography,* it moves up by capillary action.

We characterize the movement of substances by R_f **values,** which are generally proportional to solubilities in the mobile phase. An R_f value is the ratio of the distance moved by the sample to that moved by the solvent, both measured from the point of sample application, or *origin:*

$$R_f = \frac{\text{distance moved by sample}}{\text{distance moved by solvent}}$$

C.2.4. Gel-Filtration Chromatography

In **gel-filtration chromatography,** the stationary phase consists of spherical gel particles (*beads*) of controlled size and porosity, formed from cross-linked polymers (Figure C.4). Molecules are fractionated on the basis of their size and shape, two properties that determine the rate of diffusion. Smaller molecules diffuse faster than larger ones of the same general shape. Likewise, more compact and globular molecules diffuse more rapidly than elongated, asymmetric ones of the same size. The more extensive is the penetration of molecules into the beads, the more retarded their movement through the column and the greater the *elution volume.* We can use gel filtration for molecular weight determinations since, for proteins of the same general shape and for a given gel, a linear relationship exists between the elution volume and the logarithm of the molecular weight.

A special type of chromatography, called **affinity chromatography,** relies on the specific noncovalent binding (by adsorption and/or ion exchange) of a ligand to the molecule of interest (Figure C.5). The ligand is linked covalently to a column, and the technique leads to extensive purification of the molecule of interest in a single step.

Frequently, researchers perform chromatographic separations using special columns through which the mobile phase is forced under high pressure (5000–10,000 psi). Such **high-performance liquid chromatography (HPLC)** requires small amounts of sample, is rapid, and achieves high resolution. Typically, peaks are much sharper than those shown in Figure 2.6.

C.3. CENTRIFUGATION

We can separate macromolecules in solution by subjecting them to a *centrifugal force,* generated by spinning a *rotor* about its axis. The centrifugal force causes macromolecules to **sediment** through the solution and leads to their separation based on differences in molecular size, shape, and density. We use *preparative centrifuges* with *fixed-angle rotors* primarily for purification and fractionation of macromolecules. *Analytical ultracentrifuges* are equipped with optical systems that allow photographs and measurements to be made during sedimentation (Figure C.6).

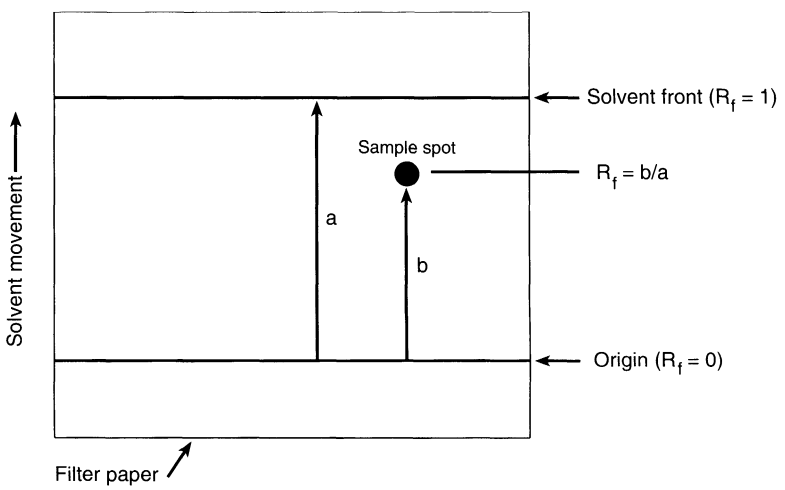


Figure C.3. Principle of paper chromatography. Typically, ascending chromatography employs a filter paper cylinder placed in a glass jar having solvent at the bottom. In descending chromatography, a filter paper sheet is hung from a solvent trough in a rectangular glass chamber.

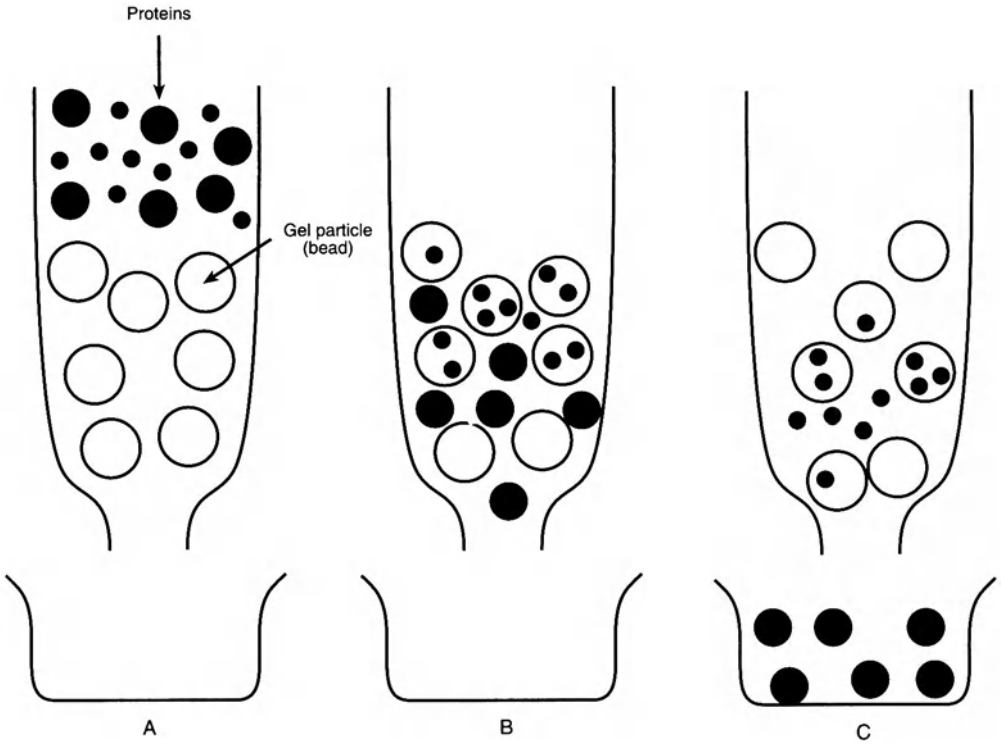


Figure C.4. Gel filtration. Smaller molecules penetrate the gel beads and are retarded; larger molecules are eluted first.

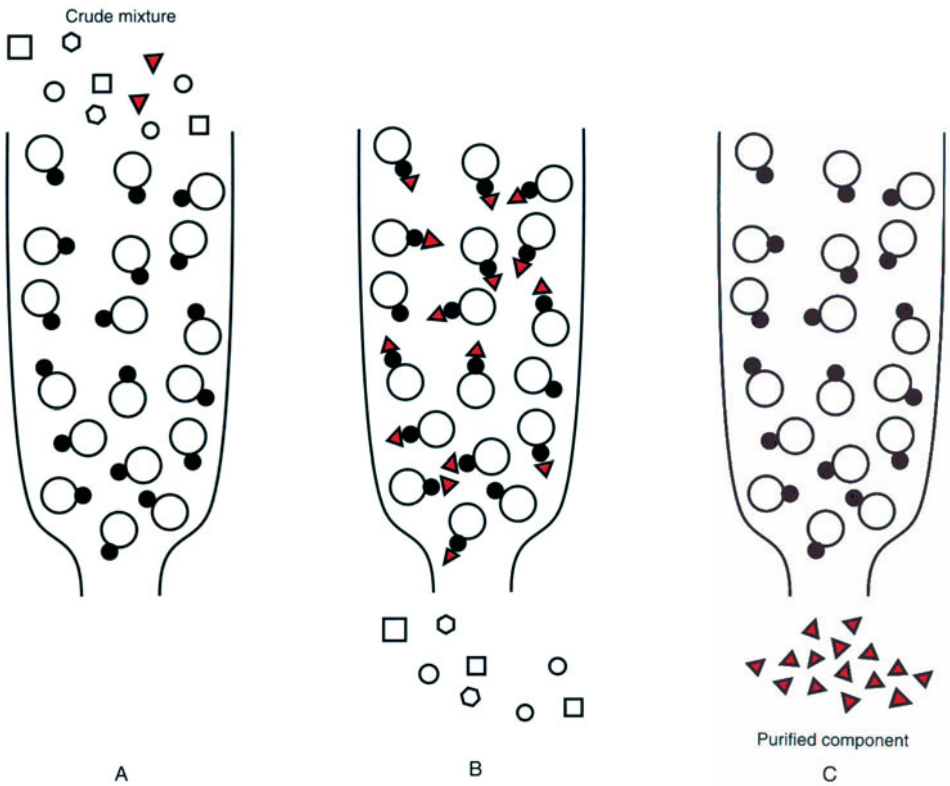


Figure C.5. Principle of affinity chromatography. A ligand, linked covalently to a column, binds a specific protein. Changing the eluting buffer disrupts the ligand–protein interactions, yielding a highly purified protein preparation.

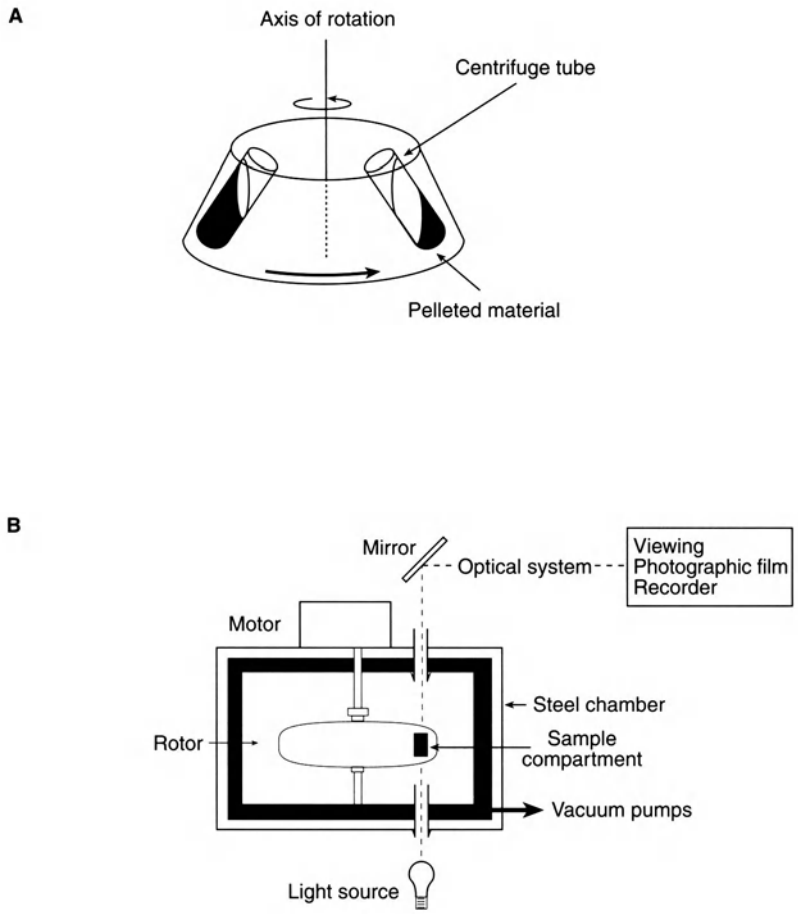


Figure C.6. Centrifugation. (A) A fixed-angle rotor; (B) diagram of an analytical ultracentrifuge.

C.3.1. Analytical Ultracentrifugation

In the analytical ultracentrifuge, each sedimenting component forms a *boundary,* a transition between solvent and solution or between two solutions. An optical system detects the boundary and causes it to appear as a peak on a photograph (Figure C.7). We measure the movement of this peak as a function of time by determining its position on successive photographs. From these measurements, we can determine the **sedimentation coefficient** (lower case *s*). A sedimentation coefficient—the velocity per unit of centrifugal field—is expressed in terms of **Svedbergs** (capital S). One Svedberg equals 10^{-13} s. We can also use analytical ultracentrifugation to determine molecular weights (M) of sedimenting particles by means of the **Svedberg equation:**

$$M = (RTs)/D(1 - \bar{V}\rho)$$

where R is the gas constant, T is the absolute temperature, s is the sedimentation coefficient, D is the diffusion coefficient, $\bar{V}$ is the partial specific volume of the particle (ml/g), and ρ is the density of the solution (g/ml).

C.3.2. Density Gradient Centrifugation

In **density gradient centrifugation,** sedimenting particles move through a density gradient, a liquid column in which the density increases from top to bottom. When we perform density gradient centrifugation in the analytical ultracentrifuge, we dissolve the sample in a dense salt solution. The salt distributes itself during centrifugation and thereby sets up a density gradient. Macromolecules of a given type band in this gradient at a position where their density equals that of the gradient. We determine the density of material in a band from the band's position. When we perform density gradient centrifugation in a preparative ultracentrifuge, we place the sample on top of a preformed density gradient, prepared by successively layering solutions of different densities in a centrifuge tube. We then place the tubes in a *swinging-bucket rotor,* where they swing out during centrifugation and line up at right angles to the axis of rotation (Figure C.8).

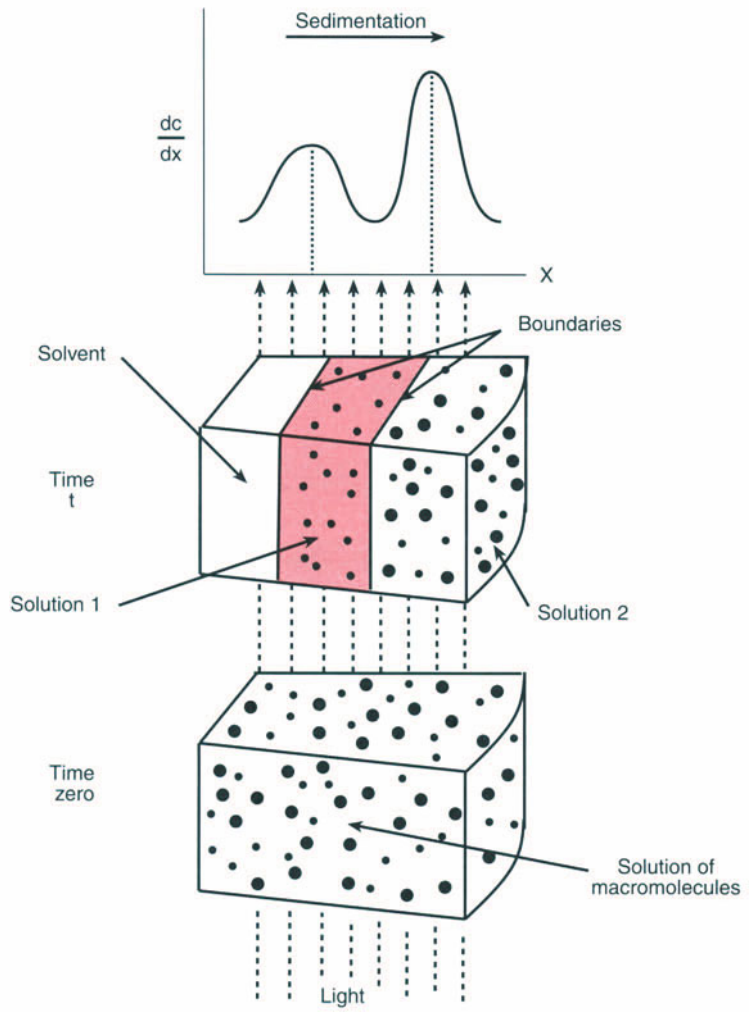


Figure C.7. Formation of ultracentrifuge patterns by light passing through a solution of sedimenting particles. Photographs provide plots of concentration gradient (dc/dx) as a function of distance (x) from the center of rotation, and boundaries appears as peaks.

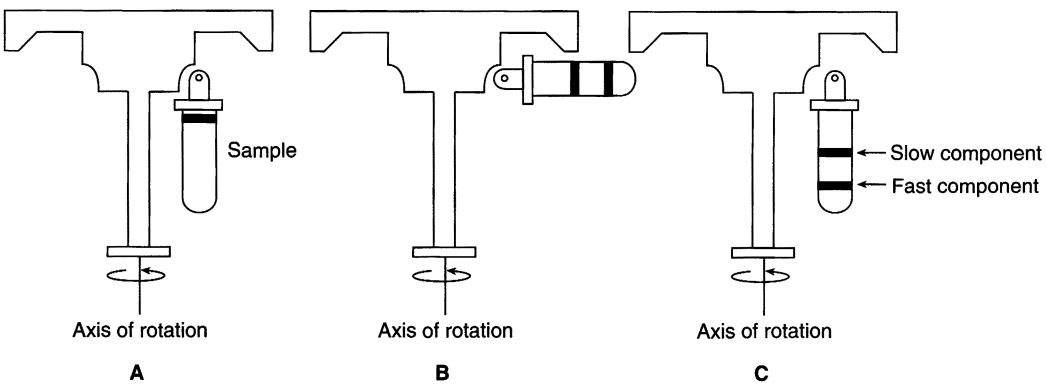


Figure C.8. Density gradient centrifugation with a swinging-bucket rotor. (A) Loading of sample on top of the gradient. (B) Components separate during centrifugation. (C) The gradient prevents mixing by convection at the end of the run. Tubes can be punctured at the bottom, and fractions collected and analyzed.

C.4. ELECTROPHORESIS

Electrophoresis represents the movement of a charged particle in an electric field. The velocity, v, of the particle is given by

$$v = Eq/f$$

where E is the electric field (V/cm), q is the net charge of the particle, and f is the frictional coefficient.

The **frictional coefficient** depends on particle size and shape. If E is constant, the velocity becomes a function of q/f, the *charge-to-mass ratio* of the particle. Of two particles having the same mass and shape, the one with greater net charge (q) will move faster in electrophoresis. Of two particles having the same mass and net charge, the one that is more more globular rather than elongated, and hence has a smaller f value, will move faster. We define **electrophoretic mobility** (u) as the velocity per unit of electric field:

$$u = v/E = q/f$$

We refer to the major type of electrophoresis employed in biochemical research as *zone electrophoresis.* Researchers perform it by placing a small aliquot of solution in contact with some support medium such as a sheet of filter paper (*paper electrophoresis*), a strip of cellulose acetate (*cellulose acetate electrophoresis*), or a slab or column of a gel (*gel electrophoresis*). Upon application of an electric field, sample components migrate as spots or zones (Figure C.9). **Polyacrylamide gel electrophoresis (PAGE)** is currently the most popular type of zone electrophoresis and utilizes gels made by polymerizing and cross-linking acrylamide ($CH_2{=}CH{-}CO{-}NH_2$). Investigators control pore size in the resultant network by varying the concentration of acrylamide. Biochemists use polyacrylamide gel electrophoresis for sequencing nucleic acids and for analysis of proteins, as described below.

