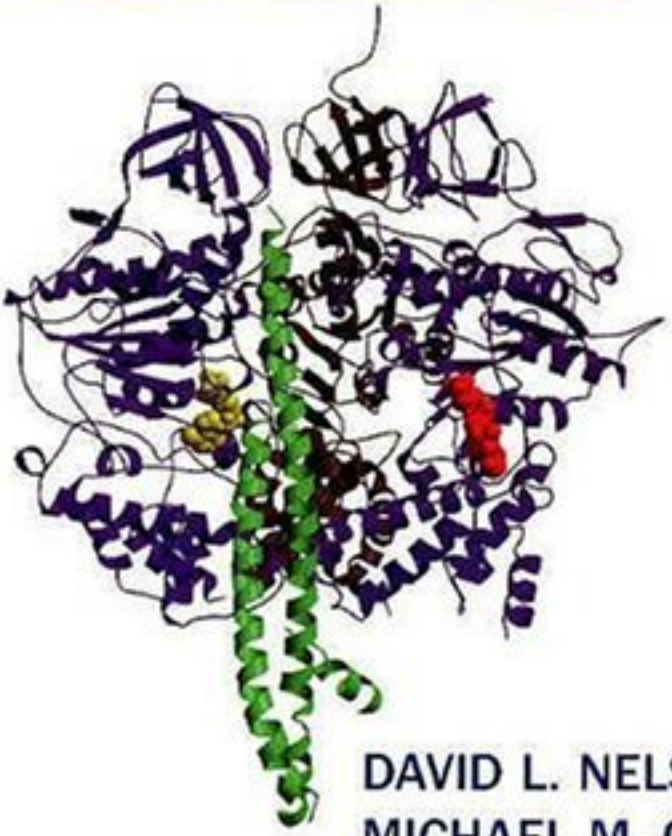
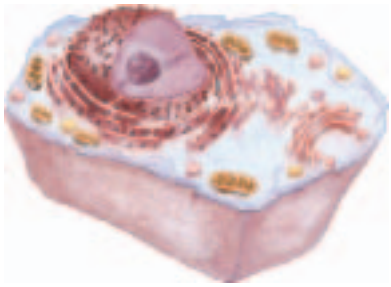
Lehninger

PRINCIPLES OF BIOCHEMISTRY

fourth edition



DAVID L. NELSON
MICHAEL M. COX



chapter 1

THE FOUNDATIONS OF BIOCHEMISTRY

With the cell, biology discovered its atom . . . To characterize life, it was henceforth essential to study the cell and analyze its structure: to single out the common denominators, necessary for the life of every cell; alternatively, to identify differences associated with the performance of special functions.

—*François Jacob,* La logique du vivant: une histoire de l'hérédité (The Logic of Life: A History of Heredity), *1970*

We must, however, acknowledge, as it seems to me, that man with all his noble qualities . . . still bears in his bodily frame the indelible stamp of his lowly origin.

—*Charles Darwin,* The Descent of Man, *1871*

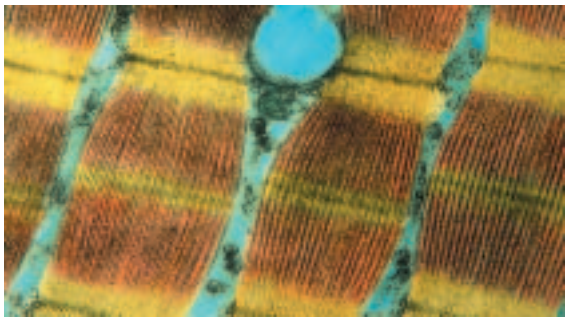
Fifteen to twenty billion years ago, the universe arose as a cataclysmic eruption of hot, energy-rich subatomic particles. Within seconds, the simplest elements (hydrogen and helium) were formed. As the universe expanded and cooled, material condensed under the influence of gravity to form stars. Some stars became enormous and then exploded as supernovae, releasing the energy needed to fuse simpler atomic nuclei into the more complex elements. Thus were produced, over billions of years, the Earth itself and the chemical elements found on the Earth today. About four billion years ago, life arose—simple microorganisms with the ability to extract energy from organic compounds or from sunlight, which they used to make a vast array of more complex **biomolecules** from the simple elements and compounds on the Earth's surface.

Biochemistry asks how the remarkable properties of living organisms arise from the thousands of different lifeless biomolecules. When these molecules are isolated and examined individually, they conform to all the physical and chemical laws that describe the behavior of inanimate matter—as do all the processes occurring in living organisms. The study of biochemistry shows how the collections of inanimate molecules that constitute living organisms interact to maintain and perpetuate life animated solely by the physical and chemical laws that govern the nonliving universe.

Yet organisms possess extraordinary attributes, properties that distinguish them from other collections of matter. What are these distinguishing features of living organisms?

A high degree of chemical complexity and microscopic organization. Thousands of different molecules make up a cell's intricate internal structures (Fig. 1–1a). Each has its characteristic sequence of subunits, its unique three-dimensional structure, and its highly specific selection of binding partners in the cell.

Systems for extracting, transforming, and using energy from the environment (Fig. 1–1b), enabling organisms to build and maintain their intricate structures and to do mechanical, chemical, osmotic, and electrical work. Inanimate matter tends, rather, to decay toward a more disordered state, to come to equilibrium with its surroundings.



(a)

(b)

(c)

FIGURE 1–1 Some characteristics of living matter. (a) Microscopic complexity and organization are apparent in this colorized thin section of vertebrate muscle tissue, viewed with the electron microscope. **(b)** A prairie falcon acquires nutrients by consuming a smaller bird. **(c)** Biological reproduction occurs with near-perfect fidelity.

A capacity for precise self-replication and self-assembly (Fig. 1–1c). A single bacterial cell placed in a sterile nutrient medium can give rise to a billion identical "daughter" cells in 24 hours. Each cell contains thousands of different molecules, some extremely complex; yet each bacterium is a faithful copy of the original, its construction directed entirely from information contained within the genetic material of the original cell.

Mechanisms for sensing and responding to alterations in their surroundings, constantly adjusting to these changes by adapting their internal chemistry.

Defined functions for each of their components and regulated interactions among them. This is true not only of macroscopic structures, such as leaves and stems or hearts and lungs, but also of microscopic intracellular structures and individual chemical compounds. The interplay among the chemical components of a living organism is dynamic; changes in one component cause coordinating or compensating changes in another, with the whole ensemble displaying a character beyond that of its individual parts. The collection of molecules carries out a program, the end result of which is reproduction of the program and self-perpetuation of that collection of molecules—in short, life.

A history of evolutionary change. Organisms change their inherited life strategies to survive in new circumstances. The result of eons of evolution is an enormous diversity of life forms, superficially very different (Fig. 1–2) but fundamentally related through their shared ancestry.

Despite these common properties, and the fundamental unity of life they reveal, very few generalizations about living organisms are absolutely correct for every organism under every condition; there is enormous diversity. The range of habitats in which organisms live, from hot springs to Arctic tundra, from animal intestines to college dormitories, is matched by a correspondingly wide range of specific biochemical adaptations, achieved

FIGURE 1–2 Diverse living organisms share common chemical features. Birds, beasts, plants, and soil microorganisms share with humans the same basic structural units (cells) and the same kinds of macromolecules (DNA, RNA, proteins) made up of the same kinds of monomeric subunits (nucleotides, amino acids). They utilize the same pathways for synthesis of cellular components, share the same genetic code, and derive from the same evolutionary ancestors. Shown here is a detail from "The Garden of Eden," by Jan van Kessel the Younger (1626–1679).

within a common chemical framework. For the sake of clarity, in this book we sometimes risk certain generalizations, which, though not perfect, remain useful; we also frequently point out the exceptions that illuminate scientific generalizations.

Biochemistry describes in molecular terms the structures, mechanisms, and chemical processes shared by all organisms and provides organizing principles that underlie life in all its diverse forms, principles we refer to collectively as *the molecular logic of life.* Although biochemistry provides important insights and practical applications in medicine, agriculture, nutrition, and industry, its ultimate concern is with the wonder of life itself.

In this introductory chapter, then, we describe (briefly!) the cellular, chemical, physical (thermodynamic), and genetic backgrounds to biochemistry and the overarching principle of evolution—the development over generations of the properties of living cells. As you read through the book, you may find it helpful to refer back to this chapter at intervals to refresh your memory of this background material.

1.1 Cellular Foundations

The unity and diversity of organisms become apparent even at the cellular level. The smallest organisms consist of single cells and are microscopic. Larger, multicellular organisms contain many different types of cells, which vary in size, shape, and specialized function. Despite these obvious differences, all cells of the simplest and most complex organisms share certain fundamental properties, which can be seen at the biochemical level.

Cells Are the Structural and Functional Units of All Living Organisms

Cells of all kinds share certain structural features (Fig. 1–3). The **plasma membrane** defines the periphery of the cell, separating its contents from the surroundings. It is composed of lipid and protein molecules that form a thin, tough, pliable, hydrophobic barrier around the cell. The membrane is a barrier to the free passage of inorganic ions and most other charged or polar compounds. Transport proteins in the plasma membrane allow the passage of certain ions and molecules; receptor proteins transmit signals into the cell; and membrane enzymes participate in some reaction pathways. Because the individual lipids and proteins of the plasma membrane are not covalently linked, the entire structure is remarkably flexible, allowing changes in the shape and size of the cell. As a cell grows, newly made lipid and protein molecules are inserted into its plasma membrane; cell division produces two cells, each with its own membrane. This growth and cell division (fission) occurs without loss of membrane integrity.

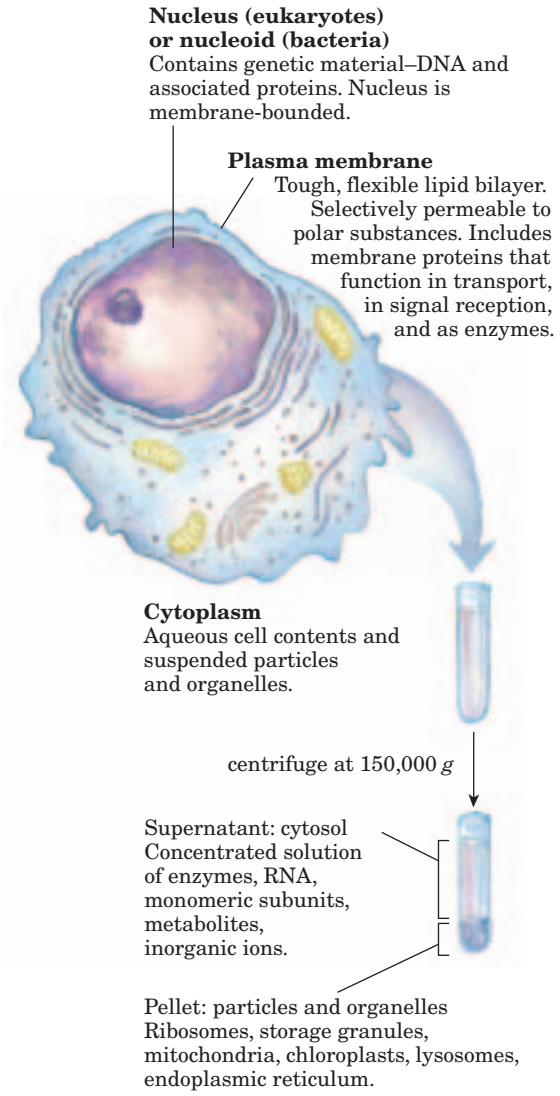


FIGURE 1–3 The universal features of living cells. All cells have a nucleus or nucleoid, a plasma membrane, and cytoplasm. The cytosol is defined as that portion of the cytoplasm that remains in the supernatant after centrifugation of a cell extract at 150,000 *g* for 1 hour.

The internal volume bounded by the plasma membrane, the **cytoplasm** (Fig. 1–3), is composed of an aqueous solution, the **cytosol,** and a variety of suspended particles with specific functions. The cytosol is a highly concentrated solution containing enzymes and the RNA molecules that encode them; the components (amino acids and nucleotides) from which these macromolecules are assembled; hundreds of small organic molecules called **metabolites,** intermediates in biosynthetic and degradative pathways; **coenzymes,** compounds essential to many enzyme-catalyzed reactions; inorganic ions; and **ribosomes,** small particles (composed of protein and RNA molecules) that are the sites of protein synthesis.

All cells have, for at least some part of their life, either a **nucleus** or a **nucleoid,** in which the **genome**—

the complete set of genes, composed of DNA—is stored and replicated. The nucleoid, in bacteria, is not separated from the cytoplasm by a membrane; the nucleus, in higher organisms, consists of nuclear material enclosed within a double membrane, the nuclear envelope. Cells with nuclear envelopes are called **eukaryotes** (Greek *eu,* "true," and *karyon,* "nucleus"); those without nuclear envelopes—bacterial cells—are **prokaryotes** (Greek *pro,* "before").

Cellular Dimensions Are Limited by Oxygen Diffusion

Most cells are microscopic, invisible to the unaided eye. Animal and plant cells are typically 5 to 100 μm in diameter, and many bacteria are only 1 to 2 μm long (see the inside back cover for information on units and their abbreviations). What limits the dimensions of a cell? The lower limit is probably set by the minimum number of each type of biomolecule required by the cell. The smallest cells, certain bacteria known as mycoplasmas, are 300 nm in diameter and have a volume of about 10^{-14} mL. A single bacterial ribosome is about 20 nm in its longest dimension, so a few ribosomes take up a substantial fraction of the volume in a mycoplasmal cell.

The upper limit of cell size is probably set by the rate of diffusion of solute molecules in aqueous systems. For example, a bacterial cell that depends upon oxygen-consuming reactions for energy production must obtain molecular oxygen by diffusion from the surrounding medium through its plasma membrane. The cell is so small, and the ratio of its surface area to its volume is so large, that every part of its cytoplasm is easily reached by O_2 diffusing into the cell. As cell size increases, however, surface-to-volume ratio decreases, until metabolism consumes O_2 faster than diffusion can supply it. Metabolism that requires O_2 thus becomes impossible as cell size increases beyond a certain point, placing a theoretical upper limit on the size of the cell.