In **disc gel electrophoresis,** *disc*continuities of pH, ionic strength, buffer composition, and gel concentration are built into the gel column. Such gels permit the separation of a large number of closely related proteins with high resolution and require only small amounts of sample. We layer the sample on top of the gel column and immerse the two ends of the column in buffer compartments containing electrodes. After electrophoresis, we stain the gel, and proteins appear as discrete thin bands.

Biochemists use **sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)** for determining molecular weights of individual polypeptide chains. We first dissociate oligomeric proteins into monomers by treatment with sodium dodecyl sulfate (SDS), a detergent that breaks hydrophobic bonds, and mercaptoethanol, a compound that breaks disulfide bonds. Under these conditions, proteins yield polypeptide chains that form random coils and that bind large amounts of SDS. The charges of the SDS–protein com-

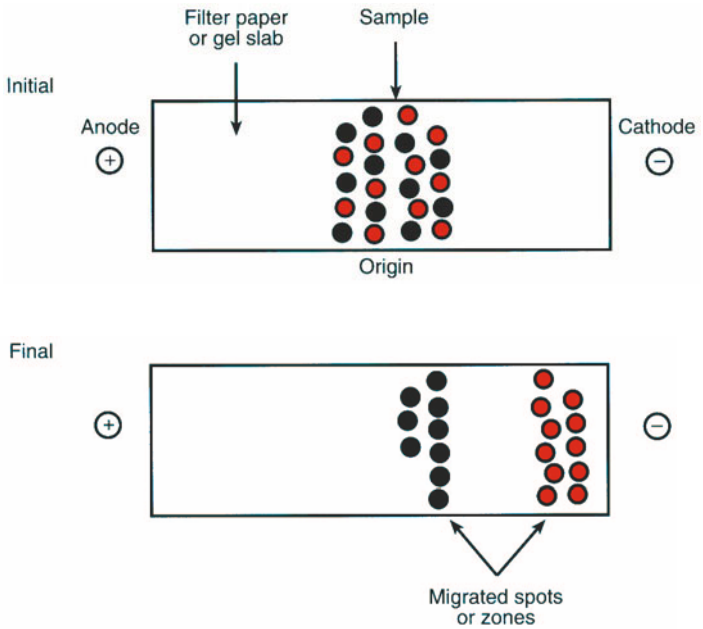


Figure C.9. Zone electrophoresis. In the illustration, both components carry a net positive charge at the pH of the experiment.

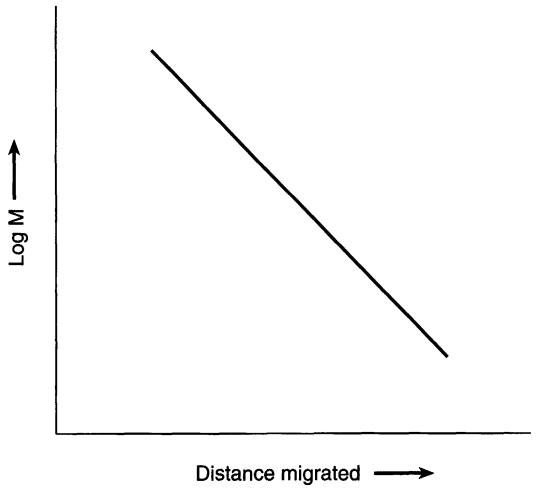


Figure C.10. Determination of molecular weights (M) by SDS-PAGE.

plexes are effectively due to the bound SDS alone since we minimize protein charges by performing the reaction at pH 7. Because all proteins bind essentially the same amount of SDS per gram of protein, all SDS–protein complexes have essentially the same charge/mass ratios and therefore the same electrophoretic mobility. When we subject these complexes to electrophoresis on a polyacrylamide gel, they are separated by the sieving action of the gel. The gel retards large molecules more than smaller ones so that large molecules move more slowly through the network. Migration distances are inversely proportional to the logarithm of the molecular weight (Figure C.10). By calibrating the gel, using migration distances of standards, we can determine molecular weights of unknown proteins.

C.5. RADIOACTIVITY

Radioactive isotopes are atoms that have unstable nuclei and that disintegrate spontaneously to produce nuclei of different elements. Radioactive **disintegration,** or **decay,** is accompanied by emission of one or more types of ionizing radiation such as *α particles* (helium nuclei), *β particles* (electrons originating in the nucleus), or *γ rays* (high-energy photons originating in the nucleus). Radioactive decay constitutes an exponential process:

$$N = N_0 e^{-\lambda t}$$

where N is the number of radioactive atoms present at time t, N_0 is the number of radioactive atoms present at time zero ($t = 0$), λ is a decay constant characteristic of the isotope, and t is time.

Table C.1. Some Isotopes of Biochemical Importance

Isotope	Radiation emitted	Half-life
^{3}H	β	12.3 years
^{14}C	β	5568 years
^{24}Na	β, γ	15 hours
^{32}P	β	14.2 days
^{35}S	β	87.1 days

The **half-life** ($t_{1/2}$) is the time required for the activity of a population of radioactive atoms to decrease by one-half ($N = N_0/2$) and can be computed from

$$t_{1/2} = \ln 2/\lambda = 0.693/\lambda$$

Table C.1 lists half-lives of some isotopes commonly used in biochemical research. Isotopes differ in the energy of the emitted particle. Thus, tritium (^{3}H) and ^{32}P represent weak and strong β emitters, respectively.

The basic unit of radioactivity, the **curie (Ci),** equals 2.2×10^{12} **disintegrations per minute (dpm).** Because we do not normally detect every disintegration by the instrument used, we call events actually measured **counts per minute (cpm).** We measure radioactivity by means of three basic methods: autoradiography, Geiger–Müller counting, and liquid scintillation counting.

In **autoradiography,** we expose films with special emulsions, by means of various techniques, to the radiation emitted by radioactively labeled compounds. Exposed silver halide grains in the film indicate locations of the radioactive isotope in the sample. In **Geiger–Müller counting,** β particles emitted by radioactive isotopes cause ionization in a gas by ejecting electrons from gas atoms, thereby producing *ion pairs* (ejected electron plus residual cation). The electrons and cations produced are collected by electrodes and are recorded as a pulse of charge or current by means of a *Geiger counter.* In **liquid scintillation counting,** β particles emitted by radioactive isotopes excite solvent molecules that emit photons as they return to the ground state. Added compounds, called *fluors,* absorb these photons and re-emit the absorbed energy in the form of photons of longer wavelength by *fluorescence.* Fluorescence of the fluors produces flashes of light, or *scintillations,* which are converted to electrical pulses and counted.

C.6. RECOMBINANT DNA TECHNOLOGY

Recombinant DNA technology , or **genetic engineering,** comprises various techniques by which researchers carry

out genetic recombination *in vitro.* This technology, initiated in the early 1970s, includes breaking and rejoining of DNA molecules from different organisms and production and isolation of modified DNA or fragments thereof. An example of recombinant DNA technology is the insertion of a human gene coding for insulin into the DNA of a bacterial plasmid followed by cloning of the plasmid to produce many identical copies of the inserted gene. The modified plasmid DNA is called a **recombinant DNA.** Recombinant DNA technology generally involves the following major steps.

C.6.1. Selection of Target DNA Fragment

Researchers usually produce a specific DNA fragment containing the gene(s) of interest by means of *restriction enzymes,* preferably those that lead to formation of *staggered* cuts. The *cohesive ends* thus formed are useful for annealing this DNA fragment to some other DNA molecule, produced with the same restriction enzymes. Investigators then link the annealed DNAs covalently by means of *DNA ligase* (Figure C.11).

Investigators can also generate cohesive ends synthetically by using the enzyme *terminal deoxynucleotidyl transferase.* This enzyme catalyzes addition of a tail of nucleotides to the 3′-end of DNA and does not require a template. Thus, we can add a tail of poly(C) to one DNA and a tail of poly(G) to another, anneal the two DNAs via their synthetic cohesive ends, and link them covalently by means of DNA ligase (Figure C.12).

C.6.2. Production of Recombinant DNA

The target DNA fragment, or **passenger,** is linked to some other DNA molecule, termed a **vector.** Linking may be accomplished by annealing the cohesive ends of the two molecules, followed by action of DNA ligase. Alternatively, researchers employ other methods to **splice** the passenger to its vector. Any molecule that has a replication origin and that can replicate after it enters a suitable cell

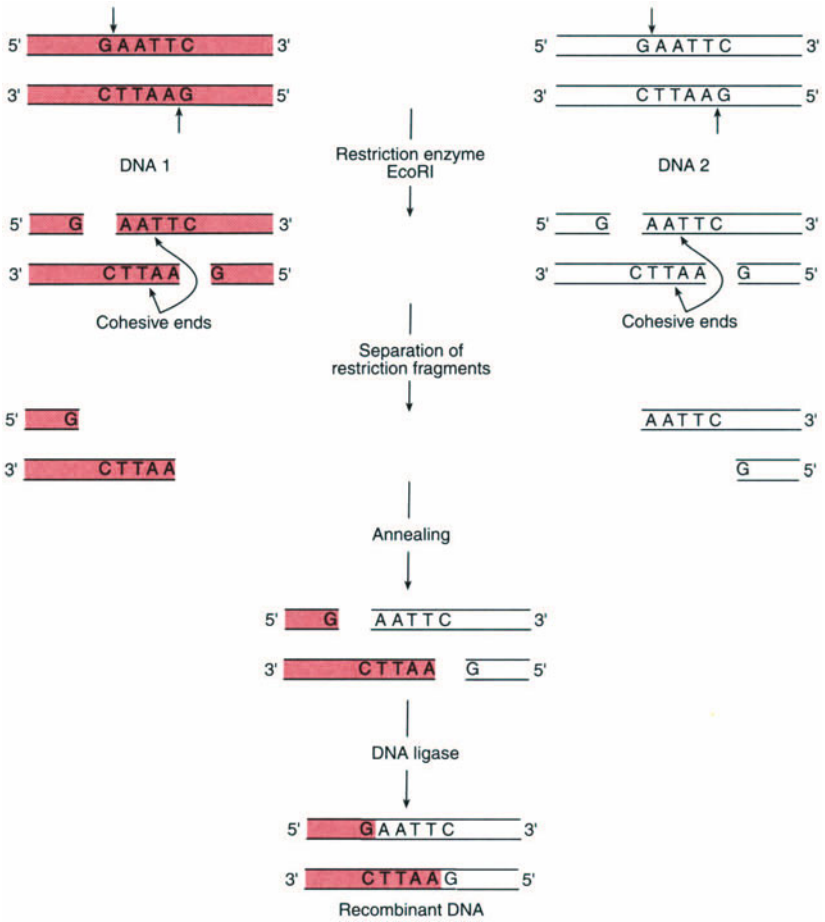


Figure C.11. Production of recombinant DNA by means of restriction enzymes.

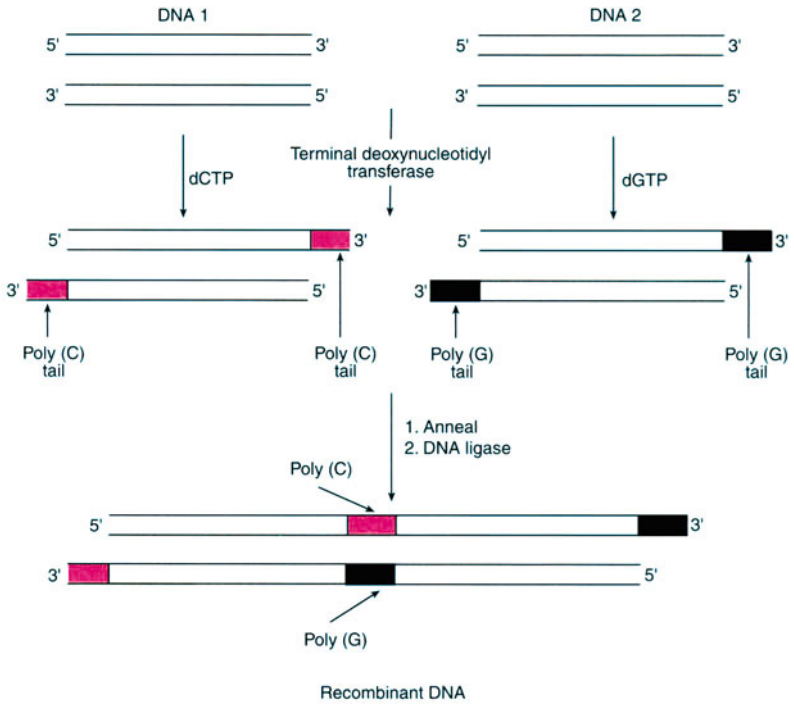


Figure C.12. Production of recombinant DNA by means of terminal deoxynucleotidyl transferase.

(e.g., plasmids or phage DNA) can be used as a vector (Figure C.13).

Plasmids, extrachromosomal genetic elements of bacteria, are circular double-stranded DNA molecules that replicate independently of the bacterial chromosome. Plasmids generally confer some evolutionary advantage onto their host cell. Thus, a plasmid may carry a marker for antibiotic resistance or for production of **colicins,** proteins that are bactericidal to other strains. A bacterium containing such a plasmid shows resistance to certain antibiotics or a capacity to produce colicins and protect itself against other bacteria.

C.6.3. Insertion of Recombinant DNA into Host Cells

Plasmids and bacteriophages constitute suitable vectors for transferring target DNA to host bacterial cells. Biochemists use other systems for insertion of recombinant DNAs into eukaryotic cells. Because recombinant DNA self-replicates independently of the host cell chromosome, many replications of the vector and its passenger can occur for any single replication of chromosomal DNA. This provides for extensive **amplification** of the target DNA. Production of large numbers of copies of target DNA in any single cell constitutes **molecular cloning.** In biology, a **clone** defines a group of genetically identical organisms, derived from a common ancestor. In molecular biology, a clone defines a population of identical molecules, derived by replication of a common ancestor.

C.6.4. Selection of Cells Containing Cloned DNA

Since not all cells in a culture contain cloned DNA, cells containing such DNA must be isolated by screening the entire culture. In the case of bacteria, screening can be done by using a plasmid that confers suitable properties on the host cell. The plasmid may, for example, confer resistance to specific antibiotics or eliminate the need for specific nutritional requirements. Cells containing the cloned gene can then be detected by growing the entire culture in the presence of the particular antibiotic or in the absence of the particular nutrient. Another method of screening consists of treating the bacterial culture with antibodies against the product of the cloned gene.

C.6.5. Polymerase Chain Reaction

The **polymerase chain reaction (PCR)** was introduced in 1987 and permits amplification of nucleic acid sequences

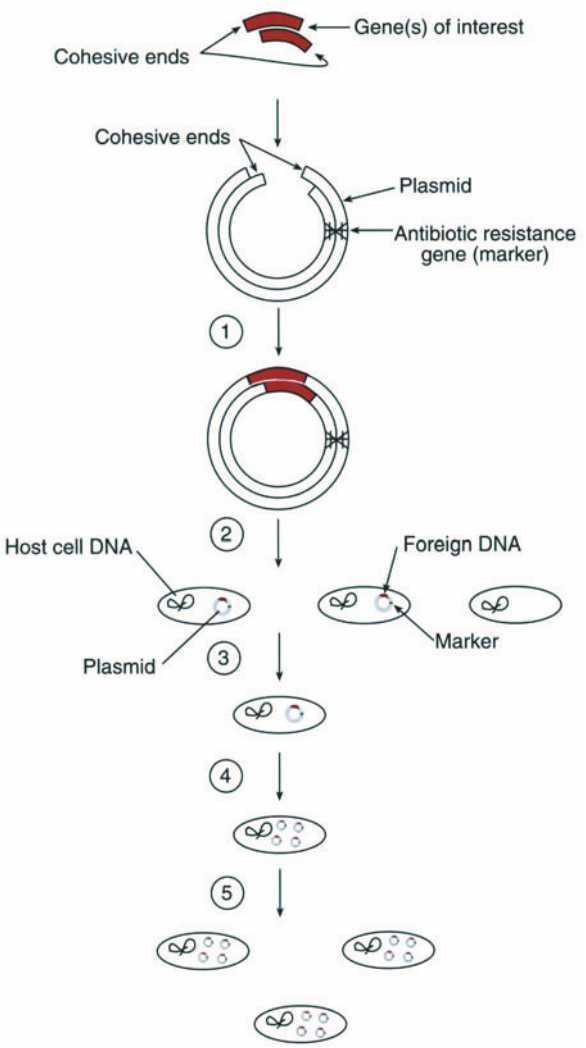


Figure C.13. Cloning of DNA in bacteria by means of plasmid vectors. (1) Insertion of passenger into vector—formation of recombinant DNA from target DNA fragment and plasmid; (2) insertion of recombinant DNA (altered plasmid) into host cells; (3) selection of host cells containing recombinant DNA; (4) intracellular replication of plasmids (cloning); (5) repeated cellular divisions (cloning).

without cloning (Figure C.14). The method requires a knowledge of the nucleotide sequences that flank the target segment in the DNA of interest. Oligonucleotides complementary to these flanking regions are produced synthetically and used as primers for a series of repetitive steps involving DNA polymerase.

To carry out the polymerase chain reaction, we heat-denature a sample of DNA containing the target segment to separate the two strands. Next, we add the synthetic primers and anneal the DNA strands to the primers. We then use DNA polymerase to extend the primers, with the original DNA serving as template. Action of DNA polymerase produces new copies of the target segment.

We follow this set of reactions by a second cycle of heat denaturation, annealing, and primer extension. Using a thermostable DNA polymerase ensures that the enzyme is not denatured during the heating step and need not be replenished for each cycle. We can obtain such thermostable enzymes from *thermophiles,* bacteria that grow at high temperatures and that occur in hot springs and geysers. We repeat the cycle of denaturation, annealing, and primer extension many times. At each cycle, the amount of the DNA of interest is roughly doubled, thereby providing exponential amplification of the target DNA.

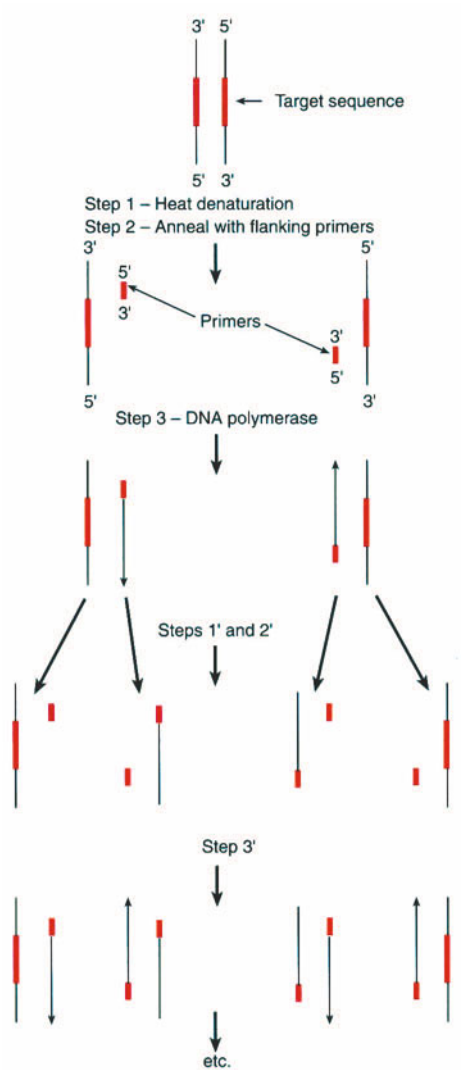


Figure C.14. Principle of the polymerase chain reaction.

Appendix D
Oxidation–Reduction Reactions

D.1. HALF-REACTIONS

Oxidation–reduction (or *redox*) reactions involve changes in the electronic structure of atoms and molecules. *Oxidation represents a loss of electrons* (e^-), and *reduction a gain of electrons*. Electron losses and gains result from direct transfer of electrons from one substance to another or from transfer of electrons in association with an H, an O, or some other atom. Because hydrogen generally acts as an electropositive element and oxygen as an electronegative element, gain of hydrogen and loss of oxygen are equivalent to reduction, and loss of hydrogen and gain of oxygen are equivalent to oxidation. Most biochemical oxidation–reduction reactions involve gain or loss of hydride ions (H^- or H:) and are catalyzed by *dehydrogenases.* Relatively few reactions—though some very important ones—involve gain or loss of oxygen and are catalyzed by *oxidases.*

Table D.1 gives some examples of oxidation–reduction reactions. Each reaction is written as a *reduction* and represents a *half-reaction,* one-half of a complete chemical reaction. An overall reaction consists of *two* half-reactions, one proceeding as a reduction, and the other as an oxidation. *There can never be oxidation without reduction, and vice versa.* This holds not only for the overall reaction but for each half-reaction as well. Even the simplest half-reaction involves an oxidized and a reduced species (for example, Ag^+ and Ag). The two form a **conjugate redox pair,** much as A^- and HA form a conjugate acid–base pair.

Two half-reactions must be combined in a way that preserves electrical neutrality. *There can be no net gain or loss of electrons* (Figure D.1).

D.2. DIRECTION OF REDOX REACTIONS

In determining the direction of an overall reaction, keep in mind the sign convention for reduction potentials. According to this convention, a half-reaction having a highly negative reduction potential involves a strong reducing agent; one having a highly positive reduction potential involves a strong oxidizing agent:

Reduction potential (volts)		
(−)	0	(+)
Strong reducing agent (weak oxidizing agent)		Weak reducing agent (strong oxidizing agent)

Of two half-reactions, the one having the *smaller* reduction potential involves a stronger reducing agent than the other half-reaction. Consequently, the half-reaction having the smaller reduction potential proceeds as an oxidation (loss of electrons) and must be written in the reverse sense (Figure D.2). The electrons generated by this half-reaction cause the second half-reaction to proceed as a reduction (gain of electrons). These relationships hold regardless of the signs of the potentials. Thus, of two half-reactions having reduction potentials of −0.52 V and −0.27 V, respectively, the former will proceed as an oxidation, the latter as a reduction. Likewise, of two half-reactions having reduction potentials of +0.18 V and +0.46 V, respectively, the former will proceed as an oxidation, the latter as a reduction. With this in mind, let us proceed to couple two half-reactions, once under standard and once under actual conditions.

Table D.1. Examples of Redox Reactions

Type of reaction	Examples
Gain/loss of electrons	$Ag^+ + e^- \rightleftarrows Ag$ Cytochrome b (Fe^{3+}) + $e^- \rightleftarrows$ cytochrome b (Fe^{2+})
Loss/gain of oxygen	$MnO_4^- + 2H_2O + 3e^- \rightleftarrows MnO_2 + 4OH^-$ $NO_3^- + 2H^+ + 2e^- \rightleftarrows NO_2^- + H_2O$
Gain/loss of hydrogen	$2CO_2 + 2H^+ + 2e^- \rightleftarrows H_2C_2O_4$ $NAD^+ + 2H^+ + 2e^- \rightleftarrows NADH + H^+$

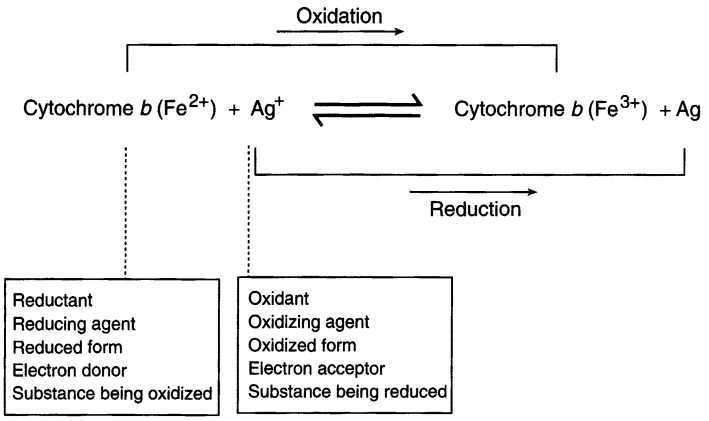


Figure D.1. Example of an overall redox reaction resulting from the combination of two half-reactions.

D.2.1. Standard Conditions

Suppose that we wanted to couple the following two half-reactions under biochemical standard conditions (concentrations are 1.0M each). We start out by *writing both reactions as reductions,* since the corresponding $E^{\circ\prime}$ values are those of reduction potentials:

$$\text{Acetaldehyde} + 2H^+ + 2e^- \rightleftarrows \text{ethanol} \qquad E^{\circ\prime} = -0.20\ \text{V}$$

$$Fe^{3+} + e^- \rightleftarrows Fe^{2+} \qquad E^{\circ\prime} = +0.77\ \text{V}$$

Comparing the $E^{\circ\prime}$ values, we see that the acetaldehyde half-reaction has a reduction potential that is *less* than that of the iron half-reaction. We conclude that ethanol constitutes a stronger reducing agent than Fe^{2+}. Hence, Fe^{3+} will be reduced to Fe^{2+} by ethanol, which undergoes oxidation to acetaldehyde. Therefore, we write the acetaldehyde reaction in the reverse sense, as an oxidation. Since we reversed the reaction we must also reverse the sign of the corresponding $E^{\circ\prime}$, just as $\Delta G^{\circ\prime}$ for a forward reaction becomes $-\Delta G^{\circ\prime}$ for the reverse reaction. Thus,

$$\text{Ethanol} \rightleftarrows \text{acetaldehyde} + 2H^+ + 2e^- \qquad E^{\circ\prime} = +0.20\ \text{V}$$

$$Fe^{3+} + e^- \rightleftarrows Fe^{2+} \qquad E^{\circ\prime} = +0.77\ \text{V}$$

Inspection shows that the overall reaction resulting from a combination of these two equations would not be balanced with respect to electrons. Because the overall reaction has to be electrically neutral, we must multiply the iron reaction by 2. Multiplication does not change the potential because $E^{\circ\prime}$ measures the *tendency* of Fe^{3+} to undergo reduction to Fe^{2+}. This tendency is the same whether one, two, or a thousand atoms undergo reaction.

We now *add* the two equations and *add* the $E^{\circ\prime}$ values likewise:

$$\text{Ethanol} \rightleftarrows \text{acetaldehyde} + 2H^+ + 2e^- \qquad E^{\circ\prime} = +0.20\ \text{V}$$

$$2Fe^{3+} + 2e^- \rightleftarrows 2Fe^{2+} \qquad E^{\circ\prime} = +0.77\ \text{V}$$

The overall reaction is

$$\text{Ethanol} + 2Fe^{3+} \rightleftarrows \text{acetaldehyde} + 2H^+ + 2Fe^{2+}$$

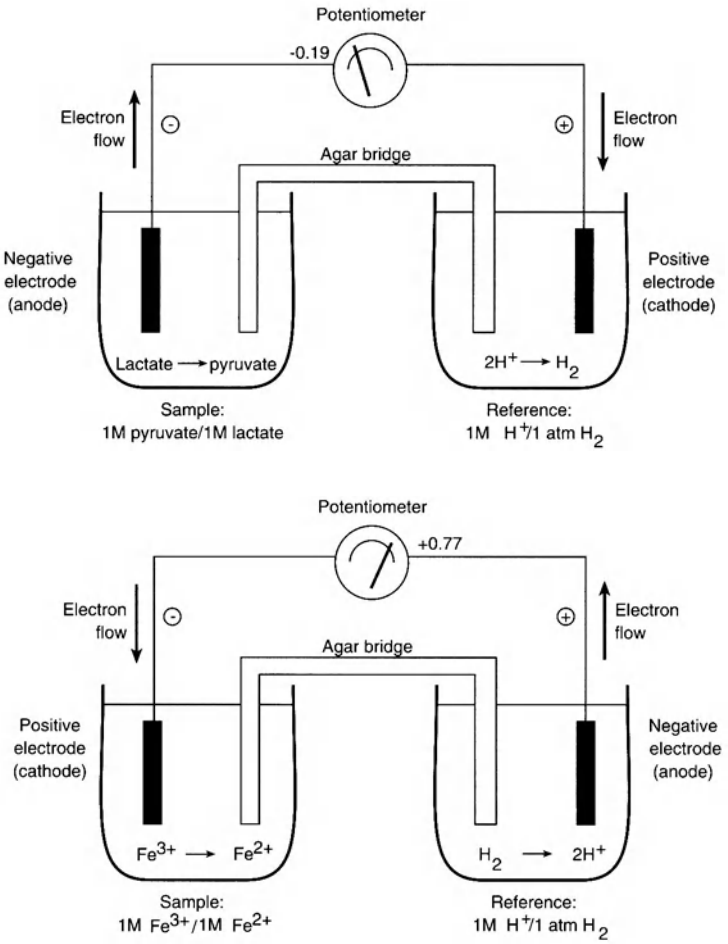


Figure D.2. Electrochemical cells for determining biochemical standard reduction potentials. Sample half-reactions are measured at 25°C and pH 7.0 against the standard hydrogen electrode.

for which $\Delta E^{\circ\prime} = (+0.20) + (+0.77) = +0.97$ V. Because $\Delta E^{\circ\prime}$ for the overall reaction is positive, $\Delta G^{\circ\prime}$ for the same reaction is negative (Eq. 12.3):

$$\Delta G^{\circ\prime} = -(2)(96{,}491)(+0.97) = -187{,}193 \text{ J mol}^{-1}$$
$$= -187 \text{ kJ mol}^{-1}$$

We conclude that we have coupled the two half-reactions properly *under biochemical standard conditions.* The negative $\Delta G^{\circ\prime}$ indicates that the overall reaction proceeds spontaneously.

D.2.2. Actual Conditions

To determine the direction of a redox reaction under nonstandard conditions, we have to use biochemical actual, rather than standard, reduction potentials. As an example, assume we wish to couple the two reactions that we considered above when the initial concentrations of oxidized and reduced forms have the following arbitrary values:

$$[\text{acetaldehyde}] = 5.0M \qquad [\text{ethanol}] = 1.0 \times 10^{-16}M$$
$$[Fe^{3+}] = 1.0 \times 10^{-16}M \qquad [Fe^{2+}] = 5.0M$$

In order to decide which species, *under these conditions,* is the stronger reducing agent, we must calculate the biochemical actual reduction potentials by means of the Nernst equation:

$$E' = E^{\circ\prime} - \frac{0.06}{2} \log \frac{[\text{reductant}]}{[\text{oxidant}]}$$

Using this equation, we have for the iron half-reaction

$$E' = +0.77 - \frac{0.06}{2} \log \frac{5.0}{1.0 \times 10^{-16}} = +0.27 \text{ V}$$

and for the acetaldehyde half-reaction

$$E' = -0.20 - \frac{0.06}{2} \log \frac{1.00 \times 10^{-16}}{5.0} = +0.30\,\text{V}$$

Note these important guidelines:

1. You must calculate the actual potentials in order to decide which species is the stronger reducing agent under nonstandard conditions. You cannot arrive at this decision by inspecting the values of the standard reduction potentials.
2. You must use tabulated values of $E^{\circ\prime}$ to calculate actual potentials, regardless of the fact that, ultimately, you will reverse one of these half-reactions and change the sign of its potential. You will have to change the sign of the actual potential; do not change the sign of $E^{\circ\prime}$.
3. You must balance the two half-reactions with respect to electrons. In the present example, you use $n = 2$ for both reactions even though the iron half-reaction involves only one electron. Proper coupling requires that the iron half-reaction be multiplied by 2.

As it turns out, under the conditions chosen, the reduction potential of the iron half-reaction is *less* than that of the acetaldehyde half-reaction. Hence, *under these conditions,* Fe^{2+} is a stronger reducing agent than ethanol. Consequently, acetaldehyde will be reduced to ethanol by Fe^{2+}, which undergoes oxidation to Fe^{3+}.

You now proceed as before. Reverse the iron half-reaction and change the sign of its potential, the actual reduction potential. Multiply the iron half-reaction by 2 in order to balance the electrons and add the two reactions; add the potentials likewise:

$$2Fe^{2+} \rightleftarrows 2Fe^{3+} + 2e^- \qquad E' = -0.27\,\text{V}$$

$$\text{Acetaldehyde} + 2H^+ + 2e^- \rightleftarrows \text{ethanol} \qquad E' = +0.30\,\text{V}$$

The overall reaction is

$$\text{Acetaldehyde} + 2Fe^{2+} + 2H^+ \rightleftarrows \text{ethanol} + 2Fe^{3+}$$

for which $\Delta E' = (-0.27) + (+0.30) = +0.03$ V, and, based on Eq. (12.4),

$$\Delta G' = -(2)(96{,}491)(+0.03) = -5789\ \text{J mol}^{-1} = -5.8\ \text{kJ mol}^{-1}$$

You have coupled the two reactions properly, this time under nonstandard conditions. Because $\Delta E'$ is positive, $\Delta G'$ is negative, and the reaction proceeds spontaneously as written. Note that the overall reaction has been reversed from that occurring under standard conditions simply as a result of changes in the concentrations of reactants and products. As pointed out in Section 9.1, *what ultimately determines the direction in which a reaction proceeds intracellularly are the concentrations of reactants and products.*

Answers to Problems

CHAPTER 1

1.1. NaCl dissociates into Na^+ and Cl^- ions, and soap micelles carry a negative surface charge. Because these particles are charged, water dipoles cluster around them. The water shells prevent the solutes from interacting with each other.

1.2. Yes, inverted micelles can form. Their inner core consists of polar carboxyl groups and is surrounded by a nonpolar shell composed of hydrocarbon chains.

1.3. (a) 1/2; (b) 1/4; (c) 1/16; (d) 1/64.

1.4. −21.8 kJ.

1.5. NaCl can disrupt electrostatic bonds, urea can disrupt hydrogen bonds, and SDS can disrupt hydrophobic interactions. Disrupting interactions results from the added compound competing with and substituting for the original solute. Urea breaks hydrogen bonds in a protein by forming hydrogen bonds with donor and acceptor groups in the protein; SDS disrupts hydrophobic interactions in a biomembrane by binding to membrane lipids; and ions bind to polar groups of an affected solute.

1.6. Charge separation occurs along each CO axis ($\overset{\delta^+}{C}{=}\overset{\delta^-}{O}$), but the molecule as a whole does not constitute a dipole; the centers of the positive and negative charges coincide.

1.7. Polar: hydroxymethyl, sulfonic acid, amide. Nonpolar: sulfide, phenyl, methyl. Capable of hydrogen bonding: hydroxymethyl, sulfonic acid, amide.

1.8. Water: 1.2×10^{-12} g; protein: 1.3×10^{-13} g; DNA: 5.3×10^{-15} g; RNA: 1.0×10^{-12} g; polysaccharide: 2.0×10^{-14} g.

1.9. Because both compounds can form hydrogen bonds with water.

1.10. Obtain 28.7 ml of concentrated acetic acid, add 178.0 ml of 1.00*M* NaOH, and dilute the solution to 2.00 liter with water.

1.11. 63.3.

1.12. pH 6.8.

1.13. The phosphate buffer, because a buffer is most effective within ±1.0 pH unit from its pK'_a.

1.14. Weigh out 2.11 g of KH_2PO_4 and 1.66 g of K_2HPO_4, and dissolve the salts in 250 ml of water.

1.15. (a) 0.157; (b) 9.55 ml.

1.16. pH 4.76.

1.17. 10^{-4}.

1.18. Because of the values calculated from the ion product of water ($K_w = 1.0 \times 10^{-14}$). When $[H^+] = 1.0 \times 10^{-14}$, pH = 14; when $[OH^-] = 1.0 \times 10^{-14}$, $[H^+] = 1.0$, and pH = 0. The pH scale can extend below 0 and above 14. For example, when $[H^+]$ = 2.00*M*, pH = −0.30; when $[OH^-]$ = 2.00*M*, $[H^+] = (1.0 \times 10^{-14})/2.00$, and pH = 14.3.

1.19. The first proton dissociating from each phosphate group has a pK'_a approximately equal to pK'_{a_1} of H_3PO_4 because in both instances dissociation occurs from an uncharged entity. In H_3PO_4, the proton dissociates from an uncharged molecule; in ATP, it dissociates from an uncharged group.

1.20. $pK'_a = 6.60$.

1.21. 0.200*M*.

1.22. Pyruvic acid, formic acid, and succinic acid (2).

1.23. According to the Henderson–Hasselbalch equation, a buffer's pH is determined by its pK'_a value and the ratio $[A^-]/[HA]$. For phosphate buffers, pK'_a is fixed, but we can vary the concentrations of A^- and HA to produce different buffers. As long as the ratio

$[A^-]/[HA]$ remains constant, the buffer's pH will remain unchanged even though its molarity will vary.

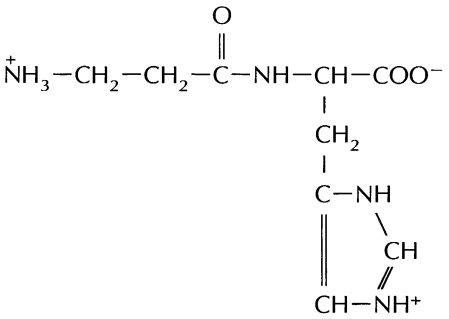
1.24. 10^{-5}–10^{-8} *M*.

1.25. 76%.

CHAPTER 2

2.1. (a) Both amino acids still classified as polar with uncharged R groups; (b) both amino acids classified as having negatively charged R groups.

2.2.