There Are Three Distinct Domains of Life

All living organisms fall into one of three large groups (kingdoms, or domains) that define three branches of evolution from a common progenitor (Fig. 1–4). Two large groups of prokaryotes can be distinguished on biochemical grounds: **archaebacteria** (Greek *archē,* "origin") and **eubacteria** (again, from Greek *eu,* "true"). Eubacteria inhabit soils, surface waters, and the tissues of other living or decaying organisms. Most of the well-studied bacteria, including *Escherichia coli,* are eubacteria. The archaebacteria, more recently discovered, are less well characterized biochemically; most inhabit extreme environments—salt lakes, hot springs, highly acidic bogs, and the ocean depths. The available evidence suggests that the archaebacteria and eubacteria diverged early in evolution and constitute two separate

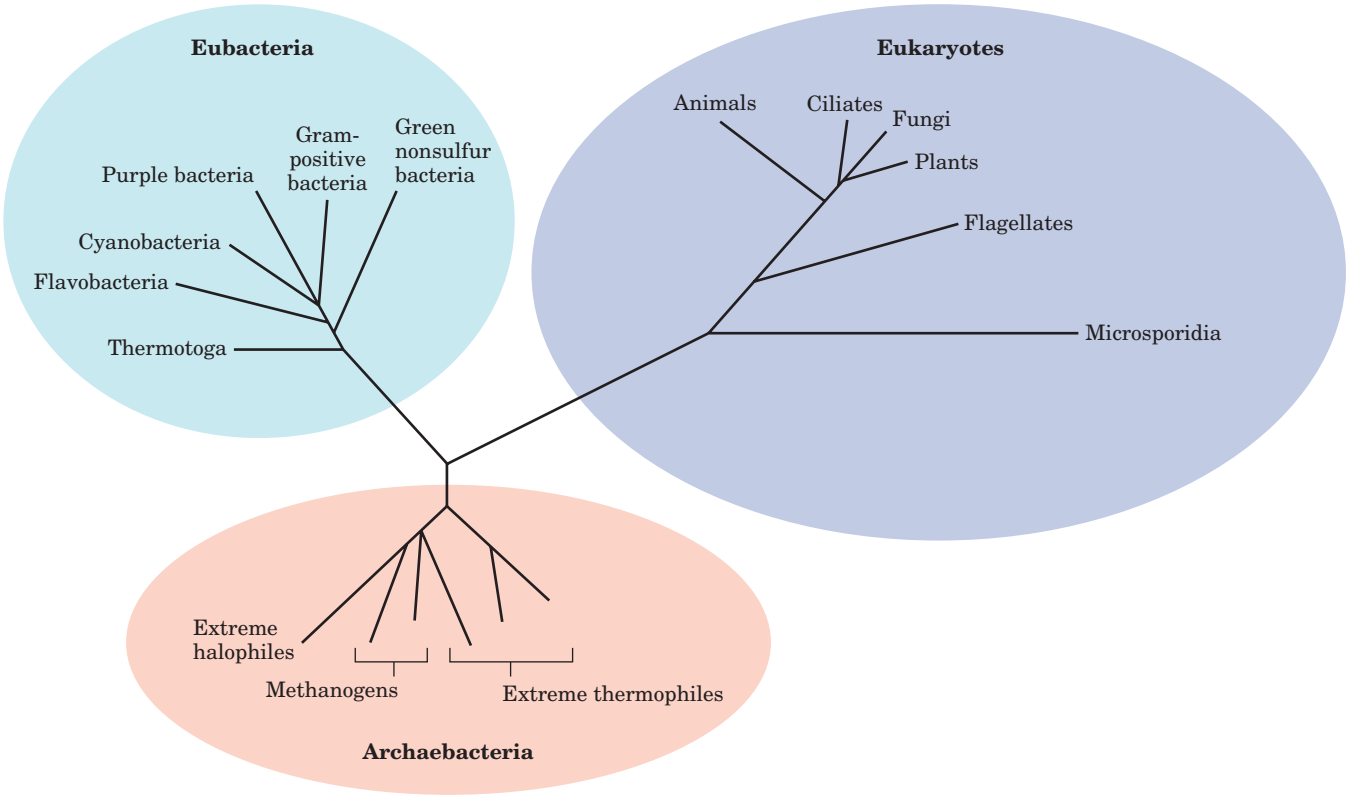


FIGURE 1–4 Phylogeny of the three domains of life. Phylogenetic relationships are often illustrated by a "family tree" of this type. The fewer the branch points between any two organisms, the closer is their evolutionary relationship.

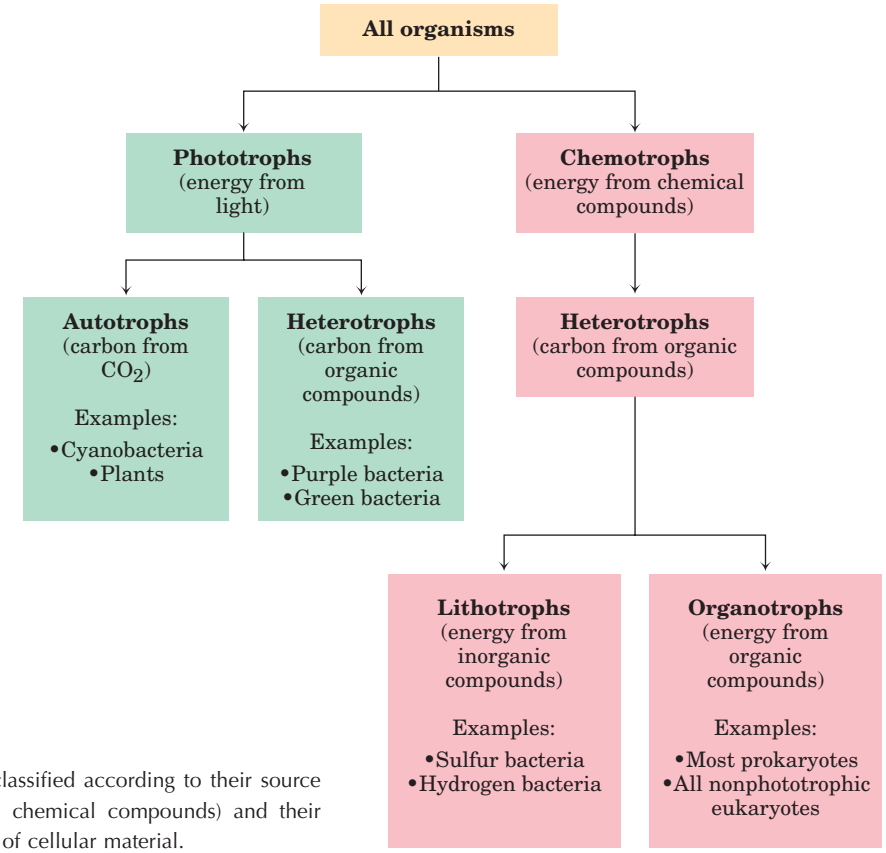


FIGURE 1–5 Organisms can be classified according to their source of energy (sunlight or oxidizable chemical compounds) and their source of carbon for the synthesis of cellular material.

domains, sometimes called Archaea and Bacteria. All eukaryotic organisms, which make up the third domain, Eukarya, evolved from the same branch that gave rise to the Archaea; archaebacteria are therefore more closely related to eukaryotes than to eubacteria.

Within the domains of Archaea and Bacteria are subgroups distinguished by the habitats in which they live. In **aerobic** habitats with a plentiful supply of oxygen, some resident organisms derive energy from the transfer of electrons from fuel molecules to oxygen. Other environments are **anaerobic,** virtually devoid of oxygen, and microorganisms adapted to these environments obtain energy by transferring electrons to nitrate (forming N_2), sulfate (forming H_2S), or CO_2 (forming CH_4). Many organisms that have evolved in anaerobic environments are *obligate anaerobes:* they die when exposed to oxygen.

We can classify organisms according to how they obtain the energy and carbon they need for synthesizing cellular material (as summarized in Fig. 1–5). There are two broad categories based on energy sources: **phototrophs** (Greek *trophē,* "nourishment") trap and use sunlight, and **chemotrophs** derive their energy from oxidation of a fuel. All chemotrophs require a source of organic nutrients; they cannot fix CO_2 into organic compounds. The phototrophs can be further divided into those that can obtain all needed carbon from CO_2 **(autotrophs)** and those that require organic nutrients **(heterotrophs).** No chemotroph can get its carbon atoms exclusively from CO_2 (that is, no chemotrophs are autotrophs), but the chemotrophs may be further classified according to a different criterion: whether the fuels they oxidize are inorganic **(lithotrophs)** or organic **(organotrophs).**

Most known organisms fall within one of these four broad categories—autotrophs or heterotrophs among the photosynthesizers, lithotrophs or organotrophs among the chemical oxidizers. The prokaryotes have several general modes of obtaining carbon and energy. *Escherichia coli,* for example, is a chemoorganoheterotroph; it requires organic compounds from its environment as fuel and as a source of carbon. Cyanobacteria are photolithoautotrophs; they use sunlight as an energy source and convert CO_2 into biomolecules. We humans, like *E. coli*, are chemoorganoheterotrophs.

Escherichia coli Is the Most-Studied Prokaryotic Cell

Bacterial cells share certain common structural features, but also show group-specific specializations (Fig. 1–6). *E. coli* is a usually harmless inhabitant of the human intestinal tract. The *E. coli* cell is about 2 μm long and a little less than 1 μm in diameter. It has a protective outer membrane and an inner plasma membrane that encloses the cytoplasm and the nucleoid. Between the inner and outer membranes is a thin but strong layer of polymers called peptidoglycans, which gives the cell its shape and rigidity. The plasma membrane and the

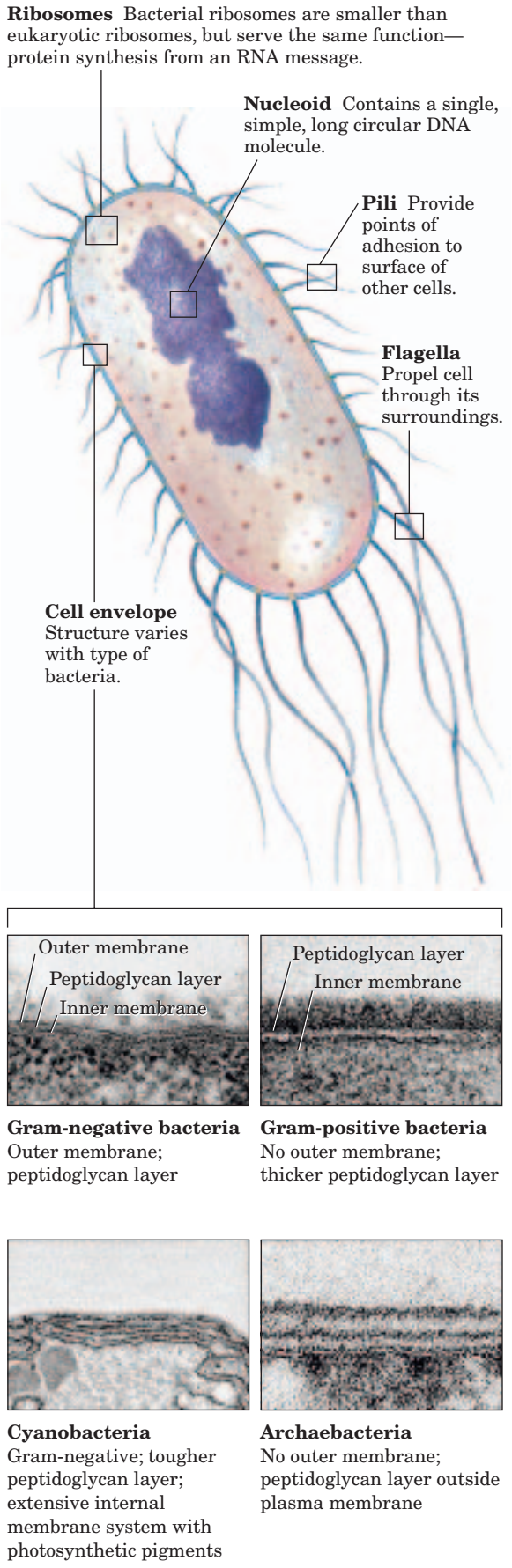


FIGURE 1–6 Common structural features of bacterial cells. Because of differences in the cell envelope structure, some eubacteria (gram-positive bacteria) retain Gram's stain, and others (gram-negative bacteria) do not. *E. coli* is gram-negative. Cyanobacteria are also eubacteria but are distinguished by their extensive internal membrane system, in which photosynthetic pigments are localized. Although the cell envelopes of archaebacteria and gram-positive eubacteria look similar under the electron microscope, the structures of the membrane lipids and the polysaccharides of the cell envelope are distinctly different in these organisms.

layers outside it constitute the **cell envelope.** In the Archaea, rigidity is conferred by a different type of polymer (pseudopeptidoglycan). The plasma membranes of eubacteria consist of a thin bilayer of lipid molecules penetrated by proteins. Archaebacterial membranes have a similar architecture, although their lipids differ strikingly from those of the eubacteria.

The cytoplasm of *E. coli* contains about 15,000 ribosomes, thousands of copies each of about 1,000 different enzymes, numerous metabolites and cofactors, and a variety of inorganic ions. The nucleoid contains a single, circular molecule of DNA, and the cytoplasm (like that of most bacteria) contains one or more smaller, circular segments of DNA called **plasmids.** In nature, some plasmids confer resistance to toxins and antibiotics in the environment. In the laboratory, these DNA segments are especially amenable to experimental manipulation and are extremely useful to molecular geneticists.

Most bacteria (including *E. coli*) lead existences as individual cells, but in some bacterial species cells tend to associate in clusters or filaments, and a few (the myxobacteria, for example) demonstrate simple social behavior.

Eukaryotic Cells Have a Variety of Membranous Organelles, Which Can Be Isolated for Study

Typical eukaryotic cells (Fig. 1–7) are much larger than prokaryotic cells—commonly 5 to 100 μm in diameter, with cell volumes a thousand to a million times larger than those of bacteria. The distinguishing characteristics of eukaryotes are the nucleus and a variety of membrane-bounded organelles with specific functions: mitochondria, endoplasmic reticulum, Golgi complexes, and lysosomes. Plant cells also contain vacuoles and chloroplasts (Fig. 1–7). Also present in the cytoplasm of many cells are granules or droplets containing stored nutrients such as starch and fat.