2.3. The equilibrium constants are:

$$K'_{a_1} = [HA^{\pm}][H^+]/[H_2A^+]$$

$$K'_{a_2} = [A^-][H^+]/[HA^{\pm}]$$

At the isoelectric point the amino acid has an equal number of positive and negative charges. Mathematically, we can express this condition by equating the concentrations of the cationic (H_2A^+) and anionic (A^-) forms of the amino acid. From the first equation,

$$[H_2A^+] = [HA^{\pm}][H^+]/K'_{a_1}$$

and from the second equation

$$[A^-] = K'_{a_2}[HA^{\pm}]/[H^+]$$

Equating the two terms,

$$[HA^{\pm}][H^+]/K'_{a_1} = K'_{a_2}[HA^{\pm}]/[H^+]$$

$$[H^+] = (K'_{a_1}K'_{a_2})^{1/2}$$

$$pH = -\log[H^+] = \tfrac{1}{2}(pK'_{a_1} + pK'_{a_2})$$

2.4. pI = 6.68.

2.5. There is no pH at which the *absolute* charge of an ordinary protein is zero; at best, the molecule can have a *net* zero charge.

2.6. The order of elution is lysine (first), glycine, aspartic acid (last).

2.7. Any amide bond contains the grouping $-\overset{\overset{O}{\|}}{C}-NH-$, but a peptide bond exists only when the CO and NH groups derive from α-carboxyl and α-amino groups of two amino acids.

2.8. (a) MW = 24,560; (b) four disulfide bonds; (c) 6.35 mg.

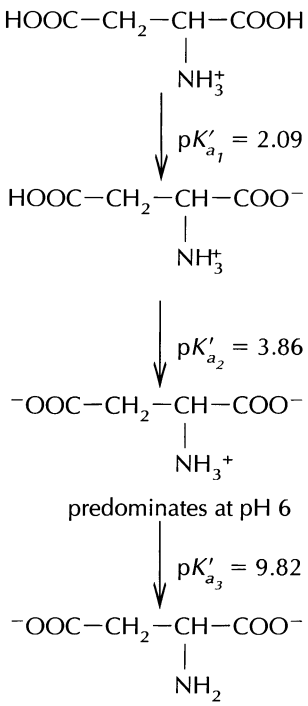
2.9. 27 peptides:

AAA
PPP
GGG

AAP, AAG	GGA, GGP	PPA, PPG
APA, GAA	GAG, GPG	PAP, PGP
PAA, AGA	AGG, PGG	APP, GPP

AGP
APG
GAP
GPA
PAG
PGA

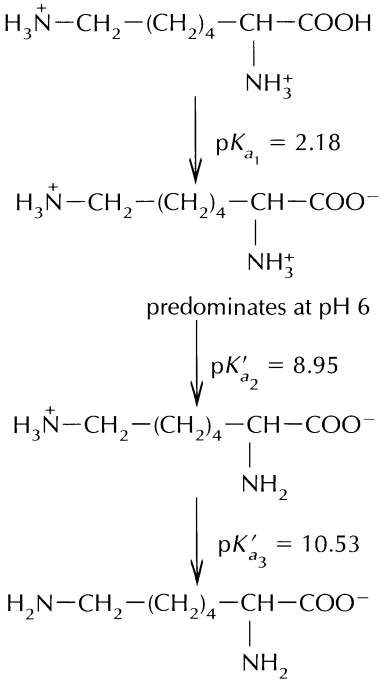
2.10. (a) Aspartic acid



predominates at pH 6

predominates at pH 12

(b) Lysine



2.11. The peptide is cyclic or has a blocked (for example, acetylated) N-terminus.

2.12. Glycine does not exist in either an L- or a D-configuration because it does not possess a chiral center.

2.13. (a) +1; (b) −3; pI ≈ 9.

2.14. $[^{+}H_3N{-}CH_2{-}COO^{-}]/[H_2N{-}CH_2{-}COOH] = 1.82 \times 10^{7}$.

2.15. (a) pH 12.0; (b) pH 13.0.

2.16. MW = 146.

2.17. A: Asp-Phe; B: Gly-Cys; C: Tyr-Lys.

2.18. Within ±1 pH unit of the pK'_a values, that is, in the range of pH 1.34–3.34 and pH 8.69–10.69. The buffers have maximum capacity at the pK'_a values (pH 2.34 and 9.69).

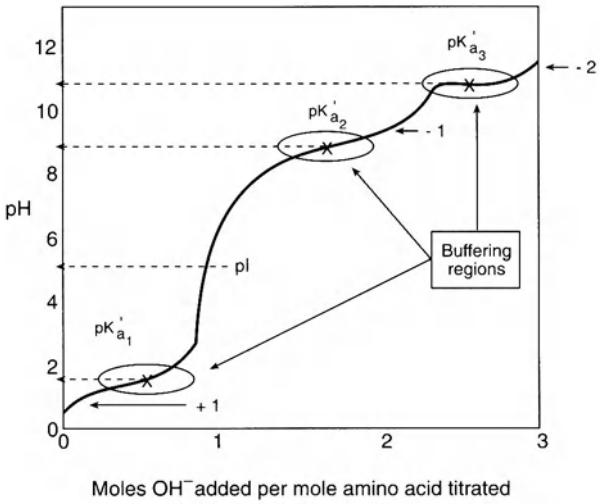
2.19. pH 7.

2.20. Aliphatic: a, e, f; aromatic: b, c, d; acidic: a; basic: e; polar: a, c, d, e; nonpolar: b, f.

2.21. MW = 50,000.

2.22. 6.0 cm.

2.23.

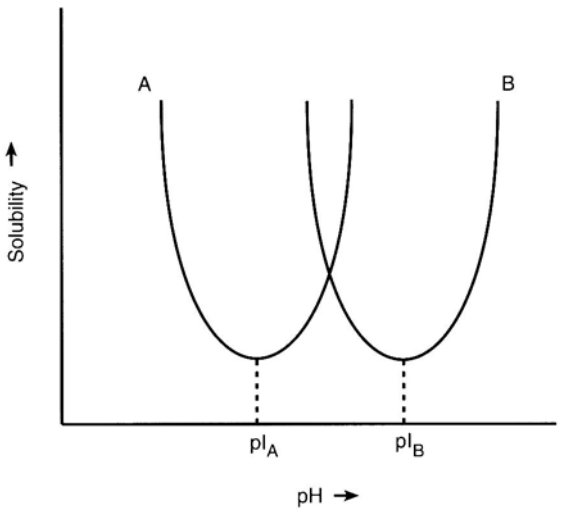


(a) The pK'_a values are approximately 1.5, 8.5, and 11. The amino acid appears to be cysteine. (b), (c) Marked on the figure.

2.24. Plot c.

CHAPTER 3

3.1.



Protein A is more acidic and has a lower pI. To separate the proteins, lower the pH to pI_A, at which point protein A precipitates. Centrifuge the solution. Protein A forms a pellet, and protein B remains in the supernatant.

3.2. MW = 27,500.

3.3. Lysozyme.

3.4. (a) 3; (b) 3; (c) 2, 3, 4.

3.5. The denatured protein has a more open structure and may have a larger number of functional groups accessible to titration. The pK'_a of a group may differ in (a) and (b) because the group may be located in different electronic environments.

3.6. MW = 20,000.

3.7. 24 H^+.

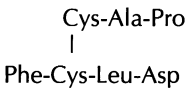
3.8. Gly-His-Pro-Arg.

3.9. MW = 50,000.

3.10. MW = 200,000.

3.11. Asp-Met-Asp-Met-His-Gly.

3.12.



3.13. Since two moles of mercaptoethanol are required to break one mole of disulfide bond, there must be three disulfide bonds per molecule of protein. The protein is an oligomer, composed of three identical subunits (MW = 20,000), joined either in series, with one and two disulfide bonds between two subunits, or in a closed triangular arrangement, with one disulfide bond at each side.

3.14. Ser-Arg-Tyr-Glu-Cys.

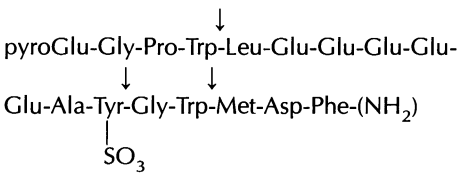
3.15. (a) At pH 7.0, the β-carboxyl group of aspartic acid is ionized and the imidazole group of histidine is uncharged. The same holds for pH 11.0 so that the interactions remain unchanged. (b) At pH 3.0, the β-carboxyl group of aspartic acid still carries a negative charge, but the imidazole group now carries a positive charge. Consequently, the attraction between the two groups increases. (c) Increasing the distance between the groups decreases their electrostatic interaction.

3.16. 4.88×10^8 years.

3.17. 4 subunits.

3.18. (a) 20; (b) 8.

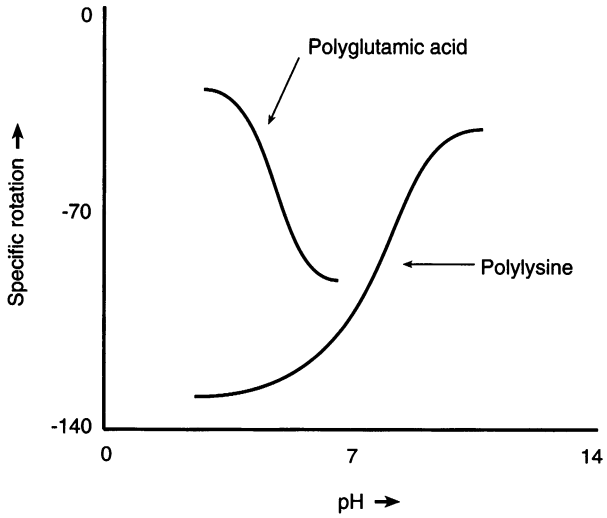
3.19. (a) No hydrolysis with trypsin. (b) Three cleavages with chymotrypsin:



3.20. Leu-Gly-Leu-Asp-His-Tyr-Gly-His-Phe.

3.21. 50 turns; 27 nm.

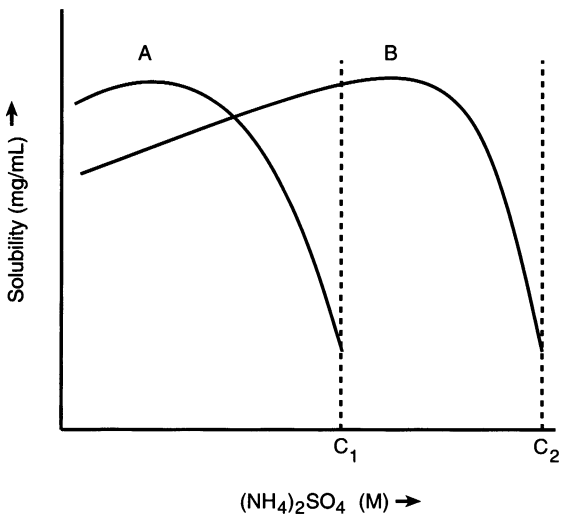
3.22.



3.23. 5 turns.

3.24. 4 forms:
α-α-β-β
α-β-α-β
α-β-β-α
β-α-α-β

3.25. Dialyze the protein solution against water, and then add enough solid $(NH_4)_2SO_4$ (to concentration c_1) to precipitate protein A. Collect protein A by centrifugation. Protein B remains in the supernatant. Add $(NH_4)_2SO_4$ (to concentration c_2), and collect the precipitated protein B by centrifugation.



3.26. (a) 5.33×10^9 amino acids; (b) 164 peptide bonds per second.

3.27. Leu-His-Arg-Ala-Met-Glu-Ser-Lys-Asp-Met-Lys-Phe-Gly-Met-Lys.

CHAPTER 4

4.1. Ordinary enzyme: 15-fold; allosteric enzyme: 2-fold.

4.2. pH 5.13.

4.3. (a) 80%; (b) 2-fold.

4.4. $Q_{10} = 1.77$.

4.5. 54.6.

4.6. (a) $[E] = [E_f] + [ES] + [EI]$
(b) $[E] = [E_f] + [ES] + [EI] + [ESI]$
(c) $[E] = [E_f] + [ES] + [ESI]$

4.7. Enzyme B.

4.8. $K_m/[S] = 0.5$.

4.9. Ala-Gly is the better substrate because it leads to a smaller K_m. The inhibitor having the smaller K_i is the stronger inhibitor.

4.10. $d[ES]/dt = k_1[E_f][S] + k_{-2}[E_f][P] - (k_{-1} + k_2)[ES] = 0$.

4.11. (a) 45°C; (b) 60°C; (c) 70°C.

4.12. According to the concerted model, increased concentrations of competitive inhibitors cause a shift of the equilibrium between the R and T enzyme forms identical to that produced by increased substrate concentrations. In both instances, large numbers of substrate binding sites are generated. Provided that the inhibitor concentration is low, some of the additional binding sites can bind substrate, thereby accelerating the reaction.

4.13. $[S] = 1.0 \times 10^{-2} M$.

4.14. $K_m = 5.0 \times 10^{-3} M$.

4.15. The second reaction has a greater energy of activation.

4.16. (a) Concentrate the solution surrounding the dialysis bag by lyophilization (freeze-drying), and add some of it back to the enzyme solution. Restoration of enzymatic activity indicates loss of an essential cofactor. (b) Dialyze the inactive enzyme solution against a solution containing high concentrations of suspected cofactors. Increasing enzymatic activity as dialysis proceeds indicates involvement of a dialyzable cofactor. (c) Determine the optical rotation (helical content) of the active and inactive enzymes. Identical optical activity indicates that the molecule has not unfolded.

4.17. MW = 25,000.

4.18. $K_i = 5.3 \times 10^{-5}$.

4.19. A is a precursor of B. C exerts feedback inhibition.

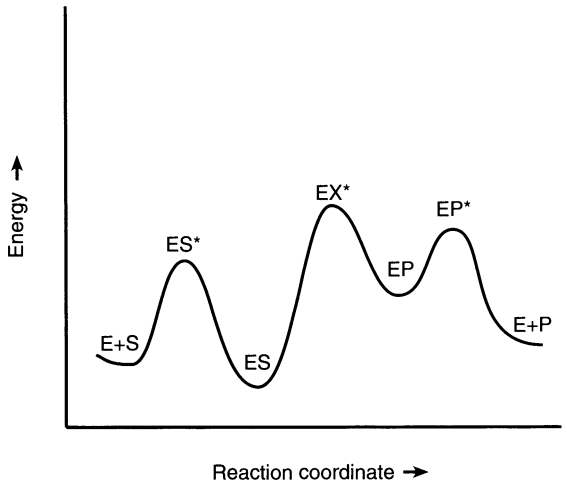
4.20. The line has a negative slope, $-(1/K_m)$, an intercept on the v/[S] ordinate of V_{max}/K_m, and an intercept on the v abscissa of V_{max}.

4.21. $[I] = 1.0 \times 10^{-1} M$.

4.22. The enzyme–substrate complex is more stable than the isolated enzyme and substrate; substrate binding stabilizes the enzyme.

4.23. Ethanol is the normal substrate of alcohol dehydrogenase, and methanol acts as a competitive inhibitor in this reaction. Since competitive inhibition can be overcome by greatly increasing the substrate concentration, the methanol inhibition is effectively eliminated by administering large doses of ethanol.

4.24. ES and EP are stable complexes, and X is an intermediate in the conversion of S to P. (*) Designates an activated complex.



4.25. 75%.

4.26. 3.3 enzyme units.

4.27. $V_{max} = 50$ nmol ml^{-1} min^{-1}; $K_m = 6.0 \times 10^{-2} M$.

4.28.

1/[S] (mM)	1/v (normal) (mmol^{-1} ml min)	1/v (inhibited) (mmol^{-1} ml min)
0.667	5.99	8.70
0.500	4.90	6.99
0.400	4.31	5.99
0.200	3.20	4.00
0.100	2.60	3.00

A plot of 1/v versus 1/[S] shows that the inhibitor is a competitive inhibitor.

	Uninhibited	Inhibited
K_m	2.94 mM	5.00 mM
V_{max}	2.00 mmol ml^{-1} min^{-1}	2.00 mmol ml^{-1} min^{-1}

4.29. IAA combines with the SH group of cysteine; DFP reacts with the OH group of serine. Inhibition with either IAA or DFP indicates that both cysteine and serine are present at the enzyme's active site. The two

amino acids must be located near each other since once an inhibitor binds to either one, it blocks the second inhibitor from binding to the other amino acid.

4.30. The rate increases exponentially rather than linearly because the preparation contains trypsinogen that is activated to trypsin in an autocatalytic manner.

4.31. Muscle tissue labeling first increases as labeled amino acids become incorporated into newly made proteins. After reaching a maximum level, the extent of tissue label begins to decrease as proteins degrade and their amino acids enter metabolism.

4.32. 80.0%; 1500-fold.

CHAPTER 5

5.1. Two molecules of 2,3-dimethylglucose, four molecules of 2,3,6-trimethylglucose, one molecule of 1,2,3,6-tetramethylglucose, and three molecules of 2,3,4,6-tetramethylglucose.

5.2. Initial: +37.4°; final +52.7°.

5.3. (a) Three: reducing end, nonreducing end, and internal residues linked via two α(1 → 4) glycosidic bonds; (b) five: reducing ends, nonreducing ends, and internal residues linked via two α(1 → 4) glycosidic bonds or two α(1 → 4) and one α(1 → 6) glycosidic bonds or one α(1 → 4) and one α(1 → 6) glycosidic bonds.

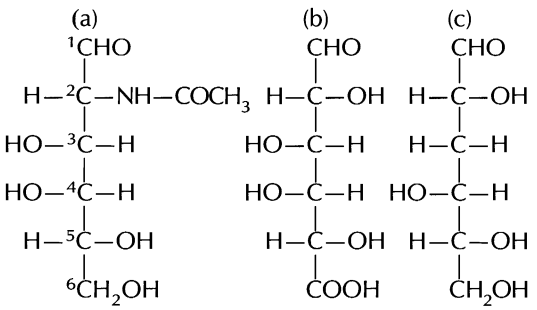
5.4. 21.6 g.

5.5. −11°.

5.6. Because the outer membrane acts as a barrier for lysozyme, thereby preventing the enzyme from digesting the cell wall.

5.7. 37% (α); 63% (β).

5.8.



5.9. The repeating unit in starch (maltose) consists of two glucose residues linked α(1 → 4); the repeating unit in cellulose (cellobiose) consists of two glucose residues linked β(1 → 4).

5.10. (a) 4; (b) 3; (c) 5.

5.11. 19 different disaccharides (the two disaccharides marked with an asterisk are identical):

α-D-glucosyl-α-D-glucose	β-D-glucosyl-β-D-glucose
(1 → 1)	(1 → 1)
(1 → 2)	(1 → 2)
(1 → 3)	(1 → 3)
(1 → 4)	(1 → 4)
(1 → 6)	(1 → 6)
α-D-glucosyl-β-D-glucose	β-D-glucosyl-α-D-glucose
(1 → 1)*	(1 → 1)*
(1 → 2)	(1 → 2)
(1 → 3)	(1 → 3)
(1 → 4)	(1 → 4)
(1 → 6)	(1 → 6)

5.12. MW = 6480.

5.13. Every amylose chain contains only one reducing end. This represents a very small fraction of reducing glucose residues relative to the total number of glucose residues present. When amylose solutions are treated with Tollens' and Benedict's reagents, the extent of reaction is so small as to be undetectable.

5.14. 75,585.

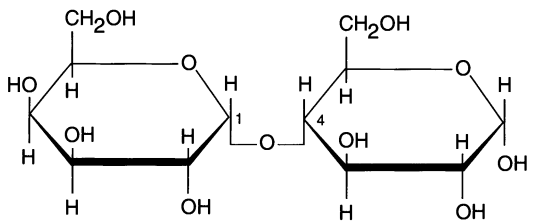
5.15. 10.

5.16. 10%.

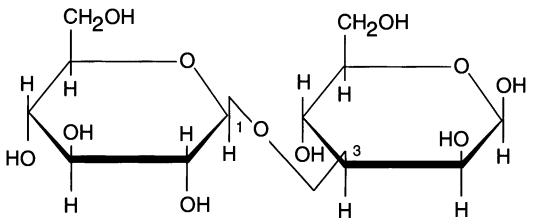
5.17. 1: C; 2: D,E; 3: B; 4: D; 5: A; 6: B.

5.18.

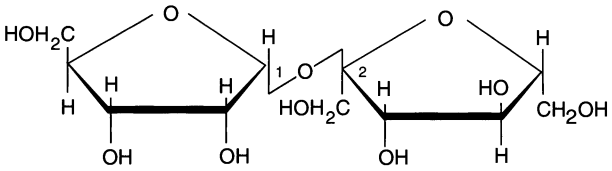
A



B



C



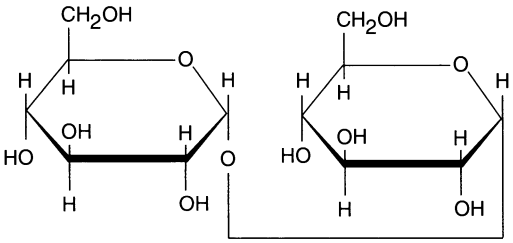
5.19. Growing cells actively synthesize cell walls in a process that includes cross-linking peptidoglycan strands. Penicillin inhibits the enzyme that catalyzes this cross-linking and prevents cell wall formation. Once the cross-links have formed, the cell wall is no longer susceptible to penicillin action.

5.20. (a) 4; (b) 2; (c) 5.

5.21. (a) Hydrolyze three consecutive glycosidic bonds at one end of the amylose chain to produce four fragments containing 1, 1, 1, and 97 residues. Average chain length: 25 residues. (b) Hydrolyze three glycosidic bonds to produce four fragments containing 25 residues each. Average chain length: 25 residues.

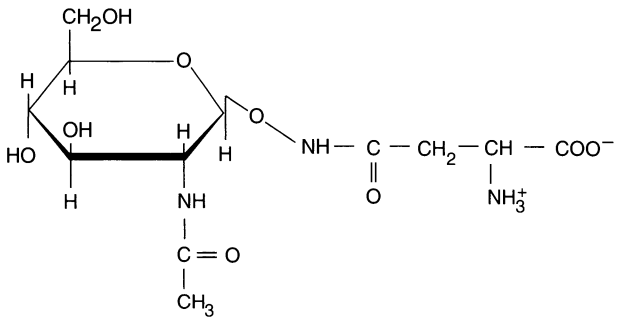
5.22. Reduction converts the optically active D-galactose to a *meso* compound.

5.23. α-D-Glucopyranosyl-(1 → 1)-α-D-glucopyranoside.

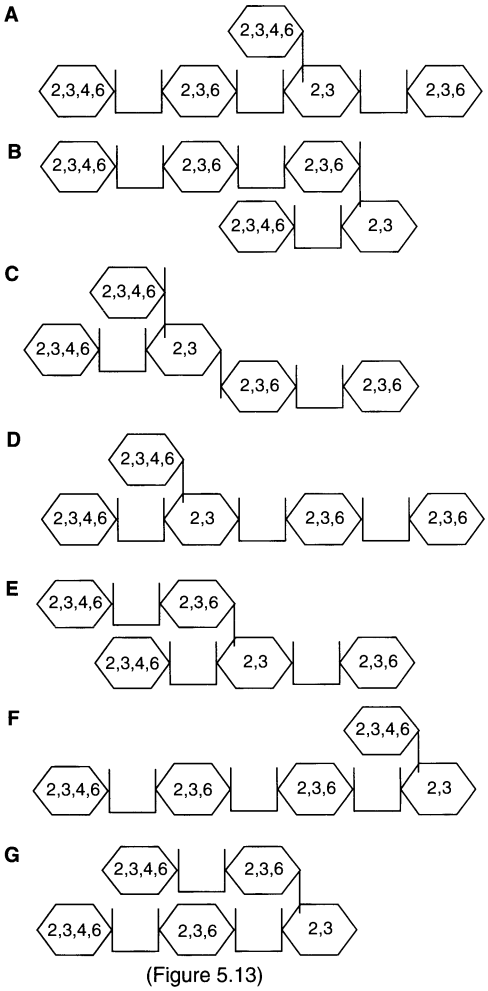


5.24. MW = 180,000.

5.25.



5.26. The following pentasaccharides all yield the same methylated products as those shown in Figure 5.13:



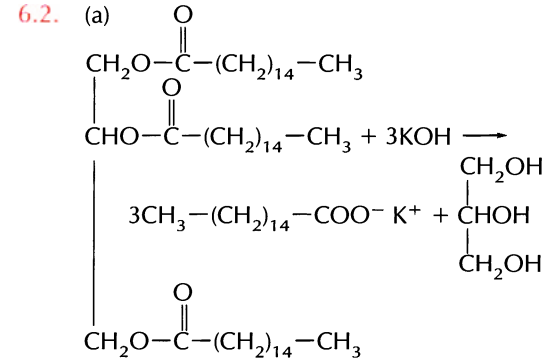
(Figure 5.13)

5.27. Because substituents linked via equatorial bonds generally have greater stability than those linked via axial bonds, the preponderance of equatorial bonds in cellulose contributes to the great stability of this substance.

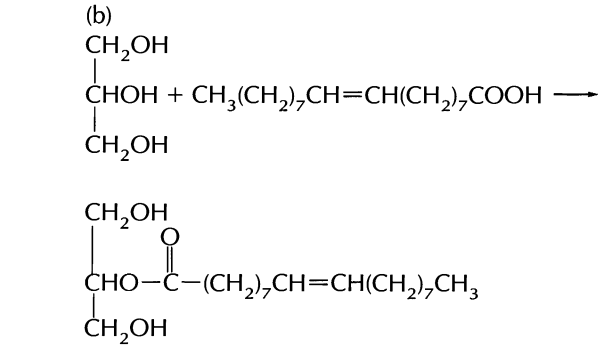
CHAPTER 6

6.1. 1.17 g/ml.

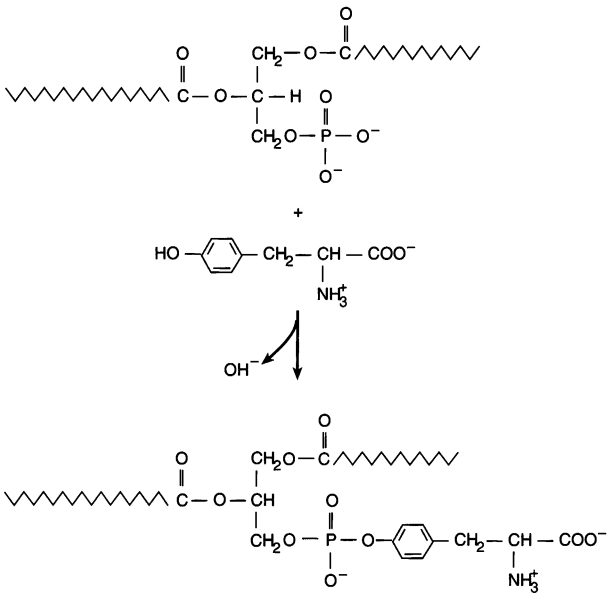
6.2. (a)



(b)



(c)



6.3. 27 triacylglycerols (18 optically active):

6 of FA_1, FA_2, and FA_3
3 of $2FA_1$, FA_2
3 of $2FA_1$, FA_3
3 of $2FA_2$, FA_1
3 of $2FA_2$, FA_3
3 of $2FA_3$, FA_1
3 of $2FA_3$, FA_2
1 of $3FA_1$
1 of $3FA_2$
1 of $3FA_3$

6.4. Lecithin, ceramide, cephalin, sphingomyelin, phosphatidic acid, and ganglioside GM_2.

6.5. 172 mg of KOH/g of fat.

6.6. 269 g of iodine/100 g of fat.

6.7. Cholesterol; phosphatidic acid; sphingosine.

6.8. An inverted micelle could form.

6.9. 3.3 ml.

6.10. At pH 12, phosphatidyl serine, phosphatidyl ethanolamine, and phosphatidyl choline have net charges of −2, −1, and 0, respectively, and can be fully separated. At that pH, phosphatidyl serine and phosphatidyl ethanolamine move toward the anode, and phosphatidyl choline remains at the origin.

6.11. A diacylglycerol.

6.12. Six double bonds.

6.13. Because of the presence of significant amounts of unsaturated fatty acids, Mb-II has a less regular structure and is more sensitive to disruption by an increase in temperature. SDS is the compound of choice to disrupt either membrane. SDS forms hydrophobic interactions with membrane lipids, thereby breaking similar interactions occurring within the lipid bilayer.

6.14. Polar amino acids, because they constitute a site on a polar membrane surface and function in the binding of polar ATP.

6.15. 0.058*M*.

6.16. The monolayer area produced by the lipids from one red blood cell is $[0.890 \times 10^{12}\ \mu m^2]/(4.74 \times 10^9) = 188\ \mu m^2$. Since the cell's surface area is only 100 μm^2, the cell's lipids cover that area 188/100 ≈ 2 times. In other words, they form a bilayer.

6.17. 2.8×10^6 molecules.

6.18. The emulsified margarine is a better source of unsaturated fatty acids because hydrogenation converts unsaturated to saturated fatty acids.

6.19. Fat-soluble vitamins accumulate in lipid tissues whereas water-soluble vitamins are readily excreted.

6.20. 21.

6.21. (a) 1; (b) 1; (c) 2.

6.22. As bile salts are amphipathic molecules, they are expected to form micelles, monolayers, and bilayers.

CHAPTER 7

7.1. The mutated DNA could hypothetically be repaired by:

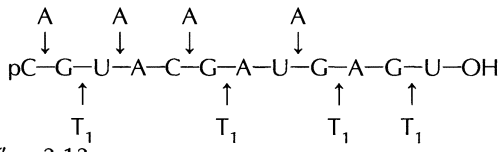
(a) Excising the mutated nucleotide, with or without adjacent nucleotides, by means of a nuclease. A nuclease can catalyze the making of two cuts in the damaged strand, one on either side of the mutated nucleotide.
(b) Filling the gap by means of DNA polymerase. DNA polymerase can catalyze synthesis of a DNA fragment, complementary and antiparallel to the intact strand.
(c) Linking the newly synthesized fragment to the remaining sections of the damaged strand by means of a third enzyme.

7.2. [A] = 30 mol %.

7.3. 2.2×10^{-3} micromoles.

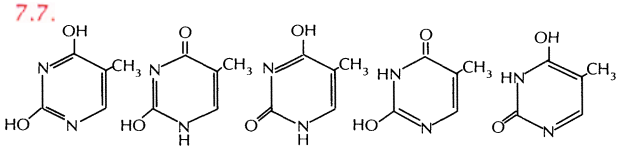
7.4. 3′-TACGGACTCATTG-5′.

7.5. (a) Cleavage at A; (b) cleavage at T_1; (c) random cleavage.



7.6. $pK'_a = 2.12$.

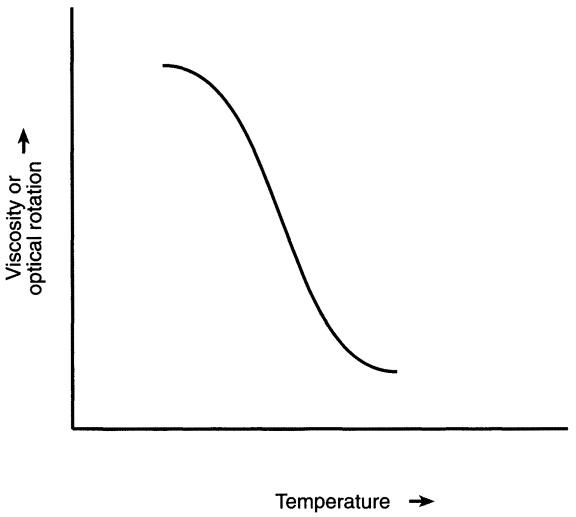
7.7.



7.8. Template: 3′ A–T–A–G–C–T–G–C–A–G–T–G–A 5′
Primer: 5′ T–A–T–C–G–A–C–G–T–C–A–C–T 3′
Composition in mol %: [A] = [T] = 26.9; [C] = [G] = 23.1.

7.9. Thermophiles might be expected to contain larger amounts of G and C than mesophiles. Because G and C are linked via three hydrogen bonds, while A and T are linked via only two, a DNA with a higher (G + C) content has a larger number of hydrogen bonds per unit length of double-stranded DNA. The more extensive the hydrogen bonding between the strands is, the greater the thermal stability of the DNA and the higher its melting temperature (T_m). Increased thermal stability of DNA may be one of the factors allowing thermophiles to grow at elevated temperatures.

7.10. Denaturation converts dsDNA from a long linear double helix to two globular random coils. The change is accompanied by decreases in both viscosity and optical rotation.



7.11. 12.5%.

7.12. A-DNA: 31.2 nm and 10.9 turns; B-DNA: 40.8 nm and 12.0 turns; Z-DNA: 44.4 nm and 10.0 turns.

7.13. Minimum: 200; maximum: 300.

7.14. No, some DNA is repetitive and some is not expressed.

7.15. The similarity of mitochondrial DNA to bacterial DNA—a circular dsDNA of comparable length—supports the endosymbiotic theory. According to this theory, mitochondria have evolved from primitive prokaryotes that took up a symbiotic relationship with early eukaryotic cells.

7.16. Since [G] does not equal [C], and [A] does not equal [T], the DNA cannot be double-stranded; it must be single-stranded.

7.17. [C] = 14.7 mol %; [T] = 31.8 mol %.

7.18. 550.

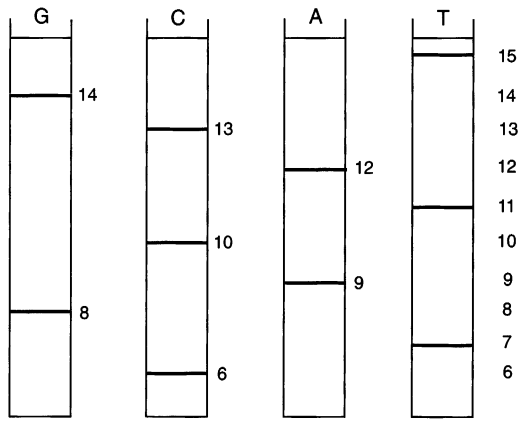
7.19. [A + T] = 73.9 mol %.

7.20. 5′ A—C—G—G—C—C—G—T 3′
3′ T—G—C—C—G—G—C—A 5′

7.21. (a) −4; (b) −5; (c) −4; (d) 0.

7.22. Mass of a base pair: 1.11×10^{-21} g; mass of the genome: 6.45×10^{-12} g or 6.45 pg (picograms).

7.23.



7.24. 2%.

7.25. 5′ A—C—A—G—G—T—C—C—A—T 3′
1 2 3 4 5 6 7 8 9 10

CHAPTER 8

8.1. Because control of the committed step determines whether or not the remaining steps of the pathway proceed.

8.2. In 150 ml: 1.35×10^{22} H^+; in 1000 ml: 9.03×10^{22} H^+.

8.3. A, Matrix; B, outer membrane; C, inner membrane.

8.4. A is contaminated by 5% of B and 10% of C; B is contaminated by 7% of A and 11% of C; C is contaminated by 8% of A and 13% of B.

8.5. 1.08 mg/h.

8.6. X → K → G → T.

8.7. Because newly synthesized polypeptide chains rapidly dissociate from ribosomes, exposure to a radioactive label must be brief so that a significant fraction of the label will tag polypeptide chains still attached to ribosomes. In organisms grown for prolonged times in the presence of a radioactive label, all newly synthesized proteins will become labeled. After the proteins dissociate from the ribosomes, their label will spread to other areas of protein metabolism, resulting in label distribution throughout the cell.

8.8. Y → S → Q → X.

8.9. 3000 ml.

8.10. The rat receiving the ^{35}S-labeled compound.

8.11. 13.6 years.

8.12. Damage to the pituitary results in a drastic decrease of vasopressin secretion. Because of the deficiency of this hormone, water is not reabsorbed by the kidneys and is excreted in large amounts as dilute urine. The resulting extensive dehydration leads to an unquenchable thirst.

8.13. 1100 kcal for men and 900 kcal for women.

CHAPTER 9

9.1. Because there can only be one set of actual conditions, regardless of the reference state used; concentrations, pH, and temperature have fixed values for any given set of conditions.

9.2. (a) The free energy change for a complex reaction sequence occurring *in vivo* must be identical to the free energy change of the simpler *in vitro* reaction, provided that the initial reactants and the final products are the same in both cases. (b) The free energy change of an enzyme-catalyzed reaction is identical to the free energy change of an uncatalyzed reaction having the same reactants and products.

9.3. Inhibitor B. The stronger the inhibitor, the greater is the equilibrium constant, $K'_{eq} = [EI]/([E][I])$, that describes the binding of inhibitor to enzyme. Since $\Delta G^\circ = -RT \ln K'_{eq}$, inhibitor B has the largest K'_{eq}.

9.4. $\Delta G' = -47.6$ kJ mol^{-1}.

9.5. $\Delta G^\circ = 13.7$ kJ mol^{-1}.

9.6.

	$\Delta G^{\circ\prime}$ (J mol^{-1})
(2) 2ADP ⇄ ATP + AMP	+2,030
(3) Glucose + ATP ⇌ glucose 6-phosphate + ADP + H^+	−19,100
(1) Glucose 6-phosphate ⇌ glucose 1-phosphate	+5,706
Glucose + ADP → glucose 1-phosphate + AMP + H^+ (overall reaction)	−11,364

Overall $\Delta G^{\circ\prime} = -11.4$ kJ mol^{-1}; overall $K'_{bio} = 98.1$.