In a major advance in biochemistry, Albert Claude, Christian de Duve, and George Palade developed methods for separating organelles from the cytosol and from each other—an essential step in isolating biomolecules and larger cell components and investigating their

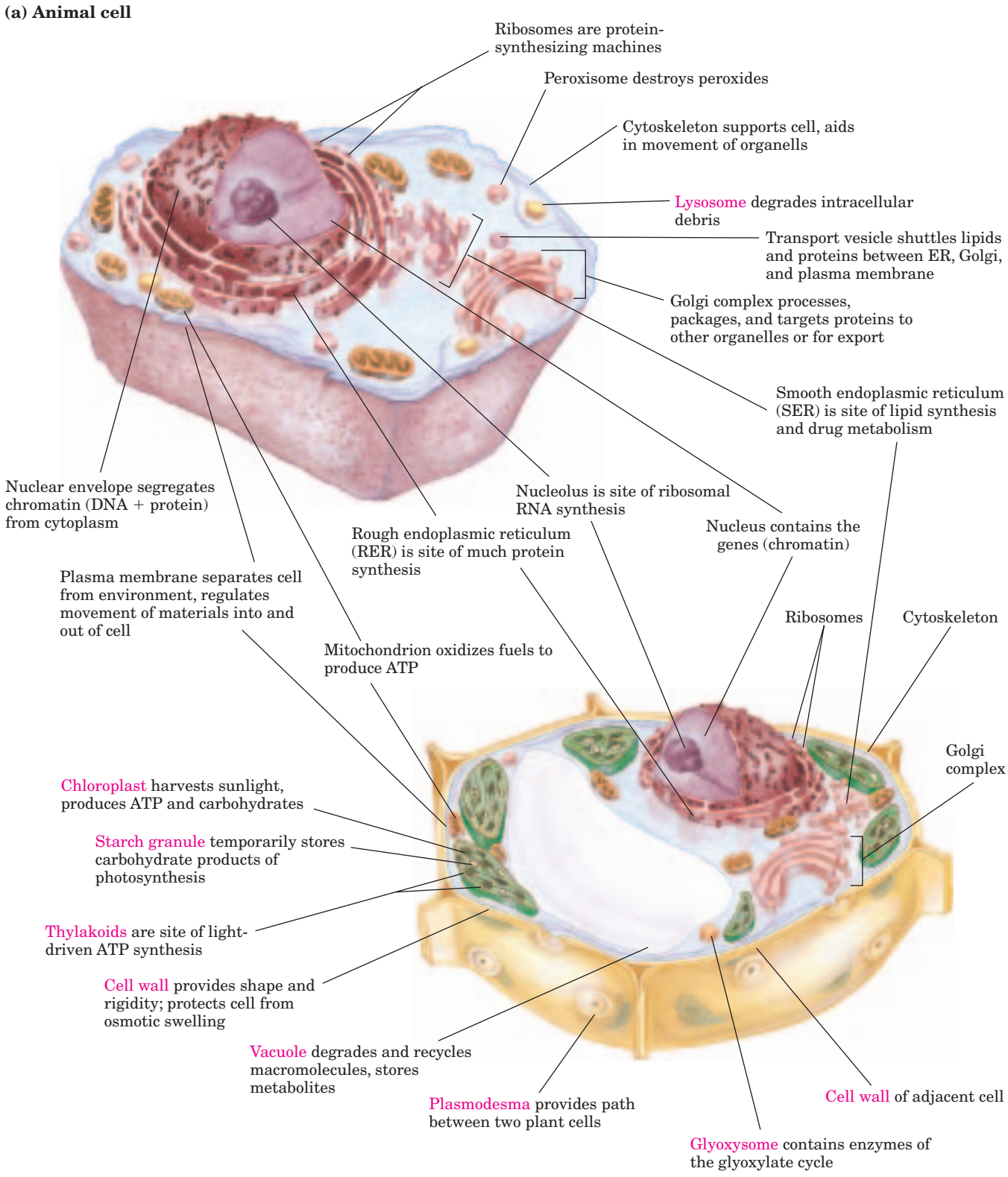
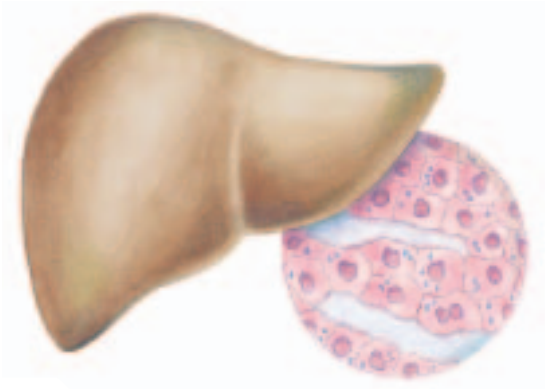
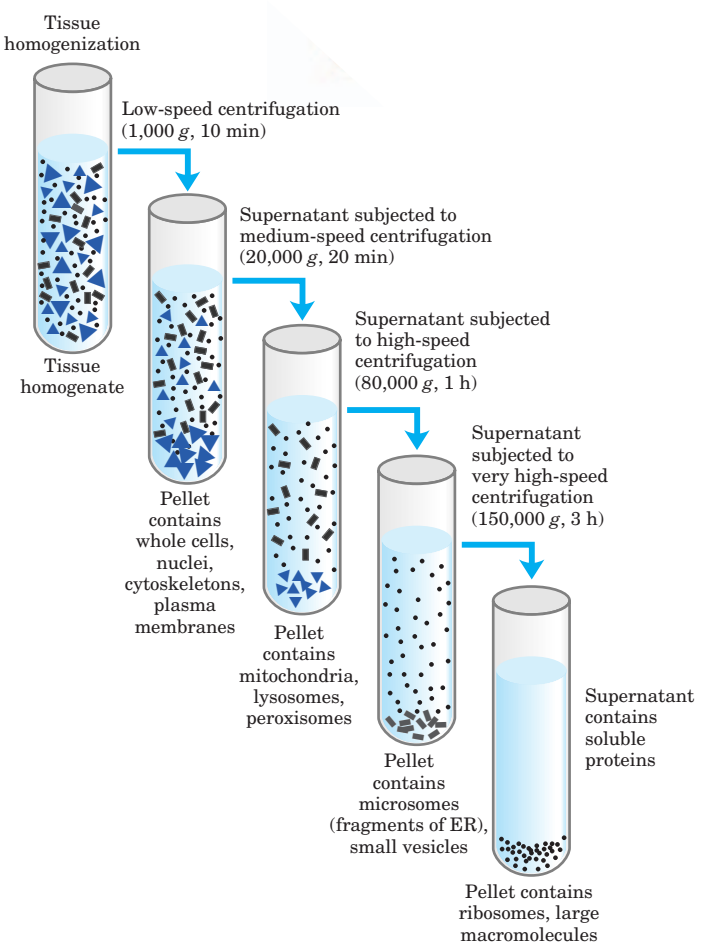


FIGURE 1–7 Eukaryotic cell structure. Schematic illustrations of the two major types of eukaryotic cell: **(a)** a representative animal cell and **(b)** a representative plant cell. Plant cells are usually 10 to 100 μm in diameter—larger than animal cells, which typically range from 5 to 30 μm. Structures labeled in red are unique to either animal or plant cells.

structures and functions. In a typical cell fractionation (Fig. 1–8), cells or tissues in solution are disrupted by gentle homogenization. This treatment ruptures the plasma membrane but leaves most of the organelles intact. The homogenate is then centrifuged; organelles such as nuclei, mitochondria, and lysosomes differ in size and therefore sediment at different rates. They also differ in specific gravity, and they "float" at different levels in a density gradient.

Differential centrifugation results in a rough fractionation of the cytoplasmic contents, which may be further purified by isopycnic ("same density") centrifugation. In this procedure, organelles of different buoyant densities (the result of different ratios of lipid and protein in each type of organelle) are separated on a density gradient. By carefully removing material from each region of the gradient and observing it with a microscope, the biochemist can establish the sedimentation position of each organelle





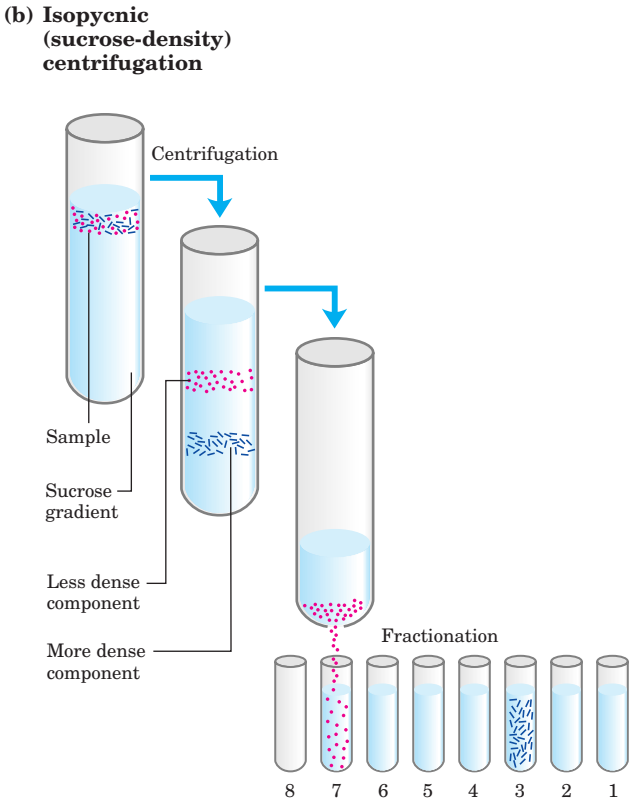


FIGURE 1–8 Subcellular fractionation of tissue. A tissue such as liver is first mechanically homogenized to break cells and disperse their contents in an aqueous buffer. The sucrose medium has an osmotic pressure similar to that in organelles, thus preventing diffusion of water into the organelles, which would swell and burst. **(a)** The large and small particles in the suspension can be separated by centrifugation at different speeds, or **(b)** particles of different density can be separated by isopycnic centrifugation. In isopycnic centrifugation, a centrifuge tube is filled with a solution, the density of which increases from top to bottom; a solute such as sucrose is dissolved at different concentrations to produce the density gradient. When a mixture of organelles is layered on top of the density gradient and the tube is centrifuged at high speed, individual organelles sediment until their buoyant density exactly matches that in the gradient. Each layer can be collected separately.

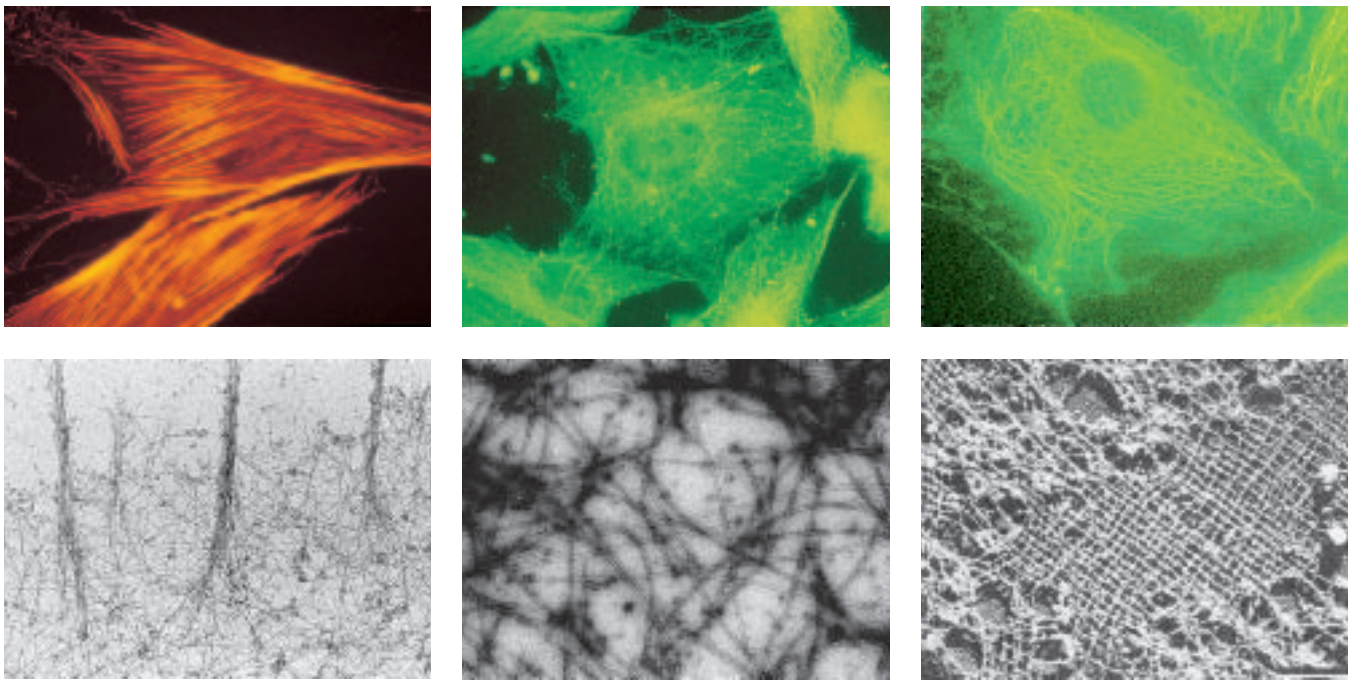
and obtain purified organelles for further study. For example, these methods were used to establish that lysosomes contain degradative enzymes, mitochondria contain oxidative enzymes, and chloroplasts contain photosynthetic pigments. The isolation of an organelle enriched in a certain enzyme is often the first step in the purification of that enzyme.

The Cytoplasm Is Organized by the Cytoskeleton and Is Highly Dynamic

Electron microscopy reveals several types of protein filaments crisscrossing the eukaryotic cell, forming an interlocking three-dimensional meshwork, the **cytoskeleton.** There are three general types of cytoplasmic filaments—actin filaments, microtubules, and intermediate filaments (Fig. 1–9)—differing in width (from about 6 to 22 nm), composition, and specific function. All types provide structure and organization to the cytoplasm and shape to the cell. Actin filaments and microtubules also help to produce the motion of organelles or of the whole cell.

Each type of cytoskeletal component is composed of simple protein subunits that polymerize to form filaments of uniform thickness. These filaments are not permanent structures; they undergo constant disassembly into their protein subunits and reassembly into filaments. Their locations in cells are not rigidly fixed but may change dramatically with mitosis, cytokinesis, amoeboid motion, or changes in cell shape. The assembly, disassembly, and location of all types of filaments are regulated by other proteins, which serve to link or bundle the filaments or to move cytoplasmic organelles along the filaments.

The picture that emerges from this brief survey of cell structure is that of a eukaryotic cell with a meshwork of structural fibers and a complex system of membrane-bounded compartments (Fig. 1–7). The filaments disassemble and then reassemble elsewhere. Membranous vesicles bud from one organelle and fuse with another. Organelles move through the cytoplasm along protein filaments, their motion powered by energy dependent motor proteins. The **endomembrane system** segregates specific metabolic processes and provides surfaces on which certain enzyme-catalyzed reactions occur. **Exocytosis** and **endocytosis,** mechanisms of transport (out of and into cells, respectively) that involve membrane fusion and fission, provide paths between the cytoplasm and surrounding medium, allowing for secretion of substances produced within the cell and uptake of extracellular materials.



Actin stress fibers **(a)** | Microtubules **(b)** | Intermediate filaments **(c)**

FIGURE 1–9 The three types of cytoskeletal filaments. The upper panels show epithelial cells photographed after treatment with antibodies that bind to and specifically stain **(a)** actin filaments bundled together to form "stress fibers," **(b)** microtubules radiating from the cell center, and **(c)** intermediate filaments extending throughout the cytoplasm. For these experiments, antibodies that specifically recognize actin, tubulin, or intermediate filament proteins are covalently attached to a fluorescent compound. When the cell is viewed with a fluorescence microscope, only the stained structures are visible. The lower panels show each type of filament as visualized by **(a, b)** transmission or **(c)** scanning electron microscopy.

Although complex, this organization of the cytoplasm is far from random. The motion and the positioning of organelles and cytoskeletal elements are under tight regulation, and at certain stages in a eukaryotic cell's life, dramatic, finely orchestrated reorganizations, such as the events of mitosis, occur. The interactions between the cytoskeleton and organelles are noncovalent, reversible, and subject to regulation in response to various intracellular and extracellular signals.