9.7. $K'_{eq,overall} = 1.0 \times 10^2$.

9.8. a and c.

9.9. $K'_{eq} = 2.59 \times 10^{-3}$.

9.10. ATP^{4-} + glucose → AMP-glucose$^-$ + PP_i^{3-}
AMP-glucose$^-$ + P_i^{2-} → glucose 6-phosphate^{2-} + AMP^{2-} + H^+

Glucose + ATP^{4-} + P_i^{2-} → glucose 6-phosphate^{2-} + AMP^{2-} + PP_i^{3-} + H^+
$\Delta G^{\circ\prime}_{overall} = -43.2$ kJ mol^{-1}.

9.11. $\Delta G' = -14.7$ kJ mol^{-1}.

9.12. $K'_{bio} = 3.30 \times 10^{-8}$.

9.13. [Creatine]/[phosphocreatine] $= 3.03 \times 10^7$.

9.14. Mechanism *b*. ATP synthesis via this mechanism requires the more or less simultaneous collision of four entities at each step: ADP, P_i, H^+, and A, B, or C. By contrast, ATP synthesis via mechanism *a* requires the more or less simultaneous collision of 10 entities: 3ADP, $3P_i$, $3H^+$, and A. Such multimolecular collisions are chemically highly unlikely.

9.15. No, ΔG cannot be equal to ΔH. Since $\Delta S > 0$, the term $T\Delta S$ has a finite value and is not equal to zero. Accordingly, $\Delta H > \Delta G$.

9.16. (a) $K'_{eq,overall} = ([B][X][Y])/[A]$; $K'_{eq_1} = ([B][C])/[A]$; $K'_{eq_2} = ([X][Y])/[C]$. Thus, $K'_{eq_1}K'_{eq_2} = ([B][X][Y])/[A] = K'_{eq,overall}$.
(b) $[C] = (K'_{eq_1}[A])/[B] = ([X][Y])/K'_{eq_2}$. Thus, $K'_{eq_1}K'_{eq_2} = ([B][X][Y])/[A] = K'_{eq,overall}$.

9.17. pH 0.76.

9.18. Yes, the organism could grow. Since $\Delta G'$ for the conversion of citrate to isocitrate is negative, citrate can serve as a nutrient for the organism.

9.19. 5.08 moles of ATP.

9.20. (a) pH = 2.00; (b) $K'_a = 1.11 \times 10^{-3}$; (c) $pK'_a = 2.96$; (d) $\Delta G^\circ = 16.9$ kJ mol^{-1}.

9.21. [Isocitrate]/[citrate] $= 6.70 \times 10^{-2}$.

9.22. 49.9 kg of ATP; 73.4% of body weight.

9.23. 5.70 kJ mol^{-1}.

9.24. $K'_{bio} = 1.99$.

9.25. $\Delta G' = -73.3$ kJ mol^{-1}.

9.26. (a) -2884 kJ mol^{-1}; (b) $+12.1$ kJ mol^{-1}; (c) $+2.5$ kJ mol^{-1}; (d) $+13.8$ kJ mol^{-1}.

CHAPTER 10

10.1. Strenuous exercise exacerbates the effect of alcohol on gluconeogenesis. Exercise lowers the levels of cellular ATP and NADH and raises that of NAD^+. Increased concentrations of NAD^+ stimulate the alcohol dehydrogenase reaction, the key to gluconeogenesis inhibition by alcohol. Additionally, exercise requires enhanced carbohydrate catabolism for energy production and leads to an even greater lowering of the blood sugar level than that produced by alcohol ingestion alone.

10.2. (a) $\Delta G^{\circ\prime} = -61.9$ kJ mol^{-1}; (b) $\Delta G^{\circ\prime}$ $+13.8$ kJ mol^{-1}.

10.3. (a) 2 ATP; (b) 2 ATP; (c) 2 ATP; (d) 1 ATP.

10.4. 4 ATP.

10.5. 3.7×10^{21} molecules of ATP.

10.6. A high K_m means that pyruvate is not a particularly good substrate for the LDH isozyme; a relatively high [S] is required to attain $\frac{1}{2}V_{max}$. A low k_{cat} also shows that conversion of pyruvate to lactate does not proceed to a great extent.

10.7. [1,3-BPG]/[3PG] $= 3.0 \times 10^{-3}$.

10.8. Activation in (a) and (c), inhibition in (b). Activation of LDH increases the conversion of pyruvate to lactate and decreases its conversion to acetyl coenzyme A. Lower levels of acetyl coenzyme A decrease the activities of the citric acid cycle and the electron transport system. Inhibition of cytochrome oxidase inhibits the operation of the electron transport system. Activation of pyruvate carboxylase increases the conversion of pyruvate to oxaloacetate and decreases its conversion to acetyl coenzyme A. As in (a), lower levels of acetyl coenzyme A would decrease the activities of the citric acid cycle and the electron transport system.

10.9. Muscle phosphorylase leads to energy production via glycolysis; liver phosphorylase leads to synthesis of blood sugar. When glycogen catabolism is needed for energy production, such energy must be generated rapidly, hence the large maximum velocity of the muscle enzyme. Restoring the level of blood sugar, on the other hand, can be carried out at a slower rate, hence the smaller maximum velocity of the liver enzyme.

10.10. Administering ethanol leads to production of NADH and acetaldehyde via the alcohol dehydrogenase reaction. The NADH can then be used by the same enzyme to reduce toxic formaldehyde back to less harmful methanol. The acetaldehyde formed from ethanol can be metabolized further to acetate.

10.11. 0.55*M* glucose.

10.12. (a) C(3), the methyl-group carbon; (b) C(1), the carboxyl-group carbon; (c) C(1); (d) C(3).

10.13. a, c.

10.14. (a) A 7-carbon aldose and an 11-carbon 2-ketose; (b) an 8-carbon aldose and a 10-carbon 2-ketose.

10.15. 30 molecules of glucose.

10.16. 6 ATP.

10.17. A rate-determining step is the slowest step in a reaction sequence (smallest rate constant). Rate constants are unaffected by concentration changes of reactants and products. However, the rate of a chemical reaction does depend on reactant concentration. When glycogen breakdown occurs, the glucose level rises and the rate of the hexokinase reaction increases, leading to more extensive glycolysis and ATP production.

10.18. A rate-determining step is the slowest step in a reaction sequence (smallest rate constant); the sequence cannot proceed faster than the rate-determining step. Hexokinase controls the overall rate of glycolysis by controlling the concentration of glucose 6-phosphate. A committed step is generally a highly exergonic and essentially irreversible step in a reaction sequence; its occurrence ensures that all subsequent steps take place. Phosphofructokinase catalyzes the formation of fructose 1,6-bisphosphate, a compound whose sole metabolic role is to serve as a glycolytic intermediate and to ensure that the remaining steps take place.

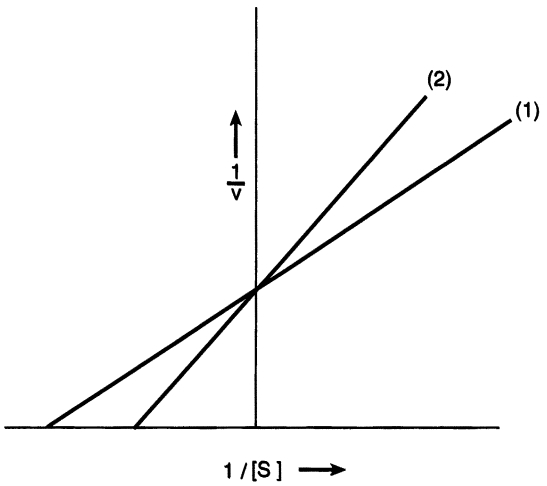
10.19. Yes, label will exchange because at equilibrium minute changes in one direction are balanced by minute changes in the opposite direction.

10.20. Because supply of the oxidized coenzyme is limited. NAD^+ must be regenerated from NADH so that subsequent substrate molecules can be oxidized.

CHAPTER 11

11.1. The ratios of NADH/NAD^+ and ATP/ADP are high during sleep and low during exercise.

11.2. (a)



(b) Yes, the inhibition can be decreased. Malonate acts as a competitive inhibitor of succinate, the substrate of succinate dehydrogenase. Competitive inhibition can be overcome by increasing the substrate concentration. Adding oxaloacetate leads to an increase in the concentrations of all citric acid cycle intermediates, of which succinate is one.

11.3. Citric acid cycle:

(a) Citrate + $2NAD^+$ + GDP + P_i + FAD + H_2O → malate + 2NADH + $2CO_2$ + GTP + $FADH_2$

(b) Acetyl CoA + oxaloacetate + H_2O → citrate + CoA—SH + H^+

(c) Isocitrate + $2NAD^+$ + GDP + P_i → succinate + 2NADH + $2CO_2$ + GTP

Glyoxylate cycle:

(a) Citrate + H_2O + acetyl CoA → succinate + CoA—SH + H^+ + malate

(b) Acetyl CoA + oxaloacetate + H_2O → citrate + CoA—SH + H^+

(c) Isocitrate → succinate + glyoxylate

11.4. [Isocitrate]/[citrate] = 0.133.

11.5. $\Delta G^{\circ\prime} = -868.4$ kJ mol^{-1}.

11.6. (a) NADH is a suitable inhibitor because it constitutes one of the main products of the citric acid cycle. ATP is likewise suitable because the NADH formed enters the electron transport system where it leads to production of ATP. (b) Yes, it is advantageous. Inhibition at the early stages of a reaction sequence automatically inhibits all of the remaining stages and eliminates unnecessary reactions.

11.7. Adding the dicarboxylic acids leads to an increase in the concentrations of all citric acid cycle intermediates and thereby stimulates the cycle's activity, resulting in increased CO_2 production.

11.8. No. The material balance of the cycle is such that two carbons are added in the form of the acetyl group of acetyl CoA and two carbons are released in the form of CO_2. No net retention of carbon atoms occurs, and hence no net synthesis of cycle intermediates is possible. The coenzyme A portion of acetyl CoA is released as coenzyme A.

11.9. Since a high [AMP] indicates a low [ATP], ATP synthesis is needed. By stimulating the pyruvate dehydrogenase complex, AMP leads to production of acetyl CoA from pyruvate. The acetyl CoA enters the citric acid cycle, and the NADH formed leads to ATP production via the electron transport system.

11.10. First turn: 0; second turn: one-half of the original intensity; third turn: one-quarter of the original intensity.

11.11. 1/4.

11.12. Competitive inhibition; a small K_i.

11.13. (a) 11 molecules of ATP per molecule of acetyl CoA; (b) 2.5 molecules of ATP per molecule of acetyl CoA.

11.4. [Isocitrate]/[citrate] = 0.133.

11.15. Feedback inhibition.

11.16. $\Delta G^{\circ\prime} = -33.8$ kJ mol^{-1}.

11.17. 45.3%.

11.18. One ATP.

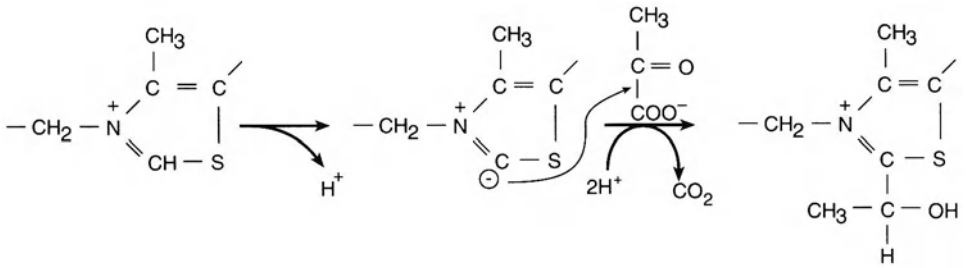
11.19. Their concentrations as cycle intermediates will decrease.

11.20. Yes, labeled CO_2 is released at step 4, the α-ketoglutarate dehydrogenase complex reaction.

11.21. Correct. Succinate, one of the intermediates of the citric acid cycle, is synthesized from acetyl CoA by the glyoxylate cycle.

11.22. Yes. Some malate would be diverted from the citric acid cycle and used for pyruvate synthesis. Increasing amounts of pyruvate yield added amounts of acetyl CoA for processing by the cycle but do not affect the levels of the cycle's intermediates.

11.23.



Vitamin B_1 constitutes the business end of TPP.

11.24. Because the citric acid cycle is obligatorily linked to the electron transport system via the NADH and $FADH_2$ produced by the cycle. The coenzymes are oxidized by the electron transport system, and this process is absolutely dependent on the presence of oxygen.

CHAPTER 12

12.1. 2,4-Dinitrophenol is an uncoupler of oxidative phosphorylation; it prevents ATP synthesis but permits electron transport to proceed. Low concentrations of 2,4-dinitrophenol decrease ATP synthesis. Accordingly, the level of electron transport activity must increase so that ATP demands of the cell can be met. Oxygen consumption increases, and the P/O ratio decreases.

12.2. 6.8×10^{-3}%.

12.3. Succinate^{2-} + dehydroascorbic acid ⇌ fumarate^{2-} + ascorbic acid; $\Delta E^{\circ\prime} = +0.03$ V, $\Delta G^{\circ\prime} = -5.8$ kJ mol^{-1}, and $K'_{bio} = 10.3$.

12.4. $\Delta G^{\circ\prime} = -0.074$ kJ.

12.5. (a) Ascorbic acid + fumarate^{2-} ⇄ dehydroascorbic acid + succinate^{2-}. (b) $\Delta E' = +0.116$ V; $\Delta G' = -22.4$ kJ mol^{-1}.

12.6. [Succinate]/[fumarate] = 40.

12.7. [Lactate]/[pyruvate] = 4.6×10^{-7}.

12.8. The nitrites administered oxidize Fe^{2+} of hemoglobin to Fe^{3+}, thereby forming methemoglobin: Hb(Fe^{2+}) → Hb(Fe^{3+}). Methemoglobin competes with cytochrome oxidase for cyanide. The competition is useful for treating cyanide poisoning because the amount of Hb(Fe^{3+}) that can be formed without impairing oxygen transport greatly exceeds the amount of poisoned cytochrome oxidase that can be tolerated.

12.9. [Succinate]/[fumarate] = 2.2×10^{-11}.

12.10. Soft drinks contain sugar that is catabolized via glycolysis, the citric acid cycle, and the electron transport system to CO_2 and H_2O. However, not all of the energy derived from these catabolic processes serves for ATP synthesis. Some energy becomes dissipated in the form of heat, thereby helping to maintain body temperature. Diet drinks lack sugar and cannot provide heat in this fashion.

12.11. 0.5.

12.12. Cells at rest have high levels of ATP and NADH and a low level of ADP. Accordingly, such cells have high energy charges. Actively metabolizing cells, on the other hand, have low levels of ATP and NADH and a high level of ADP. These cells have low energy charges.

12.13. 41%.

12.14. (a) 37%; (b) 80%; (c) 28%.

12.15. Succinate + $\frac{1}{2}O_2$ → fumarate + H_2O.

12.16. (a) 72 ATP; (b) 74 ATP.

12.17. Maximum number: 7; likely number: 2.

12.18. 2 ATP.

12.19. (a) Yes; (b) no; (c) no.

12.20. The $E^{\circ\prime}$ of NAD^+ is −0.32 V, and that of succinate is +0.03 V. In order for succinate dehydrogenase's coenzyme to oxidize succinate to fumarate, succinate must be a stronger reducing agent (have a smaller reduction potential) than the coenzyme. This relationship holds for succinate and FAD but not for succinate and NAD^+.

12.21. 2,4-Dinitrophenol is an uncoupler of oxidative phosphorylation; it prevents ATP synthesis but permits electron transport to proceed. (a) Administering 2,4-dinitrophenol leads to decreased ATP synthesis. Hence, electron transport activity and respiration must increase so that ATP demands of the cell can be met. This requires degradation of large amounts of additional metabolic fuels. Under dietary restrictions, additional metabolic fuels (carbohydrates and lipids) come from storage tissues. The breakdown of these tissues is expected to result in weight loss. (b) Some deaths occurred because so little ATP was synthesized that not enough was available to provide the energy required for maintenance of a living state. Also, metabolic fuel stores in these individuals may have been insufficient to provide the necessary additional metabolites. It is now known that 2,4-dinitrophenol is

highly toxic to humans because it is rapidly absorbed by all routes of administration (swallowing, contact with intact skin, and inhalation) and is not rapidly excreted; it has a biological half-life of 5–14 days. (c) In the presence of 2,4-dinitrophenol, there is a pronounced increase in electron transport activity, respiration, and general metabolism. Because little ATP is synthesized via oxidative phosphorylation, large amounts of heat are dissipated during operation of the electron transport system. This accounts for the rise in body temperature. Extensive electron transport activity also means that large amounts of water are produced, resulting in profuse sweating.

12.22. Electron transport and ATP synthesis are tightly coupled so that both processes are inhibited by DCCD. Adding 2,4-dinitrophenol uncouples the two processes; electron transport can proceed, but ATP synthesis is inhibited.

12.23. Three molecules of ATP per molecule of NADH.

12.24. 2,6-Dichlorophenol indophenol: between cytochrome *b* and Fe-S; TMPD: between cytochrome *c* and cytochrome *a*.

12.25. (a) 2; (b) 1; (c) 1.

12.26. 125,000 g; 2500 times.

12.27. (a) 7 ATP; (b) 37 ATP; (c) 16 ATP.

12.28. (a) 30 ATP; (b) 18 ATP.

CHAPTER 13

13.1. (a) 12 Acetyl CoA, 11 $FADH_2$, and 11 NADH; (b) 9 acetyl CoA, 9 $FADH_2$, and 9 NADH.

13.2. All of the carbons will become labeled.

13.3. 11 carbons.

13.4. 65 molecules of glucose.

13.5. β-Oxidation of fatty acids derived from adipose tissues produces large quantities of $FADH_2$ and NADH. When these coenzymes enter the electron transport system, metabolic water is produced.

13.6. (a) $CH_3(CH_2)_{18}COO^- + 9FAD + 9NAD^+ + 9H_2O + 10CoA{-}SH^{4-} + ATP^{4-} \rightarrow 10\ \text{acetyl CoA}^{4-} + 9FADH_2 + 9NADH + 9H^+ + AMP^{2-} + PP_i^{3-}$

(b) $CH_3(CH_2)_{13}COO^- + 6FAD + 6NAD^+ + 6H_2O + 7CoA{-}SH^{4-} + ATP^{4-} \rightarrow CH_3CH_2CO{-}S\text{-}CoA^{4-} + 6\ \text{acetyl CoA}^{4-} + 6FADH_2 + 6NADH + 6H^+ + AMP^{2-} + PP_i^{3-}$

13.7. All three mechanisms involve cleavage of a covalent bond by compound X such that one part of X adds to one cleavage product, and the other part of X adds to the second cleavage product. In hydrolysis, X = H_2O, and its two parts are H and OH; in phosphorolysis, X = H_3PO_4 (or some other form of P_i), and its two parts are H and H_2PO_4; in thiolysis, X = CoA—SH, and its two parts are H and CoA—S.

13.8. The enzyme cascade amplifies the original signal so that significant quantities of fatty acids can be mobilized rapidly, followed by catabolism to yield energy. Because fatty acid oxidation yields many ATPs, the cascade allows generation of large amounts of energy in a short time.

13.9. (a) 5 ATP; (b) 8 ATP; (c) 20 ATP.

13.10. (a) 42 ATP; (b) 461 ATP.

13.11. (a) 166,000 kJ; (b) 13.7 days; (c) 766 g/day.

13.12. 968 ml of H_2O.

13.13. 2.27 g of glycogen, or 8.63 g of glycogen and water.

13.14. a, b, c, e, g, i, j, and l.

13.15. Yes, the diabetic's breath is likely to contain labeled acetone. The labeled acetyl CoA administered adds to the pool of acetyl CoA. In diabetics, acetyl CoA pools tend to be large and are metabolized in part by forming ketone bodies, of which acetone is one.

13.16. (a) No $^{14}CO_2$ will be produced. (b) No $^{14}CO_2$ will be produced from the last acetyl CoA itself as it passes through one turn of the citric acid cycle. However, as this acetyl CoA passes through the cycle, $^{14}CO_2$ will be released because all of the cycle's intermediates have by now become labeled from previous turns. Labeled water will form in both (a) and (b) since steps 6 and 8 of the cycle (Figure 11.10) yield ^{3}H-labeled $FADH_2$ and (NADH + H^+). Oxidation of these coenzymes via the electron transport system produces labeled water.

13.17. (a) Glycerol + $2NAD^+$ + P_i + ADP → pyruvate + 2NADH + $2H^+$ + ATP + H_2O

(b) Propionyl CoA + HCO_3^- + ATP → succinyl CoA + ADP + P_i + H^+

(c) 2 Acetyl CoA + H_2O → acetoacetate + 2CoA—SH + H^+

13.18. The person is better off consuming odd-numbered fatty acids because their catabolism generates not only acetyl CoA but also propionyl CoA. Catabolism of acetyl CoA provides energy. Catabolism of propionyl CoA produces succinyl CoA and leads to an increase in the concentrations of the citric acid cycle's intermediates. One of these, oxaloacetate, can then be used in gluconeogenesis to provide some of the carbohydrate lacking in the person's diet.

13.19. By converting the hydrocarbons to carboxylic acids and then oxidizing these completely to CO_2 and H_2O. Because an oil spill consists of hydro-

carbons, use of such organisms will clean up the spill.

13.20. 24 ATP.

13.21. (a) Cycle 1; (b) cycle 1; (c) cycle 4; (d) cycles 1 and 2; (e) cycle 7.

13.22. $CH_3(CH_2)_6COOH + 11O_2 \rightarrow 8CO_2 + 8H_2O$

13.23. (b).

13.24. ~5 moles of palmitate/mole of glucose.

13.25. (a) 29 ATP; (b) 6 ATP and 12 (NADPH, H^+). Because NADPH can be converted to NADH, the requirement of 12 (NADPH, H^+) constitutes a potential loss of 36 ATP (assuming that no shuttle is needed), so that the total effective requirement is 42 ATP/molecule of myristic acid.

The ATP requirement for myristic acid synthesis and the ATP yield upon degradation of the fatty acid are not identical because an anabolic pathway can never be the exact reverse of the corresponding catabolic pathway. Fatty acid synthesis requires more energy than can be derived from fatty acid degradation. The same holds for glucose. Glucose synthesis from pyruvate requires 6 ATP equivalents per glucose, but glucose catabolism to pyruvate yields only 2 ATP per glucose.

13.26. Malonyl CoA will contain ^{14}C, but this is lost as $^{14}CO_2$ in step 2 of fatty acid biosynthesis, so that the palmitate produced will not be labeled.

13.27. (a) Three D at C(16); (b) two D at each of the following: C(14), C(12), C(10), C(8), C(6), C(4), and C(2).

13.28. Increased fatty acid mobilization; ketosis; fatty liver.

CHAPTER 14

14.1. (c).

14.2. Although arginine is synthesized in young organisms, its *rate* of synthesis is insufficient to meet the nutritional requirements.

14.3. (a) 9; (b) 32.

14.4. Glutamate labeled at the α-carbon.

14.5. Yes, both amino groups will become labeled. [^{15}N]Alanine can give rise to [^{15}N]aspartate by transamination and to $^{15}NH_4^+$ by oxidative deamination. $^{15}NH_4^+$ can lead to [^{15}N]carbamoyl phosphate by one of the ammonia fixation reactions. Since one of the amino groups in urea comes from aspartate while the other comes from carbamoyl phosphate, both of urea's amino groups will become labeled with ^{15}N.

14.6. D-Phenylalanine + $\frac{1}{2}O_2 \rightarrow$ phenylpyruvate + NH_4^+

14.7. (a) Nitrogenase; (b) nitrate reductase; (c) nitrite reductase; (d) carbamoyl phosphate synthase.

14.8. (a) 0; (b) +1; (c) +3; (d) +5; (e) −2; (f) −1; (g) −1.

14.9. Glucose + $2NAD^+$ + 2ADP + $2P_i$ + 2 aspartate
↓
2 alanine + 2 oxaloacetate + 2ATP + 2NADH + $2H^+$

14.10. IMP synthesis: 3 ATP; UMP synthesis: 2 ATP.

14.11. Gout is caused by abnormal metabolism of uric acid, the end product of purine catabolism in humans. Because amino acids are precursors of purines and pyrimidines, the protein-rich diet (the meat diet) is more likely to lead to excessive uric acid formation and to the development of gout.

14.12. (a) Glutamate$^-$ + NAD^+ ($NADP^+$) + H_2O
↓
α-ketoglutarate^{2-} + NADH (NADPH) + H^+ + NH_4^+

(b) Glutamate$^-$ + NH_4^+ + ATP^{4-} →
glutamine + ADP^{3-} + P_i^{2-} + H^+

(c) Phenylalanine + O_2 + tetrahydrobiopterin
↓
tyrosine + H_2O + quinonoid dihydrobiopterin

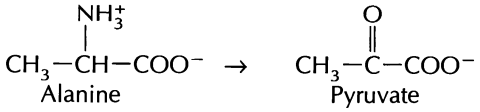
(d) Arginine$^+$ + H_2O → ornithine$^+$ + urea

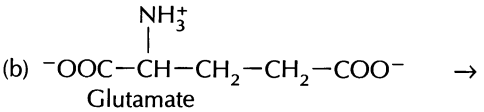
(e) Adenine + phosphoribosyl pyrophosphate^{5-}
↓
adenosine 5′-phosphate^{2-} + PP_i^{3-}

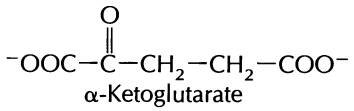
(f) Cytidine 5′-triphosphate^{4-} + reduced ribonucleotide reductase
↓
deoxycytidine 5′-triphosphate^{4-} + oxidized ribonucleotide reductase + H_2O

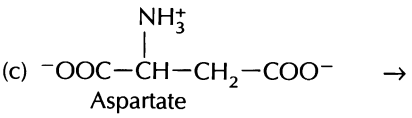
14.13. Yes for glutathione, no for iodoacetamide. Reduction of ribonucleotides requires the reduced form of ribonucleotide reductase. Because glutathione acts as an antioxidant, adding it will stabilize the enzyme in its reduced state and preserve the enzyme's critical sulfhydryl groups. Iodoacetamide is an irreversible inhibitor of the enzyme because it combines covalently with sulfhydryl groups. Adding this compound will inactivate the enzyme.

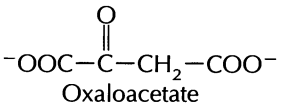
14.14. $^{14}CO_2$ does not form from [^{14}C]acetyl CoA during the first turn of the citric acid cycle but does form during subsequent turns. Because CO_2 in the form of HCO_3^- constitutes a precursor of carbamoyl phosphate, labeled carbamoyl phosphate (and hence labeled urea) will form as soon as $^{14}CO_2$ is produced.

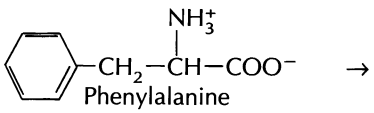
14.15. (a) 

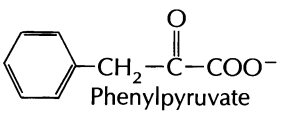
(b) 

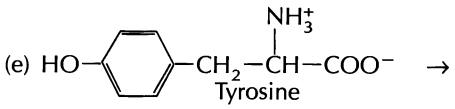


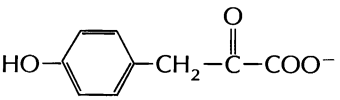
(c) 



(d) 



(e) 



p-Hydroxyphenylpyruvate

14.16. The transamination reaction yields pyruvate and glutamate. Inclusion of lactate dehydrogenase results in conversion of pyruvate to lactate, with the concomitant oxidation of NADH to NAD^+. Since NADH, but not NAD^+, absorbs at 340 nm, progress of the reaction can be followed by measuring the decrease in absorbance at that wavelength.

14.17. Three enzymatic steps convert propionyl CoA to succinyl CoA. The third enzyme in this sequence (methylmalonyl CoA mutase) requires a derivative of vitamin B_{12} as coenzyme. Because of the deficient absorption of vitamin B_{12} in pernicious anemia, methylmalonyl CoA cannot be metabolized further, and the catabolism of isoleucine, methionine, and valine is impaired. Researchers believe that the neurological deterioration associated with the disease results from interference of the accumulating methylmalonyl CoA with myelin sheath formation.

14.18. 15 ATP/molecule of alanine.

14.19. Fasting lowers the level of blood sugar. Subsequent administration of amino acids results in rapid catabolism of glucogenic amino acids. Oxidative deamination of the amino acids produces ammonia, but the arginine deficiency prevents adequate conversion of ammonia to urea. As a result, the concentration of ammonia in the blood increases sharply.

14.20. Because of their extensive nucleic acid breakdown, afflicted individuals have high levels of uric acid that can produce kidney stones and gout. Allopurinol, an inhibitor of xanthine oxidase, prevents uric acid formation. Administering allopurinol decreases uric acid accumulation and lowers the chance that patients will develop kidney stones or gout.

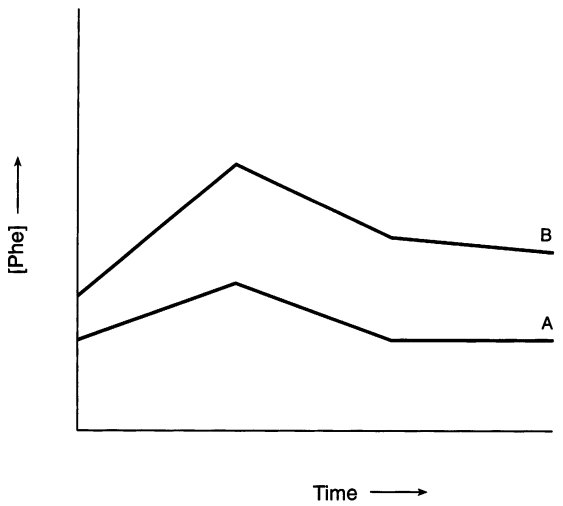
14.21. HGPRT deficiency causes an increase in the concentration of PRPP, and high levels of PRPP lead to excessive production of purines and pyrimidines. The effect is especially pronounced for purines, whose synthesis essentially begins with PRPP. High purine concentrations result in enhanced purine catabolism. The resultant accumulation of uric acid can lead to kidney stones and gout. Why HGPRT deficiency leads to the neurological and behavioral symptoms described is not known.

14.22. A high-protein diet results in increased production of urea. Drinking large amounts of water increases the volume of urine and allows for excretion of the urea from the body in the form of relatively dilute solutions. This puts less strain on the kidneys than would the excretion of more concentrated urea solutions.

14.23. Three energy-rich bonds.

14.24. Five α-amino acids participate in the urea cycle: ornithine, citrulline, aspartate, argininosuccinate, and arginine. Of these, only aspartate and arginine are used for protein biosynthesis.

14.25.



14.26. Since aspartame is a derivative of phenylalanine, the individual would be better advised to use saccharin.

14.27. Skin pigmentation results from the presence of melanins, which are polymeric subtances derived from tyrosine. Since kwashiorkor is a disease of protein deficiency, the lack of tyrosine prevents melanin formation.

14.28. 117 g/day; 16.4%.

14.29. Amino acid catabolism accounts for the difference. Compared to degradation of carbohydrates, amino acid degradation to pyruvate or to intermediates of the citric acid cycle lacks the energy-yielding steps of glycolysis. Compared to degradation of lipids, amino acid degradation generates much less acetyl CoA. Lastly, while carbohydrate and lipid catabolism are exergonic, a good part of protein catabolism—the urea cycle—is endergonic.

14.30. The rapid growth characteristic of many cancer cells requires an above-average supply of nucleotides. Consequently, such cells are particularly sensitive to inhibitors of purine or pyrimidine biosynthesis. Because viruses also replicate rapidly, viral systems may be affected similarly by action of the same inhibitors.

CHAPTER 15

15.1. No, since a free energy change calculated from $\Delta E^{\circ\prime}$ yields $\Delta G^{\circ\prime}$. To properly evaluate intracellular reactions, $\Delta G'$ values must be used.

15.2. $2H_2O \rightarrow 4H^+ + 4e^- + O_2$
$QH_2 + 2(\text{cyt } bf)_{ox} \rightarrow Q + 2(\text{cyt } bf)_{red} + 2H^+$

15.3. Since photorespiration decreases the efficiency of photosynthesis, the mouse in (a) will last longer than that in (b). The mouse in (c) will last the longest because C_4 plants carry out photosynthesis with greater efficiency than C_3 plants.

15.4. 2.5×10^{11} tons per year.

15.5. No. Both photosystems involve only visible light; they are unaffected by ultraviolet radiation, which has shorter wavelengths. Also, the two known effects of light on photosynthesis—red drop and the Emerson enhancement effect—require light in the visible range.

15.6. 2/3= 0.67.

15.7. Carotenoids: ~400–500 nm, violet-blue; phycocyanins: ~600 nm, yellow; phycoerythrins: ~500–600 nm, green-yellow.

15.8.

PSI: $3NADP^+ + 3H^- + 3H^+ \xrightarrow{6\ h\nu} 3NADPH + 3H^+$
$(6H^+ + 6e^-)$

PSII: $3H_2O \xrightarrow{6\ h\nu} O_3 + 6H^+ + 6e^-$

Overall: $3NADP^+ + 3H_2O \xrightarrow{12\ h\nu} 3NADPH + 3H^+ + O_3$

Six photons must strike each photosystem, and six electrons must flow through the Z-scheme.

15.9. 0.5.

15.10. 26.7%.

15.11. (a) ~4; (b) ~3.

15.12. No, since *Rhodopseudomonas viridis* uses a cyclic electron flow that involves only one photosystem at 870 nm.

15.13. $\Delta G^{\circ\prime} = -96.5$ kJ mol^{-1}. One photon can excite one molecule of P870.

15.14. 1/6. Fractions of ^{14}C in fructose 6-phosphate: 1/3; 1/6; 0.

15.15. The colors of the accessory pigments are normally masked by that of chlorophyll. When leaves lose their chlorophyll, colors of the accessory pigments become noticeable, accounting for the multiple colors of fall foliage.

15.16. 11.9%.

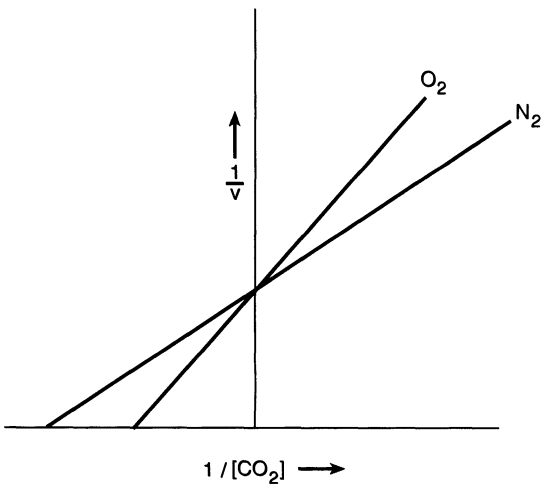
15.17. $\Delta E^{\circ\prime} = +0.11$ V; $\Delta G^{\circ\prime} = -21$ kJ mol^{-1}.

15.18. Photosynthesis of the C_3 plant leads to a decrease in the concentration of CO_2. As the concentration of CO_2 falls, photorespiration becomes more pronounced and lowers the photosynthetic efficiency. Increased photorespiration also results in appreciable water loss by the plant. By contrast, the C_4 plant avoids wasteful photorespiration and suffers minimal water loss.

15.19. No, since the sequence $Q_A \rightarrow QH_2$ is not involved in the electron transport system of cyclic photophosphorylation.

15.20. Ferricyanide acts as an artificial electron acceptor, accepting electrons from one of the plastoquinones (Q) preceding QH_2 (Figure 15.9).

15.21.



Oxygen acts as a competitive inhibitor of the enzyme. The apparent K_m (O_2 atmosphere) is greater than the true K_m (N_2 atmosphere) of the enzyme.

15.22. $\Delta G' = -55.9$ kJ mol^{-1}.

15.23. $\lambda = 385$ nm.

15.24. $\Delta E' = 0.29$ V.
15.25. $\Delta E' = 2.5$ V.
15.26. (a) Impossible. Since there is no external electron donor, net NADPH production cannot occur. (b) Possible. While the scheme is theoretically plausible, it is practically unlikely because it requires a pigment capable of generating a very large potential difference of about 2 V. (c) Possible. (d) Possible.