Cells Build Supramolecular Structures

Macromolecules and their monomeric subunits differ greatly in size (Fig. 1–10). A molecule of alanine is less than 0.5 nm long. Hemoglobin, the oxygen-carrying protein of erythrocytes (red blood cells), consists of nearly 600 amino acid subunits in four long chains, folded into globular shapes and associated in a structure 5.5 nm in diameter. In turn, proteins are much smaller than ribosomes (about 20 nm in diameter), which are in turn much smaller than organelles such as mitochondria, typically 1,000 nm in diameter. It is a long jump from simple biomolecules to cellular structures that can be seen

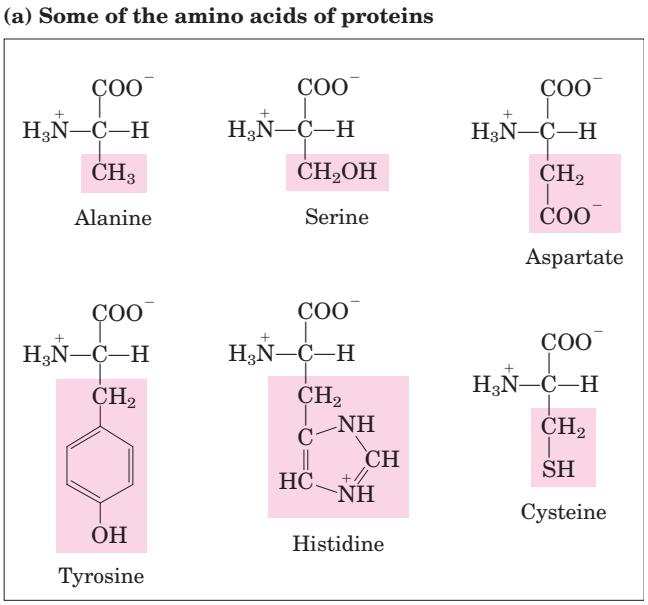


FIGURE 1–10 The organic compounds from which most cellular materials are constructed: the ABCs of biochemistry. Shown here are **(a)** six of the 20 amino acids from which all proteins are built (the side chains are shaded pink); **(b)** the five nitrogenous bases, two five-carbon sugars, and phosphoric acid from which all nucleic acids are built; **(c)** five components of membrane lipids; and **(d)** D-glucose, the parent sugar from which most carbohydrates are derived. Note that phosphoric acid is a component of both nucleic acids and membrane lipids.

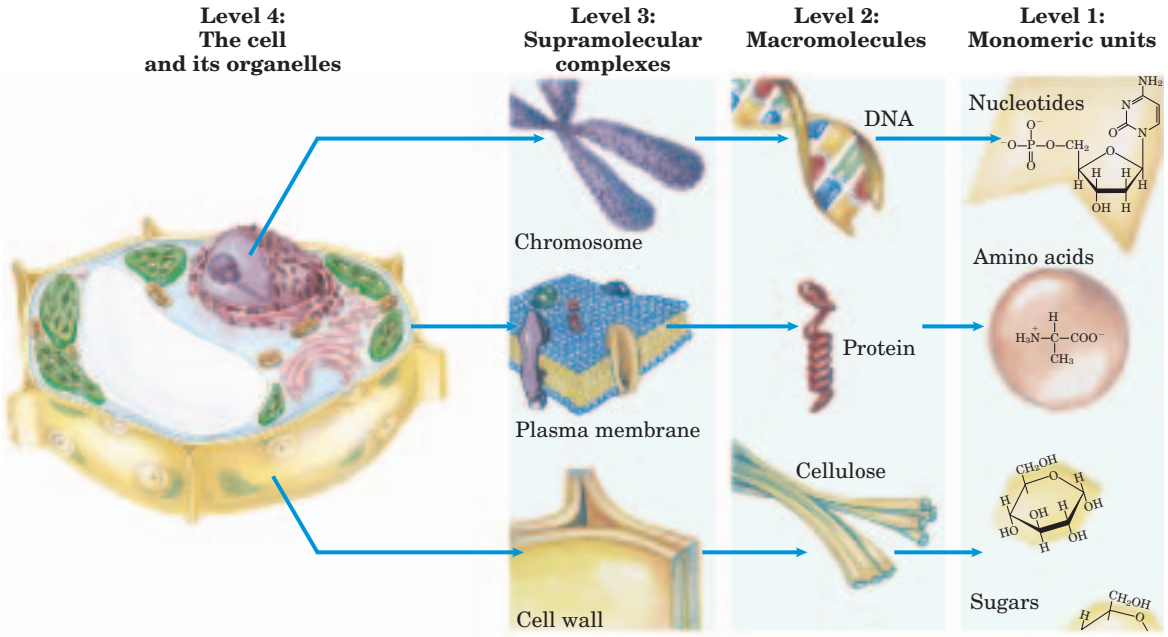


FIGURE 1–11 Structural hierarchy in the molecular organization of cells. In this plant cell, the nucleus is an organelle containing several types of supramolecular complexes, including chromosomes. Chromosomes consist of macromolecules of DNA and many different proteins. Each type of macromolecule is made up of simple subunits—DNA of nucleotides (deoxyribonucleotides), for example.

with the light microscope. Figure 1–11 illustrates the structural hierarchy in cellular organization.

The monomeric subunits in proteins, nucleic acids, and polysaccharides are joined by covalent bonds. In supramolecular complexes, however, macromolecules are held together by noncovalent interactions—much weaker, individually, than covalent bonds. Among these noncovalent interactions are hydrogen bonds (between polar groups), ionic interactions (between charged groups), hydrophobic interactions (among nonpolar groups in aqueous solution), and van der Waals interactions—all of which have energies substantially smaller than those of covalent bonds (Table 1–1). The nature of these noncovalent interactions is described in Chapter 2. The large numbers of weak interactions between macromolecules in supramolecular complexes stabilize these assemblies, producing their unique structures.

In Vitro Studies May Overlook Important Interactions among Molecules

One approach to understanding a biological process is to study purified molecules in vitro ("in glass"—in the test tube), without interference from other molecules present in the intact cell—that is, in vivo ("in the living"). Although this approach has been remarkably revealing, we must keep in mind that the inside of a cell is quite different from the inside of a test tube. The "interfering" components eliminated by purification may be critical to the biological function or regulation of the molecule purified. For example, in vitro studies of pure enzymes are commonly done at very low enzyme concentrations in thoroughly stirred aqueous solutions. In the cell, an enzyme is dissolved or suspended in a gel-like cytosol with thousands of other proteins, some of which bind to that enzyme and influence its activity.

TABLE 1–1 Strengths of Bonds Common in Biomolecules

Type of bond	*Bond dissociation energy* (kJ/mol)*	*Type of bond*	*Bond dissociation energy (kJ/mol)*
Single bonds		**Double bonds**	
O—H	470	C═O	712
H—H	435	C═N	615
P—O	419	C═C	611
C—H	414	P═O	502
N—H	389		
C—O	352	**Triple bonds**	
C—C	348	C≡C	816
S—H	339	N≡N	930
C—N	293		
C—S	260		
N—O	222		
S—S	214		

*The greater the energy required for bond dissociation (breakage), the stronger the bond.

Some enzymes are parts of multienzyme complexes in which reactants are channeled from one enzyme to another without ever entering the bulk solvent. Diffusion is hindered in the gel-like cytosol, and the cytosolic composition varies in different regions of the cell. In short, a given molecule may function quite differently in the cell than in vitro. A central challenge of biochemistry is to understand the influences of cellular organization and macromolecular associations on the function of individual enzymes and other biomolecules—to understand function in vivo as well as in vitro.

SUMMARY 1.1 Cellular Foundations

- All cells are bounded by a plasma membrane; have a cytosol containing metabolites, coenzymes, inorganic ions, and enzymes; and have a set of genes contained within a nucleoid (prokaryotes) or nucleus (eukaryotes).
- Phototrophs use sunlight to do work; chemotrophs oxidize fuels, passing electrons to good electron acceptors: inorganic compounds, organic compounds, or molecular oxygen.
- Bacterial cells contain cytosol, a nucleoid, and plasmids. Eukaryotic cells have a nucleus and are multicompartmented, segregating certain processes in specific organelles, which can be separated and studied in isolation.
- Cytoskeletal proteins assemble into long filaments that give cells shape and rigidity and serve as rails along which cellular organelles move throughout the cell.
- Supramolecular complexes are held together by noncovalent interactions and form a hierarchy of structures, some visible with the light microscope. When individual molecules are removed from these complexes to be studied in vitro, interactions important in the living cell may be lost.

1.2 Chemical Foundations

Biochemistry aims to explain biological form and function in chemical terms. As we noted earlier, one of the most fruitful approaches to understanding biological phenomena has been to purify an individual chemical component, such as a protein, from a living organism and to characterize its structural and chemical characteristics. By the late eighteenth century, chemists had concluded that the composition of living matter is strikingly different from that of the inanimate world. Antoine Lavoisier (1743–1794) noted the relative chemical simplicity of the "mineral world" and contrasted it with the complexity of the "plant and animal worlds"; the latter, he knew, were composed of compounds rich in the elements carbon, oxygen, nitrogen, and phosphorus.

During the first half of the twentieth century, parallel biochemical investigations of glucose breakdown in yeast and in animal muscle cells revealed remarkable chemical similarities in these two apparently very different cell types; the breakdown of glucose in yeast and muscle cells involved the same ten chemical intermediates. Subsequent studies of many other biochemical processes in many different organisms have confirmed the generality of this observation, neatly summarized by Jacques Monod: "What is true of *E. coli* is true of the elephant." The current understanding that all organisms share a common evolutionary origin is based in part on this observed universality of chemical intermediates and transformations.

Only about 30 of the more than 90 naturally occurring chemical elements are essential to organisms. Most of the elements in living matter have relatively low atomic numbers; only five have atomic numbers above that of selenium, 34 (Fig. 1–12). The four most abundant elements in living organisms, in terms of percentage of total number of atoms, are hydrogen, oxygen, nitrogen, and carbon, which together make up more than 99% of the mass of most cells. They are the lightest elements capable of forming one, two, three, and four bonds, respectively; in general, the lightest elements

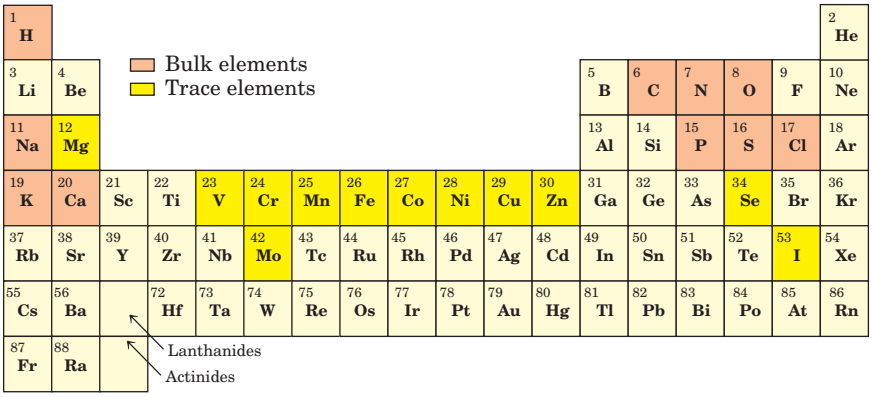


FIGURE 1–12 Elements essential to animal life and health. Bulk elements (shaded orange) are structural components of cells and tissues and are required in the diet in gram quantities daily. For trace elements (shaded bright yellow), the requirements are much smaller: for humans, a few milligrams per day of Fe, Cu, and Zn, even less of the others. The elemental requirements for plants and microorganisms are similar to those shown here; the ways in which they acquire these elements vary.

form the strongest bonds. The trace elements (Fig. 1–12) represent a miniscule fraction of the weight of the human body, but all are essential to life, usually because they are essential to the function of specific proteins, including enzymes. The oxygen-transporting capacity of the hemoglobin molecule, for example, is absolutely dependent on four iron ions that make up only 0.3% of its mass.

Biomolecules Are Compounds of Carbon with a Variety of Functional Groups

The chemistry of living organisms is organized around carbon, which accounts for more than half the dry weight of cells. Carbon can form single bonds with hydrogen atoms, and both single and double bonds with oxygen and nitrogen atoms (Fig. 1–13). Of greatest significance in biology is the ability of carbon atoms to form very stable carbon–carbon single bonds. Each carbon atom can form single bonds with up to four other carbon atoms. Two carbon atoms also can share two (or three) electron pairs, thus forming double (or triple) bonds.

The four single bonds that can be formed by a carbon atom are arranged tetrahedrally, with an angle of

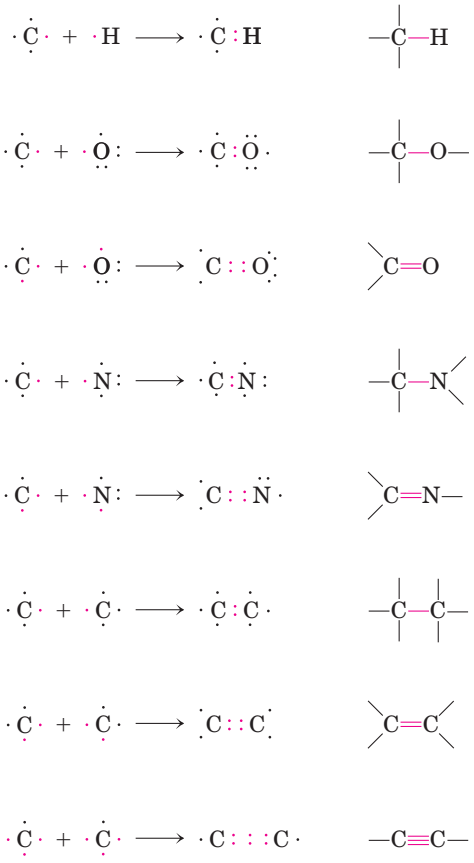


FIGURE 1–13 **Versatility of carbon bonding.** Carbon can form covalent single, double, and triple bonds (in red), particularly with other carbon atoms. Triple bonds are rare in biomolecules.

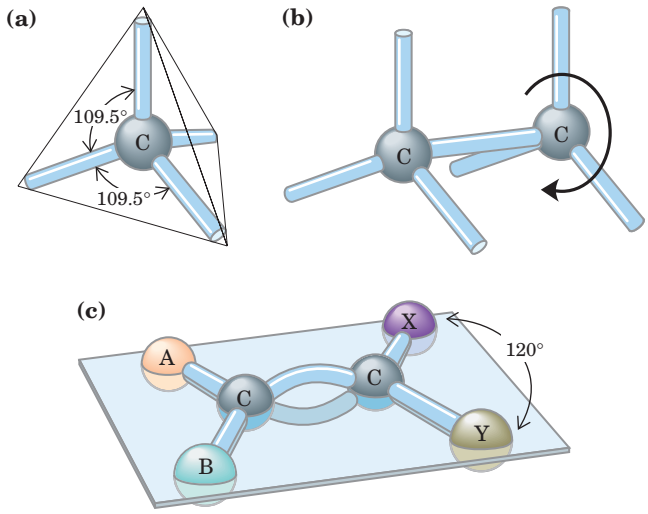


FIGURE 1–14 **Geometry of carbon bonding.** **(a)** Carbon atoms have a characteristic tetrahedral arrangement of their four single bonds. **(b)** Carbon–carbon single bonds have freedom of rotation, as shown for the compound ethane (CH_3—CH_3). **(c)** Double bonds are shorter and do not allow free rotation. The two doubly bonded carbons and the atoms designated A, B, X, and Y all lie in the same rigid plane.

about 109.5° between any two bonds (Fig. 1–14) and an average length of 0.154 nm. There is free rotation around each single bond, unless very large or highly charged groups are attached to both carbon atoms, in which case rotation may be restricted. A double bond is shorter (about 0.134 nm) and rigid and allows little rotation about its axis.

Covalently linked carbon atoms in biomolecules can form linear chains, branched chains, and cyclic structures. To these carbon skeletons are added groups of other atoms, called **functional groups,** which confer specific chemical properties on the molecule. It seems likely that the bonding versatility of carbon was a major factor in the selection of carbon compounds for the molecular machinery of cells during the origin and evolution of living organisms. No other chemical element can form molecules of such widely different sizes and shapes or with such a variety of functional groups.

Most biomolecules can be regarded as derivatives of hydrocarbons, with hydrogen atoms replaced by a variety of functional groups to yield different families of organic compounds. Typical of these are alcohols, which have one or more hydroxyl groups; amines, with amino groups; aldehydes and ketones, with carbonyl groups; and carboxylic acids, with carboxyl groups (Fig. 1–15). Many biomolecules are polyfunctional, containing two or more different kinds of functional groups (Fig. 1–16), each with its own chemical characteristics and reactions. The chemical "personality" of a compound is determined by the chemistry of its functional groups and their disposition in three-dimensional space.

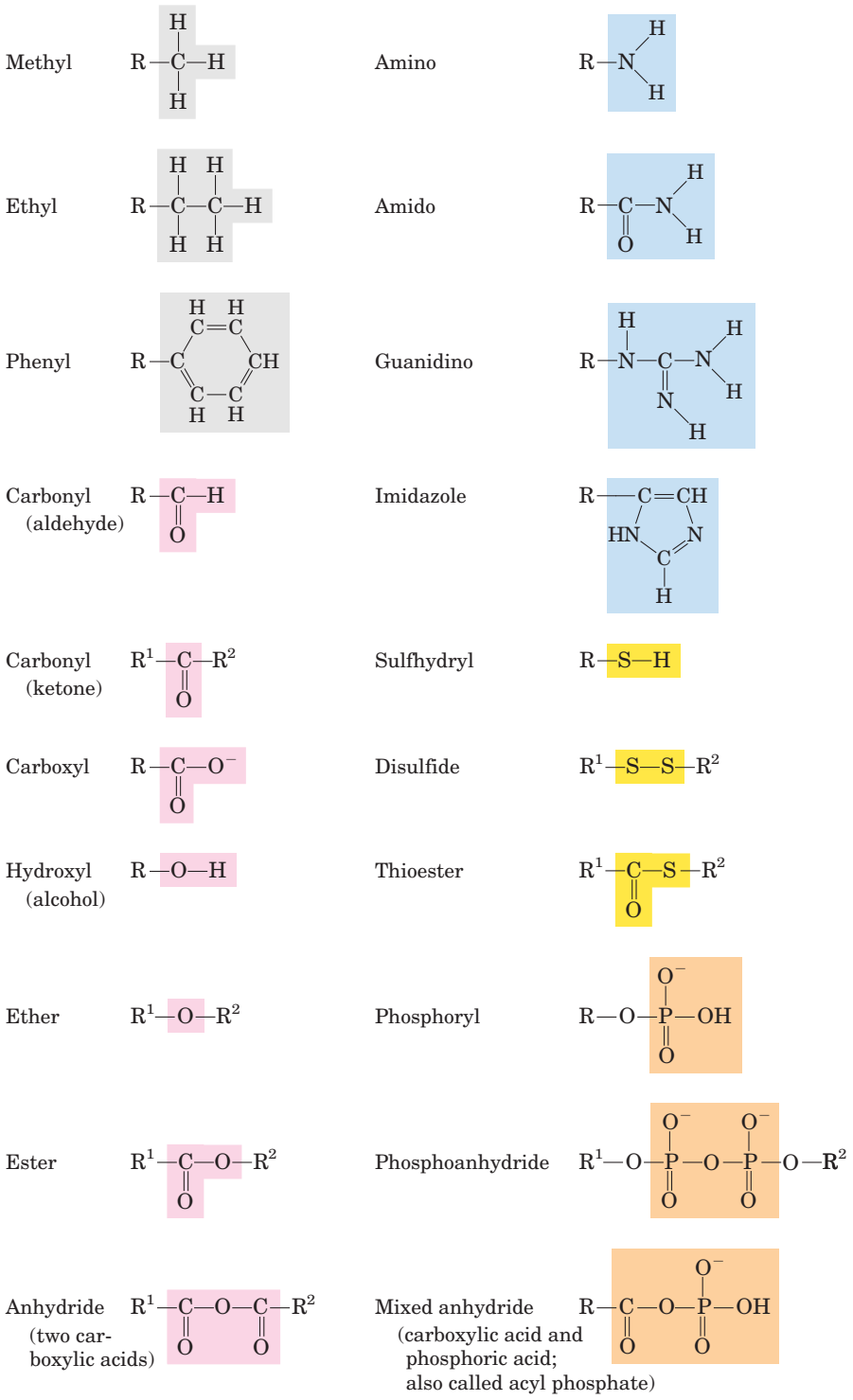


FIGURE 1–15 Some common functional groups of biomolecules. In this figure and throughout the book, we use R to represent "any substituent." It may be as simple as a hydrogen atom, but typically it is a carbon-containing moiety. When two or more substituents are shown in a molecule, we designate them R^1, R^2, and so forth.

Cells Contain a Universal Set of Small Molecules

Dissolved in the aqueous phase (cytosol) of all cells is a collection of 100 to 200 different small organic molecules (M_r ~100 to ~500), the central metabolites in the major pathways occurring in nearly every cell—the metabolites and pathways that have been conserved throughout the course of evolution. (See Box 1–1 for an explanation of the various ways of referring to molecular weight.) This collection of molecules includes the common amino acids, nucleotides, sugars and their phosphorylated derivatives, and a number of mono-, di-, and tricarboxylic acids. The molecules are polar or charged, water soluble, and present in micromolar to millimolar concentrations. They are trapped within the cell because the plasma membrane is impermeable to them—although specific membrane transporters can catalyze the movement of some molecules into and out

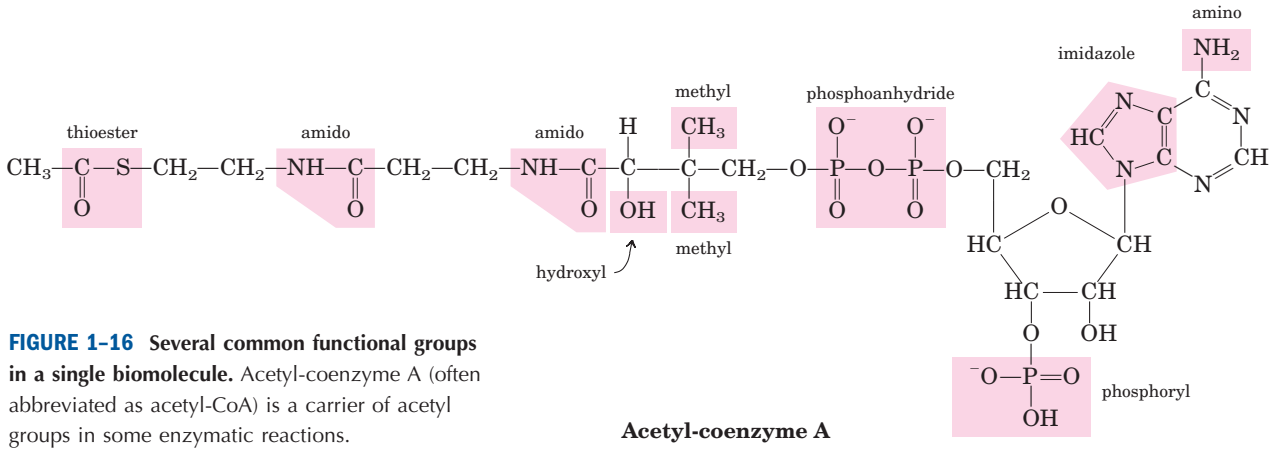


FIGURE 1–16 Several common functional groups in a single biomolecule. Acetyl-coenzyme A (often abbreviated as acetyl-CoA) is a carrier of acetyl groups in some enzymatic reactions.

of the cell or between compartments in eukaryotic cells. The universal occurrence of the same set of compounds in living cells is a manifestation of the universality of metabolic design, reflecting the evolutionary conservation of metabolic pathways that developed in the earliest cells.

There are other small biomolecules, specific to certain types of cells or organisms. For example, vascular plants contain, in addition to the universal set, small molecules called **secondary metabolites,** which play a role specific to plant life. These metabolites include compounds that give plants their characteristic scents, and compounds such as morphine, quinine, nicotine, and caffeine that are valued for their physiological effects on humans but used for other purposes by plants. The entire collection of small molecules in a given cell has been called that cell's **metabolome,** in parallel with the term "genome" (defined earlier and expanded on in Section 1.4). If we knew the composition of a cell's metabolome, we could predict which enzymes and metabolic pathways were active in that cell.

Macromolecules Are the Major Constituents of Cells

Many biological molecules are macromolecules, polymers of high molecular weight assembled from relatively simple precursors. Proteins, nucleic acids, and polysaccharides are produced by the polymerization of relatively small compounds with molecular weights of 500 or less. The number of polymerized units can range from tens to millions. Synthesis of macromolecules is a major energy-consuming activity of cells. Macromolecules themselves may be further assembled into supramolecular complexes, forming functional units such as ribosomes. Table 1–2 shows the major classes of biomolecules in the bacterium *E. coli.*

BOX 1–1 WORKING IN BIOCHEMISTRY

Molecular Weight, Molecular Mass, and Their Correct Units

There are two common (and equivalent) ways to describe molecular mass; both are used in this text. The first is *molecular weight,* or *relative molecular mass,* denoted M_r. The molecular weight of a substance is defined as the ratio of the mass of a molecule of that substance to one-twelfth the mass of carbon-12 (^{12}C). Since M_r is a ratio, it is dimensionless—it has no associated units. The second is *molecular mass,* denoted m. This is simply the mass of one molecule, or the molar mass divided by Avogadro's number. The molecular mass, m, is expressed in daltons (abbreviated Da). One dalton is equivalent to one-twelfth the mass of carbon-12; a kilodalton (kDa) is 1,000 daltons; a megadalton (MDa) is 1 million daltons.

Consider, for example, a molecule with a mass 1,000 times that of water. We can say of this molecule either $M_r = 18{,}000$ or $m = 18{,}000$ daltons. We can also describe it as an "18 kDa molecule." However, the expression $M_r = 18{,}000$ daltons is incorrect.

Another convenient unit for describing the mass of a single atom or molecule is the atomic mass unit (formerly amu, now commonly denoted u). One atomic mass unit (1 u) is defined as one-twelfth the mass of an atom of carbon-12. Since the experimentally measured mass of an atom of carbon-12 is 1.9926×10^{-23} g, 1 u = 1.6606×10^{-24} g. The atomic mass unit is convenient for describing the mass of a peak observed by mass spectrometry (see Box 3–2).

TABLE 1–2 Molecular Components of an *E. coli* Cell

	Percentage of total weight of cell	*Approximate number of different molecular species*
Water	70	1
Proteins	15	3,000
Nucleic acids		
DNA	1	1
RNA	6	>3,000
Polysaccharides	3	5
Lipids	2	20
Monomeric subunits and intermediates	2	500
Inorganic ions	1	20

Proteins, long polymers of amino acids, constitute the largest fraction (besides water) of cells. Some proteins have catalytic activity and function as enzymes; others serve as structural elements, signal receptors, or transporters that carry specific substances into or out of cells. Proteins are perhaps the most versatile of all biomolecules. The **nucleic acids,** DNA and RNA, are polymers of nucleotides. They store and transmit genetic information, and some RNA molecules have structural and catalytic roles in supramolecular complexes. The **polysaccharides,** polymers of simple sugars such as glucose, have two major functions: as energy-yielding fuel stores and as extracellular structural elements with specific binding sites for particular proteins. Shorter polymers of sugars (oligosaccharides) attached to proteins or lipids at the cell surface serve as specific cellular signals. The **lipids,** greasy or oily hydrocarbon derivatives, serve as structural components of membranes, energy-rich fuel stores, pigments, and intracellular signals. In proteins, nucleotides, polysaccharides, and lipids, the number of monomeric subunits is very large: molecular weights in the range of 5,000 to more than 1 million for proteins, up to several billion for nucleic acids, and in the millions for polysaccharides such as starch. Individual lipid molecules are much smaller (M_r 750 to 1,500) and are not classified as macromolecules. However, large numbers of lipid molecules can associate noncovalently into very large structures. Cellular membranes are built of enormous noncovalent aggregates of lipid and protein molecules.

Proteins and nucleic acids are **informational macromolecules:** each protein and each nucleic acid has a characteristic information-rich subunit sequence. Some oligosaccharides, with six or more different sugars connected in branched chains, also carry information; on the outer surface of cells they serve as highly specific points of recognition in many cellular processes (as described in Chapter 7).

Three-Dimensional Structure Is Described by Configuration and Conformation

The covalent bonds and functional groups of a biomolecule are, of course, central to its function, but so also is the arrangement of the molecule's constituent atoms in three-dimensional space—its stereochemistry. A carbon-containing compound commonly exists as **stereoisomers,** molecules with the same chemical bonds but different stereochemistry—that is, different **configuration,** the fixed spatial arrangement of atoms. Interactions between biomolecules are invariably **stereospecific,** requiring specific stereochemistry in the interacting molecules.