CHAPTER 16

16.1. Using the layout of Figure 16.11, you can show that the minimum number of different tRNAs required to "recognize" the amino acid codons is 32:

	U	C	A	G
U	2	2	1	2
C	2	2	2	2
A	3	2	2	2
G	2	2	2	2

16.2. In the mitochondrial code, a single codon is used for each of the eight four-codon families (Leu, Val, Ser, Pro, Thr, Ala, Arg, and Gly). Other changes are shown in Table 16.4. Using the same layout as in the previous problem, you can show that the minimum number of different tRNAs required to "recognize" the amino acid codons is 22:

	U	C	A	G
U	2	1	1	2
C	1	1	2	1
A	2	1	2	1
G	1	1	2	1

16.3. 5′-CTG-3′.
16.4. 3′-GAC-5′.
16.5. Three different polypeptide chains form depending on the point at which translation of the oligonucleotide begins. The bulk of all three polypeptide chains is an alternating sequence of serine and leucine residues, but the chains differ in the structures of their N- and C-termini:

$$\begin{array}{l} \text{AGC}-(\text{UCU}-\text{CUC})_n-\cdots-\text{CUG}-\text{G} \\ \text{Ser}\quad\ (\text{Ser}\quad\text{Leu})_n\qquad\qquad\ \text{Leu} \\ \text{A}-\text{GCU}-(\text{CUC}-\text{UCU})_n-\cdots-\text{UGG} \\ \quad\text{Ala}\quad\ (\text{Leu}\quad\text{Ser})_n\qquad\qquad\ \text{Trp} \\ \text{AG}-(\text{CUC}-\text{UCU})_n-\cdots-\text{UCU}-\text{GG} \\ \quad\ (\text{Leu}\quad\text{Ser})_n\qquad\qquad\ \text{Ser} \end{array}$$

16.6. $p = 0.0338$ or 3.4%.
16.7.

Base composition	Sequence	Probability	(%)
[UUC]: Two U, one C; probability = (0.80)(0.80)(0.20)	UCU	12.8	
	UUC	12.8	
	CUU	12.8	
			38.4
[CCU]: Two C, one U; probability = (0.20)(0.20)(0.80)	CCU	3.2	
	CUC	3.2	
	UCC	3.2	
			9.6
UUU: Three U; probability = (0.80)(0.80)(0.80)	UUU	51.2	
			51.2
CCC: Three C; probability = (0.20)(0.20)(0.20)	CCC	0.8	
			0.8
			100.0

16.8. If the oligonucleotide is translated in the 5′ → 3′ direction, the product is oligoalanine having threonine as its N-terminus. If translation is in the 3′ → 5′ direction, the product is oligoproline having threonine as its C-terminus.
16.9. Translation of the oligonucleotide in the 5′ → 3′ direction yields the oligopeptide Cys(Leu)_n. Because the leucine sequence increases in length with increasing translation time, the data prove that the polypeptide chain grows from the N- to the C-terminus.
16.10. Lack of base complementarity indicates that the DNA is single-stranded. Resistance to attack by either phosphodiesterase I or II indicates that the DNA has neither a free 5′-OH nor a free 3′-OH. You conclude that the DNA is single-stranded and circular.
16.11. N-Thr-Leu-Thr-Asp-Cys-Pro-Arg-C.
16.12. (a) N-Thr-Leu-Asp-Gly-Leu-Pro-C.
(b) N-Thr-Leu-Thr-Ala-Ala-Leu-C.
16.13. The nature of the peptide product depends on the point at which translation of the oligonucleotide begins: (a) Polyarginine, polyglutamic acid, and polylysine; (b) polyglutamic acid and polylysine.
16.14. Changing normal to sickle-cell hemoglobin in-

volves replacing a glutamic acid (codons GAA and GAG) by valine (codons GUU, GUC, GUA, and GUG). Two codon conversions can be accomplished by single base mutations: GAA → GUA and GAG → GUG.

16.15. 51 bp.

16.16. Four types of intact ribosomes can form:

"All heavy": Heavy 30S + heavy 50S
"All light": Light 30S + light 50S
Hybrid: Heavy 30S + light 50S
Hybrid: Light 30S + heavy 50S

16.17. Phenylalanine.

16.18. 5′-AAUCGUC-3′.

16.19. Isoleucine, threonine, or methionine.

16.20. In the overlapping code of Problem 16.19, most bases are used three times. Each base occurs in three different codons and at a different position in each: 5′-end, middle, 3′-end. In these cases, a single base mutation leads to three altered codons. Unless two of the altered codons are synonym codons, the mutation results in three amino acid changes in the corresponding protein.

16.21. a, b, and c.

16.22. (a) 5′-CCU-3′; Pro
5′-CCC-3′; Pro
5′-CCA-3′; Pro
(b) 5′-UCA-3′; Ser
5′-UCG-3′; Ser
(c) 5′-UGU-3′; Cys
5′-UGC-3′; Cys

16.23. a, c, e, and f.

16.24. mRNA: [U] = 33.3 mol %, [G] = 66.7 mol %; DNA: [A] = [T] = 16.7 mol %, [C] = [G] = 33.3 mol %.

16.25. Working backward from tryptophan, you can establish the following underlined sequence of single point mutations:

Leu	→	Pro	→	Ala	→	Gly	→	Trp
CUU		CCU		GCU		GGU		UGG
CUC		CCC		GCC		GGC		
CUA		CCA		GCA		GGA		
CUG	→	CCG	→	GCG	→	GGG		
UUA								
UUG								

16.26. 5′-ACT-3′.

16.27. Five times.

16.28. Arginine: [AGG]; lysine: [AAA].

16.29. 312 base pairs.

CHAPTER 17

17.1. Since A/T-rich segments have fewer hydrogen bonds than G/C-rich segments, they are easier to unwind; local unwinding of the double helix is a prerequisite of replication.

17.2. DNA replication requires the participation of RNA; it cannot proceed without the aid of RNA primers. This suggests that a simpler mechanism involving RNA replication evolved first and subsequently changed to the more complex mechanism of DNA replication.

17.3. (a) 5′ U—G—A—C—A—G—C—U—A—C—C—U—G—G—U 3′
(b) 5′ T—G—A—C—A—G—C—T—A—C—C—T—G—G—T 3′
(c) 5′ U—A—A—C—A—A—C—U—A—C—C—U—A—A—U 3′

17.4. At t_1: 1000 nucleotides; at t_2: 1818 nucleotides.

17.5. 6580 cells.

17.6. Heme and testosterone because they have planar structures.

17.7. 4000 Okazaki fragments.

17.8. 1.28 years.

17.9. Because DNA is entirely double-stranded whereas RNA is either single-stranded or only partially double-stranded. The double-stranded structure allows for accurate DNA replication through synthesis of two complementary and antiparallel strands. It also allows for ready DNA repair.

17.10. Hybrid DNA (ds): 2/8 or 25%; "all light" DNA (ds): 6/8 or 75%; "heavy" single strands: 2/16 or 12.5%; "light" single strands: 14/16 or 87.5%.

17.11. a, d, and f. The β- and γ-phosphate groups of dTTP are cleaved out in the form of PP_i as dTMP is incorporated into the growing polynucleotide strand. Hence, only dTTP forms containing labeled α-phosphate groups are useful for the assay.

17.12. (a) ••••••·····; (b) ·····••••••·····

17.13. B-DNA: 4.0×10^5 turns; Z-DNA: 3.3×10^5 turns.

17.14. *E. coli:* 0.31 s; eukaryotes: 4.7 s.

17.15. 0.12 cm/h.

17.16. 5.6×10^4 molecules of glucose.

17.17. "All heavy" DNA (ds): none; hybrid DNA (ds): 2/16 or 12.5%; "all light" DNA (ds): 14/16 or 87.5%; "heavy" single strands: 2/32 or 6.25%; "light" single strands: 30/32 or 93.75%.

17.18. 250 replication forks.

17.19. 5 mg; 2.5%.

17.20. Phenylalanine and leucine.

17.21. (a) 5.0×10^4 bp/min; (b) 17 μm/min.

17.22. DNA polymerase I and DNA ligase.

17.23. G·C.

17.24. 5′ G—A—T—C—A—T—A—T—G—A—T—C 3′
3′ C—T—A—G—T—A—T—A—C—T—A—G 5′

17.25. Because of its structural similarity to deoxyribonu-

cleoside triphosphates (the normal substrates of DNA polymerase), 5-fluorouridine triphosphate is likely to act as a competitive inhibitor of the enzyme.

17.26. Because nuclear DNA controls most cell functions whereas mitochondrial DNA is limited to coding for mitochondrial protein synthesis and a few other proteins.

17.27. FAD, FMN, CoA—SH, or $NADP^+$ might function in the reaction because each of these molecules contains an AMP moiety. GTP might substitute for ATP as the two compounds differ only in the nature of their purine ring. Lastly, AMP itself might interact directly with the enzyme.

17.28. 1.733 g/cm^3.

CHAPTER 18

18.1. (a) An inducer/repressor complex in enzyme induction and a co-repressor/repressor complex in enzyme repression. The inducer/repressor complex prevents the repressor from binding to the operator, thereby allowing transcription to proceed during enzyme induction. The co-repressor/repressor complex binds to the operator and turns it off, thereby stopping transcription during enzyme repression. (b) A cAMP–CAP complex. This complex binds to the promoter, stimulates attachment of RNA polymerase, and increases transcription of the *lac* operon so that lactose is metabolized.

18.2. (a) 40; (b) 59; (c) 60.

18.3. 5.04×10^6.

18.4. 2.51×10^{-15} g.

18.5. 1.89×10^6.

18.6. DNA base sequence:

3′-TAC-TAC-ACC-ACC-TAC-ACC-ACC-5′

5′-ATG-ATG-TGG-TGG-ATG-TGG-TGG-3′

DNA base composition: [A] = [T] = 10/42 = 23.8 mol %; [C] = [G] = 11/42 = 26.2 mol %.

18.7. Determine the base composition of the 16S rRNA. RNA transcribed from both strands of DNA must show base complementarity so that [A] = [U] and [G] = [C].

18.8. Both terms may be applied to the inducer of the *lac* operon. The inducer counteracts the repressor; it functions as an antirepressor. In the absence of inducer, the operator is blocked and the operon is repressed. Thus, the inducer acts as a de-repressor. The co-repressor of the *trp* operon also functions as an antirepressor in the sense that it counteracts the repressor, but it does not function as a de-repressor.

18.9. An incorrect conclusion. The three genes are transcribed as a unit, and the mRNA is translated as a unit, resulting in synthesis of the three gene products in equal numbers. However, the accumulation of each protein in the cell is a function of its rate of degradation or half-life.

18.10. Because attenuation requires tight coupling of transcription and translation. This coupling, which occurs in prokaryotes, allows the two processes to proceed simultaneously in the same subcellular compartment. Coupling of transcription and translation is impossible in eukaryotes because the two processes take place in separate subcellular compartments.

18.11. The *trp* attenuator contains two tryptophan codons, but the *his* attenuator contains seven histidine codons. Accordingly, a larger number of amino acid molecules are needed to terminate transcription by attenuation for the *his* operon than for the *trp* operon. If you add an equal number of amino acid molecules, attenuation of the *trp* operon will be more pronounced.

18.12. 5′-AGC-CAU-UCG-AGG-UGU-UCA-CGU-AAA-3′.

18.13. Ser-His-Ser-Arg-Cys-Ser-Arg-Lys.

18.14. For any given section of double-stranded DNA, only one of the two strands is transcribed into RNA.

18.15. (a) 24.0 s; (b) 53.3 s.

18.16. Since the mRNA has a short half-life, the label remaining after a prolonged time (40%) represents that taken up by other types of RNA. Accordingly, the amount taken up by mRNA is approximately 100% − 40% = 60%. All of the mRNA label is lost within about 5 min (300 s) from the start of the experiment. About half of the amount incorporated into mRNA (30%) is lost within about 180 s from the start of the experiment or within (180 − 60) = 120 s from the time the chase was added. By this time, the incorporation into mRNA has dropped by 30% (one-half of the maximum 60%) from 100% to 70%, indicating an approximate half-life of 120 s.

18.17. Further transcription is inhibited because cordycepin 5′-triphosphate lacks the free 3′-OH required for elongation of the strand. The inhibition indicates that the polynucleotide strand grows in the 5′ → 3′ direction.

18.18. The sigma subunit would compete with the holoenzyme for binding to the promoter. Because the holoenzyme is essential for initiation of transcription, the sigma subunit would act as a competitive inhibitor and lead to a decrease in the rate of transcription.

18.19. The core enzyme by itself must bind more tightly to the DNA than the core enzyme associated with the sigma subunit. Failure of the sigma subunit to dissociate from the holoenzyme would make the mutant enzyme less processive, resulting in slower RNA synthesis.

18.20. Deletion of the regulatory gene means that the repressor is not synthesized. Consequently, both lac-

tose and tryptophan synthesis would be continuous. In the case of the *lac* operon, the operator is not turned off because of the absence of a repressor. In the case of the *trp* operon, the operator is not turned off because of the absence of a co-repressor/repressor complex.

18.21. Determine the distribution of poly(A) tail lengths in bulk mRNA. Finding that the tail in all mRNAs has a uniform size would argue against both parts of the proposed hypothesis. Finding that the tails have varying lengths would support part (a). Finding that the distribution of tail lengths is cut off at a particular minimal value (5 or some other value) would support part (b).

18.22. [A] = [T] = 20 mol %; [C] = [G] = 30 mol %.

18.23. (a) In the absence of active β-galactoside permease, lactose cannot be transported from the medium into the cells. Consequently, lactose cannot be utilized to any significant extent. (b) Lactose utilization by the cells will increase.

18.24. (a) Synthesis of a protein from an mRNA transcript of genes C, B, and A. (b) No tryptophan synthesis.

18.25. [A] = 25 mol %; [U] = 19 mol %; [C] = [G] = 28 mol %.

18.26. mRNA generally has a short half-life or high turnover; it is synthesized and degraded rapidly.

18.27. The intact double-stranded segment most likely represents that section of the DNA that binds the RNA polymerase holoenzyme. Binding of the RNA polymerase protects the DNA against digestion by deoxyribonuclease. Determining the base sequence of the intact segment helps to identify the promoter site of the DNA.

CHAPTER 19

19.1. Codon UGU.

19.2. By inhibiting a specific protein-synthesizing system with puromycin (for example, reticulocytes synthesizing hemoglobin, or bacterial cells induced to produce β-galactosidase) and assaying the peptidyl puromycin formed for protein structure and/or activity. If the protein begins to fold before its synthesis is complete, the peptidyl puromycin may possess a portion of the protein's final three-dimensional structure. Depending on what section of the molecule has folded, the peptidyl puromycin may also possess some or all of the protein's activity.

19.3. 0.25 mg.

19.4. (a) 3; (b) 8; (c) 5; (d) 4; (e) 12; (f) 1; (g) 9; (h) 2; (i) 6; (j) 11; (k) 10; (l) 7.

19.5. 18 amino acids.

19.6. 9 amino acids.

19.7. (b).

19.8. No energy-rich bonds are expended at this step; $\Delta G' \approx 0$.

19.9. Both act on the 50S subunit of prokaryotic ribosomes and block chain elongation.

19.10. 90 nucleotides per second.

19.11. 600 energy-rich bonds.

19.12. 3600 energy-rich bonds.

19.13. (a) All protein molecules from both experiments have the same N-terminal amino acid. (b) In the absence of puromycin, all protein molecules have the same C-terminal amino acid. In the presence of puromycin, partially completed chains of the protein end in various C-terminal amino acids. (c) In the presence of puromycin, the average chain length is shorter because polypeptide chains are prematurely terminated.

19.14. 6.4 kg.

19.15. (a) 0.990; (b) 0.912.

19.16. Met-Pro-Asp-Phe-Met-Val-Gln-Leu-Ser-Ser-Glu-Met.

19.17. (a) Met-Ser-Asp-Phe-Met-Val-Gln-Leu-Ser-Ser-Glu-Met.
(b) Met-Pro-Asp-Phe-Met-Val-Gln-Leu-Ser-Ser-Glu-Met.
(c) Met-Pro-Asp-Phe-Met-Val.

19.18. 0.42 min (*E. coli*); 8.3 min (rabbit reticulocytes).

19.19. Their mechanisms of action involve a kinase that catalyzes the phosphorylation of initiation factor eIF-2, thereby blocking translation. Heme-controlled inhibitor is a kinase of this type, whereas interferons may induce the enzyme.

19.20. (a) and (c). In preparation (b), protein synthesis will not be stimulated because eukaryotic initiation requires methionyl-tRNA. In preparation (c), protein synthesis will be stimulated because mitochondrial initiation also proceeds with *N*-formylmethionyl-tRNA.

19.21. (a) 3′-CAG-5′; (b) 3′-CUG-5′.

19.22. The enzyme must have at least three binding sites for an incoming amino acid, ATP, and tRNA. In addition, the enzyme must bind the two products, aminoacyl adenylate and aminoacyl-tRNA.

19.23. (a) A, C, B; (b) X, Z, Y.

19.24. (d) and (e).

19.25. Based on the wobble hypothesis (Table 16.5):

Codons:	5′-CGU-3′	5′-CGC-3′	5′-CGA-3′	5′-CGG-3′
Anticodons:	3′-GCA-5′	<u>3′-GCG-5′</u>	3′-GCU-5′	<u>3′-GCC-5′</u>
	3′-GCG-5′	3′-GCI-5′	3′-GCI-5′	3′-GCU-5′
	3′-GCI-5′			

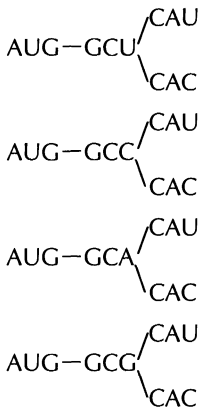
GC base pairs are linked via three hydrogen bonds; all others (AU, AI, UI, CI, and GU) are linked via two hydrogen bonds. Accordingly, the underlined anticodons result in the strongest hydrogen bonding with their codons. These anticodon–codon pairings release the tRNA more slowly, resulting in slower rates of protein synthesis.

19.26. Met-Leu-Asn-Leu-Arg-Tyr-Pro.

19.27. Modify the bacterial DNA so that the protein's gene

leads to synthesis of a polypeptide chain enriched in nonpolar amino acids at the N-terminus. If at least 10–15 of the protein's first 30 amino acids are hydrophobic in nature, the N-terminal segment might function as a signal peptide. In that case, synthesis of the protein would be completed by membrane-bound ribosomes, and, once synthesized, the protein would be secreted across the cell membrane.

19.28. There are eight possible mRNA sequences:



19.29. Since leucine has six codons, adding it to the tripeptide leads to $6 \times 8 = 48$ possible mRNA sequences. Adding valine (four codons) to the tetrapeptide raises the number of possible sequences to $48 \times 4 = 192$.

19.30. You should purify phenylalanyl-tRNA synthetase. It is thought that an aminoacyl-tRNA synthetase can screen out, by hydrolysis, mischarged amino acids *smaller* than the correct amino acid. Since glycine is the smallest amino acid, glycyl-tRNA synthetase may not possess such a hydrolytic site, and studying this enzyme might not be useful.

Index

Page numbers in **boldface** refer to structural formulas. F indicates a figure, T a table. Positional and configurational designations in chemical names are either omitted or disregarded in alphabetizing