Figure 1–17 shows three ways to illustrate the stereochemical structures of simple molecules. The perspective diagram specifies stereochemistry unambiguously, but bond angles and center-to-center bond lengths are better represented with ball-and-stick models. In space-

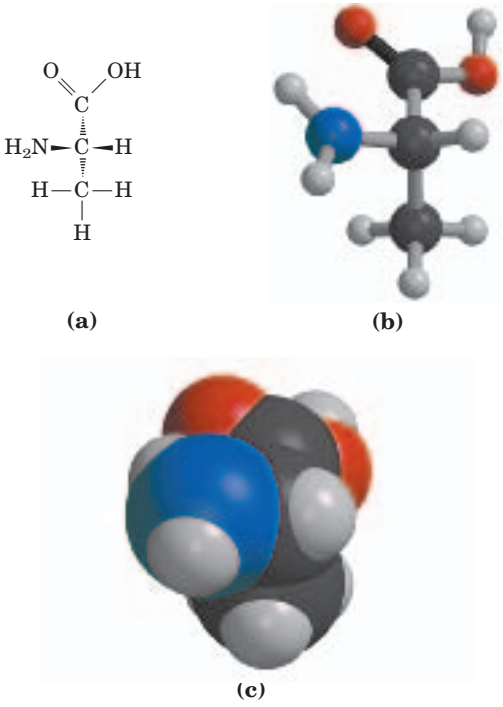
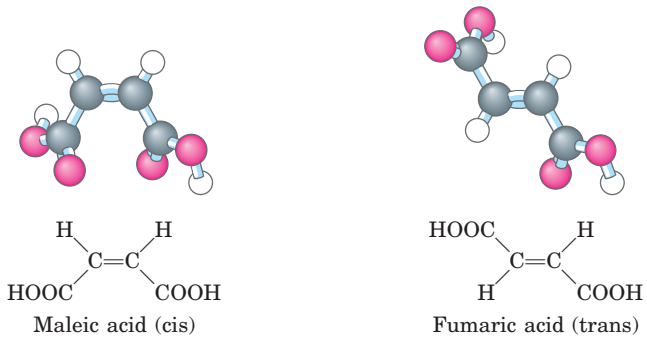
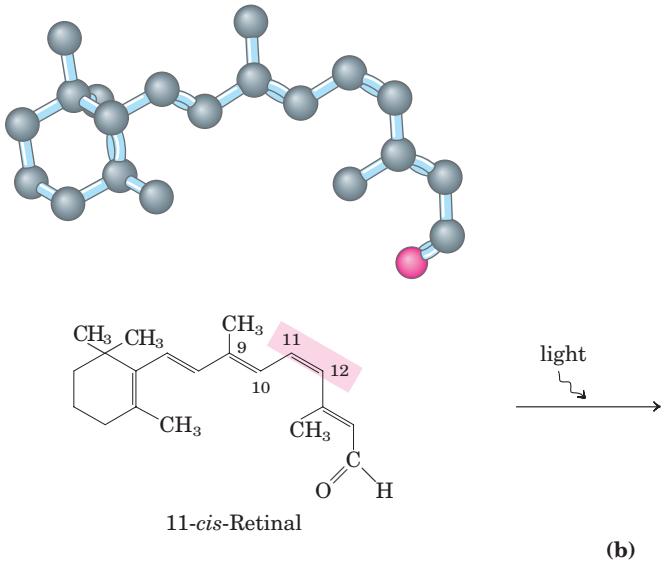


FIGURE 1–17 Representations of molecules. Three ways to represent the structure of the amino acid alanine. **(a)** Structural formula in perspective form: a solid wedge (—) represents a bond in which the atom at the wide end projects out of the plane of the paper, toward the reader; a dashed wedge (┈) represents a bond extending behind the plane of the paper. **(b)** Ball-and-stick model, showing relative bond lengths and the bond angles. **(c)** Space-filling model, in which each atom is shown with its correct relative van der Waals radius.



(a)



(b)

FIGURE 1–18 Configurations of geometric isomers. (a) Isomers such as maleic acid and fumaric acid cannot be interconverted without breaking covalent bonds, which requires the input of much energy. **(b)** In the vertebrate retina, the initial event in light detection is the absorption of visible light by 11-*cis*-retinal. The energy of the absorbed light (about 250 kJ/mol) converts 11-*cis*-retinal to all-*trans*-retinal, triggering electrical changes in the retinal cell that lead to a nerve impulse. (Note that the hydrogen atoms are omitted from the ball-and-stick models for the retinals.)

filling models, the radius of each atom is proportional to its van der Waals radius, and the contours of the model define the space occupied by the molecule (the volume of space from which atoms of other molecules are excluded).

Configuration is conferred by the presence of either (1) double bonds, around which there is no freedom of rotation, or (2) chiral centers, around which substituent groups are arranged in a specific sequence. The identifying characteristic of configurational isomers is that they cannot be interconverted without temporarily breaking one or more covalent bonds. Figure 1–18a shows the configurations of maleic acid and its isomer, fumaric acid. These compounds are **geometric,** or **cis-trans, isomers;** they differ in the arrangement of their substituent groups with respect to the nonrotating double bond (Latin *cis,* "on this side"—groups on the same side of the double bond; *trans,* "across"—groups on opposite sides). Maleic acid is the cis isomer and fumaric acid the trans isomer; each is a well-defined compound that can be separated from the other, and each has its own unique chemical properties. A binding site (on an enzyme, for example) that is complementary to one of these molecules would not be a suitable binding site for the other, which explains why the two compounds have distinct biological roles despite their similar chemistry.

In the second type of configurational isomer, four different substituents bonded to a tetrahedral carbon atom may be arranged two different ways in space—that is, have two configurations (Fig. 1–19)—yielding two stereoisomers with similar or identical chemical properties but differing in certain physical and biological properties. A carbon atom with four different substituents is said to be asymmetric, and asymmetric carbons are called **chiral centers** (Greek *chiros,* "hand"; some stereoisomers are related structurally as the right hand is to the left). A molecule with only one chiral carbon can have two stereoisomers; when two or more (n) chiral carbons are present, there can be 2^n stereoisomers. Some stereoisomers are mirror images of each other; they are called **enantiomers** (Fig. 1–19). Pairs of stereoisomers that are not mirror images of each other are called **diastereomers** (Fig. 1–20).

As Louis Pasteur first observed (Box 1–2), enantiomers have nearly identical chemical properties but differ in a characteristic physical property, their interaction with plane-polarized light. In separate solutions, two enantiomers rotate the plane of plane-polarized light in opposite directions, but an equimolar solution of the two enantiomers (a **racemic mixture**) shows no optical rotation. Compounds without chiral centers do not rotate the plane of plane-polarized light.

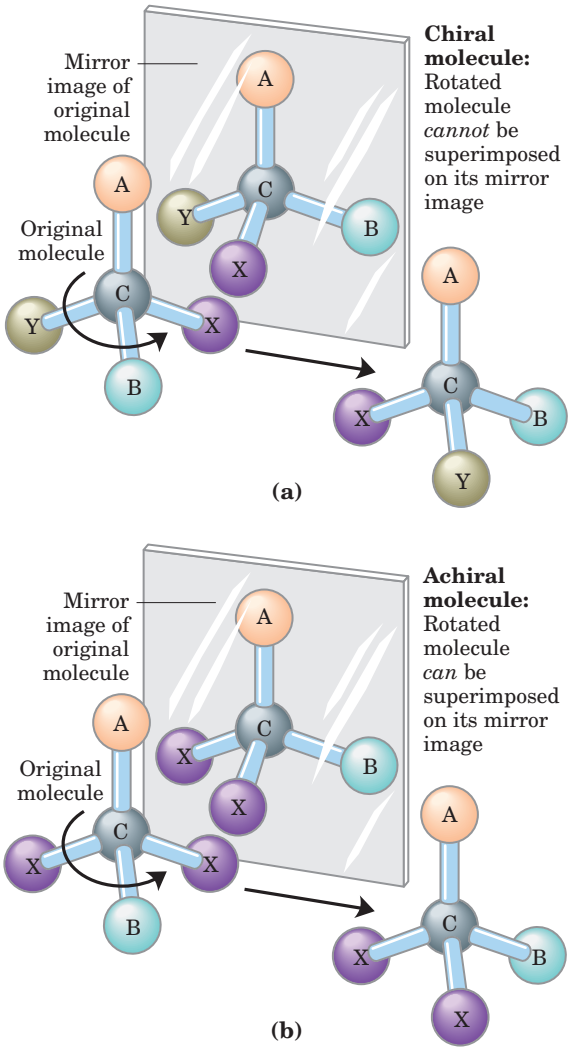


FIGURE 1–19 Molecular asymmetry: chiral and achiral molecules. **(a)** When a carbon atom has four different substituent groups (A, B, X, Y), they can be arranged in two ways that represent nonsuperimposable mirror images of each other (enantiomers). This asymmetric carbon atom is called a chiral atom or chiral center. **(b)** When a tetrahedral carbon has only three dissimilar groups (i.e., the same group occurs twice), only one configuration is possible and the molecule is symmetric, or achiral. In this case the molecule is superimposable on its mirror image: the molecule on the left can be rotated counterclockwise (when looking down the vertical bond from A to C) to create the molecule in the mirror.

Given the importance of stereochemistry in reactions between biomolecules (see below), biochemists must name and represent the structure of each biomolecule so that its stereochemistry is unambiguous. For compounds with more than one chiral center, the most useful system of nomenclature is the RS system. In this system, each group attached to a chiral carbon is assigned a *priority.* The priorities of some common substituents are

$$-OCH_2 > -OH > -NH_2 > -COOH > -CHO > -CH_2OH > -CH_3 > -H$$

For naming in the RS system, the chiral atom is viewed with the group of lowest priority (4 in the diagram on the next page) pointing away from the viewer. If the priority of the other three groups (1 to 3) decreases in clockwise order, the configuration is (*R*) (Latin *rectus*, "right"); if in counterclockwise order, the configuration

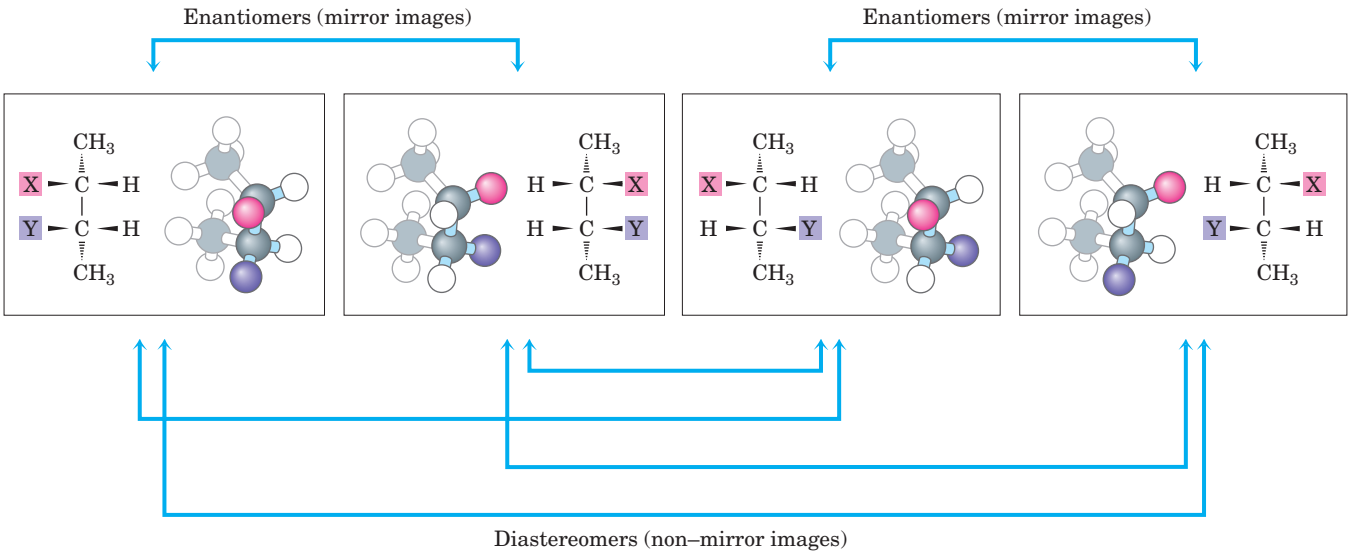


FIGURE 1–20 Two types of stereoisomers. There are four different 2,3-disubstituted butanes ($n = 2$ asymmetric carbons, hence $2^n = 4$ stereoisomers). Each is shown in a box as a perspective formula and a ball-and-stick model, which has been rotated to allow the reader to view all the groups. Some pairs of stereoisomers are mirror images of each other, or enantiomers. Other pairs are not mirror images; these are diastereomers.

BOX 1–2 WORKING IN BIOCHEMISTRY

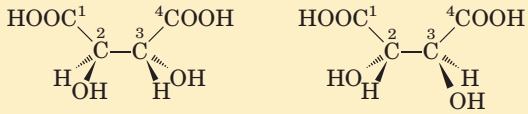
Louis Pasteur and Optical Activity: *In Vino, Veritas*

Louis Pasteur
1822–1895

Louis Pasteur encountered the phenomenon of **optical activity** in 1843, during his investigation of the crystalline sediment that accumulated in wine casks (a form of tartaric acid called paratartaric acid—also called racemic acid, from Latin *racemus,* "bunch of grapes"). He used fine forceps to separate two types of crystals identical in shape but mirror images of each other. Both types proved to have all the chemical properties of tartaric acid, but in solution one type rotated polarized light to the left (levorotatory), the other to the right (dextrorotatory). Pasteur later described the experiment and its interpretation:

> In isomeric bodies, the elements and the proportions in which they are combined are the same, only the arrangement of the atoms is different . . . We know, on the one hand, that the molecular arrangements of the two tartaric acids are asymmetric, and, on the other hand, that these arrangements are absolutely identical, excepting that they exhibit asymmetry in opposite directions. Are the atoms of the dextro acid grouped in the form of a right-handed spiral, or are they placed at the apex of an irregular tetrahedron, or are they disposed according to this or that asymmetric arrangement? We do not know.*

Now we do know. X-ray crystallographic studies in 1951 confirmed that the levorotatory and dextrorotatory forms of tartaric acid are mirror images of each other at the molecular level and established the absolute configuration of each (Fig. 1). The same approach has been used to demonstrate that although the amino acid alanine has two stereoisomeric forms (designated D and L), alanine in proteins exists exclusively in one form (the L isomer; see Chapter 3).

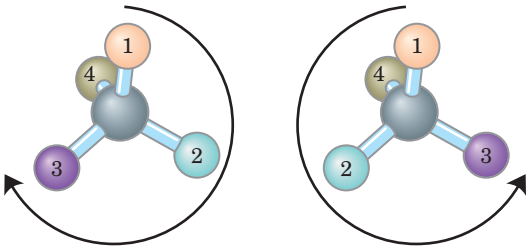


(2*R*,3*R*)-Tartaric acid (dextrorotatory) (2*S*,3*S*)-Tartaric acid (levorotatory)

FIGURE 1 Pasteur separated crystals of two stereoisomers of tartaric acid and showed that solutions of the separated forms rotated polarized light to the same extent but in opposite directions. These dextrorotatory and levorotatory forms were later shown to be the (*R,R*) and (*S,S*) isomers represented here. The RS system of nomenclature is explained in the text.

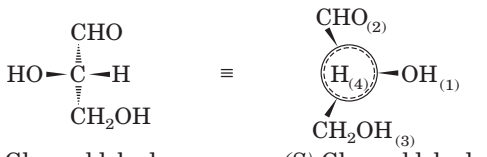
*From Pasteur's lecture to the Société Chimique de Paris in 1883, quoted in DuBos, R. (1976) *Louis Pasteur: Free Lance of Science,* p. 95, Charles Scribner's Sons, New York.

is (*S*) (Latin *sinister,* "left"). In this way each chiral carbon is designated either (*R*) or (*S*), and the inclusion of these designations in the name of the compound provides an unambiguous description of the stereochemistry at each chiral center.



Clockwise (*R*) Counterclockwise (*S*)

Another naming system for stereoisomers, the D and L system, is described in Chapter 3. A molecule with a single chiral center (glyceraldehydes, for example) can be named unambiguously by either system.



L-Glyceraldehyde (*S*)-Glyceraldehyde

Distinct from configuration is molecular **conformation,** the spatial arrangement of substituent groups that, without breaking any bonds, are free to assume different positions in space because of the freedom of rotation about single bonds. In the simple hydrocarbon ethane, for example, there is nearly complete freedom of rotation around the C—C bond. Many different, interconvertible conformations of ethane are possible, depending on the degree of rotation (Fig. 1–21). Two conformations are of special interest: the staggered, which is more stable than all others and thus predominates, and the eclipsed, which is least stable. We cannot isolate either of these conformational forms, because

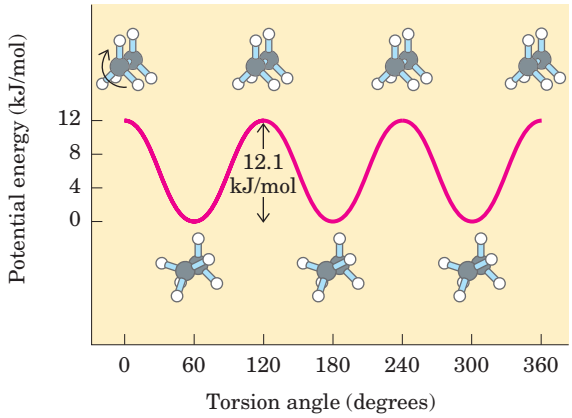


FIGURE 1–21 Conformations. Many conformations of ethane are possible because of freedom of rotation around the C—C bond. In the ball-and-stick model, when the front carbon atom (as viewed by the reader) with its three attached hydrogens is rotated relative to the rear carbon atom, the potential energy of the molecule rises to a maximum in the fully eclipsed conformation (torsion angle 0°, 120°, etc.), then falls to a minimum in the fully staggered conformation (torsion angle 60°, 180°, etc.). Because the energy differences are small enough to allow rapid interconversion of the two forms (millions of times per second), the eclipsed and staggered forms cannot be separately isolated.

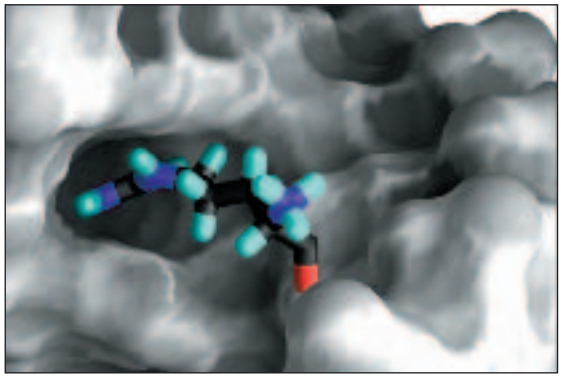


FIGURE 1–22 Complementary fit between a macromolecule and a small molecule. A segment of RNA from the regulatory region TAR of the human immunodeficiency virus genome (gray) with a bound argininamide molecule (colored), representing one residue of a protein that binds to this region. The argininamide fits into a pocket on the RNA surface and is held in this orientation by several noncovalent interactions with the RNA. This representation of the RNA molecule is produced with the computer program GRASP, which can calculate the shape of the outer surface of a macromolecule, defined either by the van der Waals radii of all the atoms in the molecule or by the "solvent exclusion volume," into which a water molecule cannot penetrate.

they are freely interconvertible. However, when one or more of the hydrogen atoms on each carbon is replaced by a functional group that is either very large or electrically charged, freedom of rotation around the C—C bond is hindered. This limits the number of stable conformations of the ethane derivative.

Interactions between Biomolecules Are Stereospecific

Biological interactions between molecules are stereospecific: the "fit" in such interactions must be stereochemically correct. The three-dimensional structure of biomolecules large and small—the combination of configuration and conformation—is of the utmost importance in their biological interactions: reactant with enzyme, hormone with its receptor on a cell surface, antigen with its specific antibody, for example (Fig. 1–22). The study of biomolecular stereochemistry with precise physical methods is an important part of modern research on cell structure and biochemical function.

In living organisms, chiral molecules are usually present in only one of their chiral forms. For example, the amino acids in proteins occur only as their L isomers; glucose occurs only as its D isomer. (The conventions for naming stereoisomers of the amino acids are described in Chapter 3; those for sugars, in Chapter 7; the RS system, described above, is the most useful for some biomolecules.) In contrast, when a compound with an asymmetric carbon atom is chemically synthesized in the laboratory, the reaction usually produces all possible chiral forms: a mixture of the D and L forms, for example. Living cells produce only one chiral form of biomolecules because the enzymes that synthesize them are also chiral.

Stereospecificity, the ability to distinguish between stereoisomers, is a property of enzymes and other proteins and a characteristic feature of the molecular logic of living cells. If the binding site on a protein is complementary to one isomer of a chiral compound, it will not be complementary to the other isomer, for the same reason that a left glove does not fit a right hand. Two striking examples of the ability of biological systems to distinguish stereoisomers are shown in Figure 1–23.

SUMMARY 1.2 Chemical Foundations

- Because of its bonding versatility, carbon can produce a broad array of carbon–carbon skeletons with a variety of functional groups; these groups give biomolecules their biological and chemical personalities.
- A nearly universal set of several hundred small molecules is found in living cells; the interconversions of these molecules in the central metabolic pathways have been conserved in evolution.
- Proteins and nucleic acids are linear polymers of simple monomeric subunits; their sequences contain the information that gives each molecule its three-dimensional structure and its biological functions.

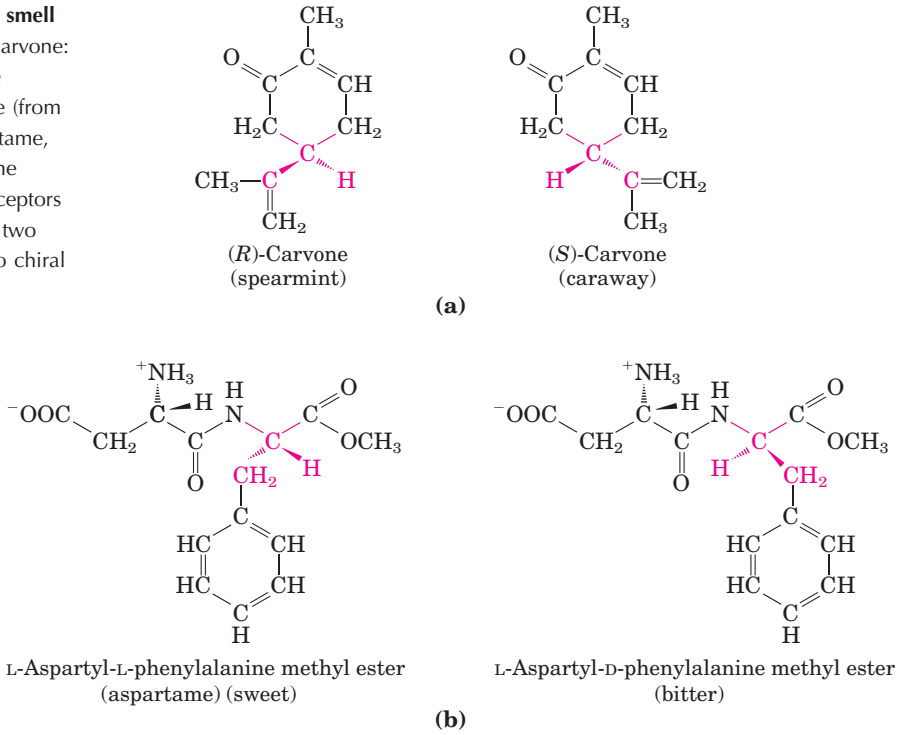


FIGURE 1–23 Stereoisomers distinguishable by smell and taste in humans. (a) Two stereoisomers of carvone: (*R*)-carvone (isolated from spearmint oil) has the characteristic fragrance of spearmint; (*S*)-carvone (from caraway seed oil) smells like caraway. **(b)** Aspartame, the artificial sweetener sold under the trade name NutraSweet, is easily distinguishable by taste receptors from its bitter-tasting stereoisomer, although the two differ only in the configuration at one of the two chiral carbon atoms.

- Molecular configuration can be changed only by breaking covalent bonds. For a carbon atom with four different substituents (a chiral carbon), the substituent groups can be arranged in two different ways, generating stereoisomers with distinct properties. Only one stereoisomer is biologically active. Molecular conformation is the position of atoms in space that can be changed by rotation about single bonds, without breaking covalent bonds.
- Interactions between biological molecules are almost invariably stereospecific: they require a complementary match between the interacting molecules.

1.3 Physical Foundations

Living cells and organisms must perform work to stay alive and to reproduce themselves. The synthetic reactions that occur within cells, like the synthetic processes in any factory, require the input of energy. Energy is also consumed in the motion of a bacterium or an Olympic sprinter, in the flashing of a firefly or the electrical discharge of an eel. And the storage and expression of information require energy, without which structures rich in information inevitably become disordered and meaningless.

In the course of evolution, cells have developed highly efficient mechanisms for coupling the energy obtained from sunlight or fuels to the many energy-consuming processes they must carry out. One goal of biochemistry is to understand, in quantitative and chemical terms, the means by which energy is extracted, channeled, and consumed in living cells. We can consider cellular energy conversions—like all other energy conversions—in the context of the laws of thermodynamics.

Living Organisms Exist in a Dynamic Steady State, Never at Equilibrium with Their Surroundings

The molecules and ions contained within a living organism differ in kind and in concentration from those in the organism's surroundings. A *Paramecium* in a pond, a shark in the ocean, an erythrocyte in the human bloodstream, an apple tree in an orchard—all are different in composition from their surroundings and, once they have reached maturity, all (to a first approximation) maintain a constant composition in the face of constantly changing surroundings.

Although the characteristic composition of an organism changes little through time, the population of molecules within the organism is far from static. Small molecules, macromolecules, and supramolecular complexes are continuously synthesized and then broken down in chemical reactions that involve a constant flux of mass and energy through the system. The hemoglobin molecules carrying oxygen from your lungs to your brain at this moment were synthesized within the past month; by next month they will have been degraded and entirely replaced by new hemoglobin molecules. The glucose you ingested with your most recent meal is now circulating in your bloodstream; before the day is over these particular glucose molecules will have been

converted into something else—carbon dioxide or fat, perhaps—and will have been replaced with a fresh supply of glucose, so that your blood glucose concentration is more or less constant over the whole day. The amounts of hemoglobin and glucose in the blood remain nearly constant because the rate of synthesis or intake of each just balances the rate of its breakdown, consumption, or conversion into some other product. The constancy of concentration is the result of a *dynamic steady state,* a steady state that is far from equilibrium. Maintaining this steady state requires the constant investment of energy; when the cell can no longer generate energy, it dies and begins to decay toward equilibrium with its surroundings. We consider below exactly what is meant by "steady state" and "equilibrium."

Organisms Transform Energy and Matter from Their Surroundings

For chemical reactions occurring in solution, we can define a **system** as all the reactants and products present, the solvent that contains them, and the immediate atmosphere—in short, everything within a defined region of space. The system and its surroundings together constitute the **universe.** If the system exchanges neither matter nor energy with its surroundings, it is said to be **isolated.** If the system exchanges energy but not matter with its surroundings, it is a **closed** system; if it exchanges both energy and matter with its surroundings, it is an **open** system.

A living organism is an open system; it exchanges both matter and energy with its surroundings. Living organisms derive energy from their surroundings in two ways: (1) they take up chemical fuels (such as glucose) from the environment and extract energy by oxidizing them (see Box 1–3, Case 2); or (2) they absorb energy from sunlight.

The first law of thermodynamics, developed from physics and chemistry but fully valid for biological systems as well, describes the principle of the conservation of energy: *in any physical or chemical change, the total amount of energy in the universe remains constant, although the form of the energy may change.* Cells are consummate transducers of energy, capable of interconverting chemical, electromagnetic, mechanical, and osmotic energy with great efficiency (Fig. 1–24).

The Flow of Electrons Provides Energy for Organisms

Nearly all living organisms derive their energy, directly or indirectly, from the radiant energy of sunlight, which arises from thermonuclear fusion reactions carried out in the sun. Photosynthetic cells absorb light energy and use it to drive electrons from water to carbon dioxide, forming energy-rich products such as glucose ($C_6H_{12}O_6$), starch, and sucrose and releasing O_2 into the atmosphere:

$$6CO_2 + 6H_2O \xrightarrow{\text{light}} C_6H_{12}O_6 + 6O_2$$

(light-driven reduction of CO_2)

Nonphotosynthetic cells and organisms obtain the energy they need by oxidizing the energy-rich products of photosynthesis and then passing electrons to atmos-

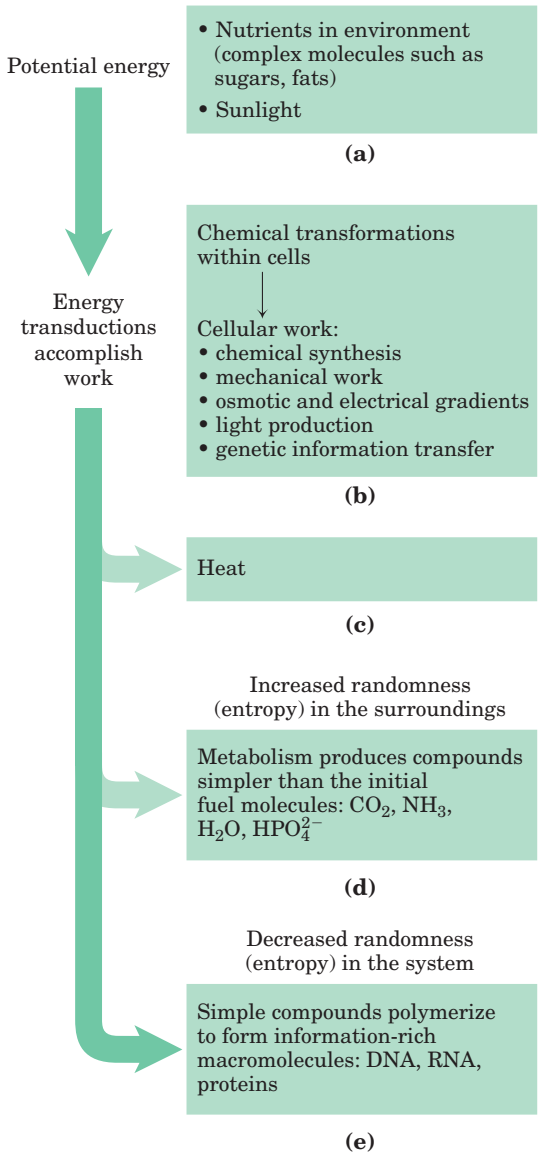


FIGURE 1–24 Some energy interconversion in living organisms. During metabolic energy transductions, the randomness of the system plus surroundings (expressed quantitatively as entropy) increases as the potential energy of complex nutrient molecules decreases. **(a)** Living organisms extract energy from their surroundings; **(b)** convert some of it into useful forms of energy to produce work; **(c)** return some energy to the surroundings as heat; and **(d)** release end-product molecules that are less well organized than the starting fuel, increasing the entropy of the universe. One effect of all these transformations is **(e)** increased order (decreased randomness) in the system in the form of complex macromolecules. We return to a quantitative treatment of entropy in Chapter 13.

pheric O_2 to form water, carbon dioxide, and other end products, which are recycled in the environment:

$$C_6H_{12}O_6 + O_2 \longrightarrow 6CO_2 + 6H_2O + \text{energy}$$
(energy-yielding oxidation of glucose)

Virtually all energy transductions in cells can be traced to this flow of electrons from one molecule to another, in a "downhill" flow from higher to lower electrochemical potential; as such, this is formally analogous to the flow of electrons in a battery-driven electric circuit. All these reactions involving electron flow are **oxidation-reduction reactions:** one reactant is oxidized (loses electrons) as another is reduced (gains electrons).

Creating and Maintaining Order Requires Work and Energy

DNA, RNA, and proteins are informational macromolecules. In addition to using chemical energy to form the covalent bonds between the subunits in these polymers, the cell must invest energy to order the subunits in their correct sequence. It is extremely improbable that amino acids in a mixture would spontaneously condense into a single type of protein, with a unique sequence. This would represent increased order in a population of molecules; but according to the second law of thermodynamics, the tendency in nature is toward ever-greater disorder in the universe: *the total entropy of the universe is continually increasing.* To bring about the synthesis of macromolecules from their monomeric units, free energy must be supplied to the system (in this case, the cell).

The randomness or disorder of the components of a chemical system is expressed as **entropy, *S*** (Box 1–3).

J. Willard Gibbs, 1839–1903

Any change in randomness of the system is expressed as entropy change, ΔS, which by convention has a positive value when randomness increases. J. Willard Gibbs, who developed the theory of energy changes during chemical reactions, showed that the **free-energy content, *G,*** of any closed system can be defined in terms of three quantities: **enthalpy, *H,*** reflecting the number and kinds of bonds; entropy, S; and the absolute temperature, T (in degrees Kelvin). The definition of free energy is $G = H - TS$. When a chemical reaction occurs at constant temperature, the **free-energy change, ΔG,** is determined by the enthalpy change, ΔH, reflecting the kinds and numbers of chemical bonds and noncovalent interactions broken and formed, and the entropy change, ΔS, describing the change in the system's randomness:

$$\Delta G = \Delta H - T\,\Delta S$$

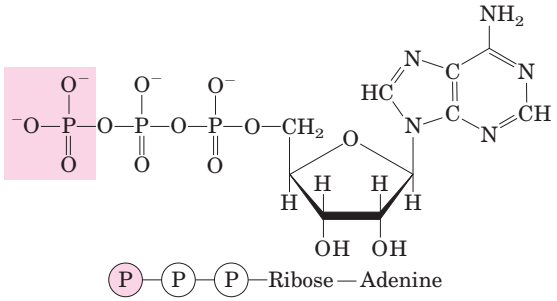


FIGURE 1–25 Adenosine triphosphate (ATP). The removal of the terminal phosphoryl group (shaded pink) of ATP, by breakage of a phosphoanhydride bond, is highly exergonic, and this reaction is coupled to many endergonic reactions in the cell (as in the example in Fig. 1–26b).

A process tends to occur spontaneously only if ΔG is negative. Yet cell function depends largely on molecules, such as proteins and nucleic acids, for which the free energy of formation is positive: the molecules are less stable and more highly ordered than a mixture of their monomeric components. To carry out these thermodynamically unfavorable, energy-requiring **(endergonic)** reactions, cells couple them to other reactions that liberate free energy (**exergonic** reactions), so that the overall process is exergonic: the *sum* of the free-energy changes is negative. The usual source of free energy in coupled biological reactions is the energy released by hydrolysis of phosphoanhydride bonds such as those in adenosine triphosphate (ATP; Fig. 1–25; see also Fig. 1–15). Here, each Ⓟ represents a phosphoryl group:

Amino acids ⟶ polymer ΔG_1 is positive (endergonic)

—Ⓟ—Ⓟ ⟶ —Ⓟ + Ⓟ ΔG_2 is negative (exergonic)

When these reactions are coupled, the sum of ΔG_1 and ΔG_2 is negative—the overall process is exergonic. By this coupling strategy, cells are able to synthesize and maintain the information-rich polymers essential to life.

Energy Coupling Links Reactions in Biology

The central issue in *bioenergetics* (the study of energy transformations in living systems) is the means by which energy from fuel metabolism or light capture is coupled to a cell's energy-requiring reactions. In thinking about energy coupling, it is useful to consider a simple mechanical example, as shown in Figure 1–26a. An object at the top of an inclined plane has a certain amount of potential energy as a result of its elevation. It tends to slide down the plane, losing its potential energy of position as it approaches the ground. When an appropriate string-and-pulley device couples the falling object to another, smaller object, the spontaneous downward motion of the larger can lift the smaller, accomplishing a

BOX 1–3 WORKING IN BIOCHEMISTRY

Entropy: The Advantages of Being Disorganized

The term "entropy," which literally means "a change within," was first used in 1851 by Rudolf Clausius, one of the formulators of the second law of thermodynamics. A rigorous quantitative definition of entropy involves statistical and probability considerations. However, its nature can be illustrated qualitatively by three simple examples, each demonstrating one aspect of entropy. The key descriptors of entropy are *randomness* and *disorder,* manifested in different ways.

Case 1: The Teakettle and the Randomization of Heat

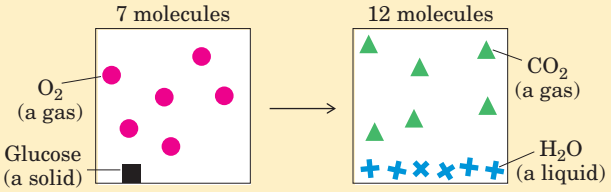
We know that steam generated from boiling water can do useful work. But suppose we turn off the burner under a teakettle full of water at 100 °C (the "system") in the kitchen (the "surroundings") and allow the teakettle to cool. As it cools, no work is done, but heat passes from the teakettle to the surroundings, raising the temperature of the surroundings (the kitchen) by an infinitesimally small amount until complete equilibrium is attained. At this point all parts of the teakettle and the kitchen are at precisely the same temperature. The free energy that was once concentrated in the teakettle of hot water at 100 °C, *potentially* capable of doing work, has disappeared. Its equivalent in heat energy is still present in the teakettle + kitchen (i.e., the "universe") but has become completely randomized throughout. This energy is no longer available to do work because there is no temperature differential within the kitchen. Moreover, the increase in entropy of the kitchen (the surroundings) is irreversible. We know from everyday experience that heat never spontaneously passes back from the kitchen into the teakettle to raise the temperature of the water to 100 °C again.

Case 2: The Oxidation of Glucose

Entropy is a state not only of energy but of matter. Aerobic (heterotrophic) organisms extract free energy from glucose obtained from their surroundings by oxidizing the glucose with O_2, also obtained from the surroundings. The end products of this oxidative metabolism, CO_2 and H_2O, are returned to the surroundings. In this process the surroundings undergo an increase in entropy, whereas the organism itself remains in a steady state and undergoes no change in its internal order. Although some entropy arises from the dissipation of heat, entropy also arises from another kind of disorder, illustrated by the equation for the oxidation of glucose:

$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O$$

We can represent this schematically as



The atoms contained in 1 molecule of glucose plus 6 molecules of oxygen, a total of 7 molecules, are more randomly dispersed by the oxidation reaction and are now present in a total of 12 molecules ($6CO_2 + 6H_2O$).

Whenever a chemical reaction results in an increase in the number of molecules—or when a solid substance is converted into liquid or gaseous products, which allow more freedom of molecular movement than solids—molecular disorder, and thus entropy, increases.

Case 3: Information and Entropy

The following short passage from *Julius Caesar,* Act IV, Scene 3, is spoken by Brutus, when he realizes that he must face Mark Antony's army. It is an information-rich nonrandom arrangement of 125 letters of the English alphabet:

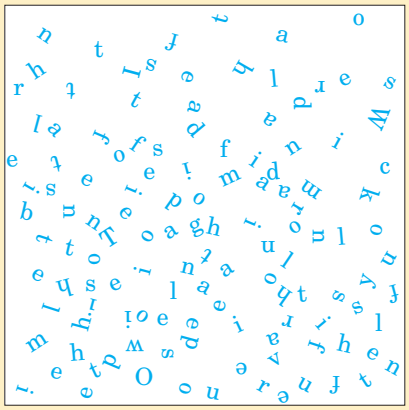
certain amount of work. The amount of energy available to do work is the **free-energy change, ΔG;** this is always somewhat less than the theoretical amount of energy released, because some energy is dissipated as the heat of friction. The greater the elevation of the larger object, the greater the energy released (ΔG) as the object slides downward and the greater the amount of work that can be accomplished.

How does this apply in chemical reactions? In closed systems, chemical reactions proceed spontaneously until **equilibrium** is reached. When a system is at equilibrium, the rate of product formation exactly equals the rate at which product is converted to reactant. Thus there is no net change in the concentration of reactants and products; a *steady state* is achieved. The energy change as the system moves from its initial state to equilibrium, with no changes in temperature or pressure, is given by the free-energy change, ΔG. The magnitude of ΔG depends on the particular chemical reaction and on how far from equilibrium the system is initially. Each compound involved in a chemical reaction contains a certain amount of potential energy, related to the kind and number of its bonds. In reactions that occur spontaneously, the products have less free energy than the re-

There is a tide in the affairs of men,
Which, taken at the flood, leads on to fortune;
Omitted, all the voyage of their life
Is bound in shallows and in miseries.

In addition to what this passage says overtly, it has many hidden meanings. It not only reflects a complex sequence of events in the play, it also echoes the play's ideas on conflict, ambition, and the demands of leadership. Permeated with Shakespeare's understanding of human nature, it is very rich in information.

However, if the 125 letters making up this quotation were allowed to fall into a completely random, chaotic pattern, as shown in the following box, they would have no meaning whatsoever.



In this form the 125 letters contain little or no information, but they are very rich in entropy. Such considerations have led to the conclusion that information is a form of energy; information has been called "negative entropy." In fact, the branch of mathematics called information theory, which is basic to the programming logic of computers, is closely related to thermodynamic theory. Living organisms are highly ordered, nonrandom structures, immensely rich in information and thus entropy-poor.

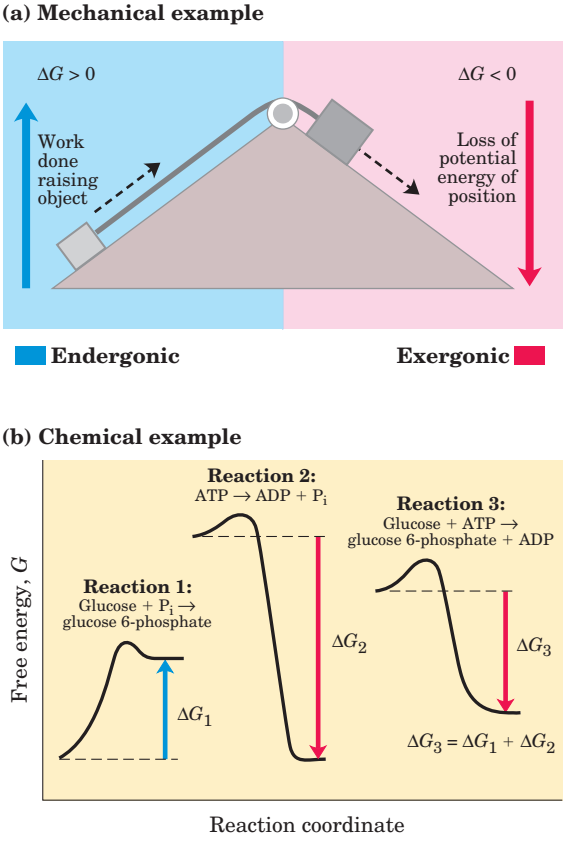


FIGURE 1–26 Energy coupling in mechanical and chemical processes. (a) The downward motion of an object releases potential energy that can do mechanical work. The potential energy made available by spontaneous downward motion, an exergonic process (pink), can be coupled to the endergonic upward movement of another object (blue). **(b)** In reaction 1, the formation of glucose 6-phosphate from glucose and inorganic phosphate (P_i) yields a product of higher energy than the two reactants. For this endergonic reaction, ΔG is positive. In reaction 2, the exergonic breakdown of adenosine triphosphate (ATP) can drive an endergonic reaction when the two reactions are coupled. The exergonic reaction has a large, negative free-energy change (ΔG_2), and the endergonic reaction has a smaller, positive free-energy change (ΔG_1). The third reaction accomplishes the sum of reactions 1 and 2, and the free-energy change, ΔG_3, is the arithmetic sum of ΔG_1 and ΔG_2. Because ΔG_3 is negative, the overall reaction is exergonic and proceeds spontaneously.

actants, thus the reaction releases free energy, which is then available to do work. Such reactions are exergonic; the decline in free energy from reactants to products is expressed as a negative value. Endergonic reactions require an input of energy, and their ΔG values are positive. As in mechanical processes, only part of the energy released in exergonic chemical reactions can be used to accomplish work. In living systems some energy is dissipated as heat or lost to increasing entropy.

In living organisms, as in the mechanical example in Figure 1–26a, an exergonic reaction can be coupled to an endergonic reaction to drive otherwise unfavorable reactions. Figure 1–26b (a type of graph called a reaction coordinate diagram) illustrates this principle for the conversion of glucose to glucose 6-phosphate, the first step in the pathway for oxidation of glucose. The simplest way to produce glucose 6-phosphate would be:

Reaction 1: Glucose + P_i ⟶ glucose 6-phosphate
(endergonic; ΔG_1 is positive)

(P_i is an abbreviation for inorganic phosphate, HPO_4^{2-}. Don't be concerned about the structure of these compounds now; we describe them in detail later in the book.) This reaction does not occur spontaneously; ΔG

is positive. A second, very exergonic reaction can occur in all cells:

Reaction 2: $ATP \longrightarrow ADP + P_i$ (exergonic; ΔG_2 is negative)

These two chemical reactions share a common intermediate, P_i, which is consumed in reaction 1 and produced in reaction 2. The two reactions can therefore be coupled in the form of a third reaction, which we can write as the sum of reactions 1 and 2, with the common intermediate, P_i, omitted from both sides of the equation:

Reaction 3: Glucose + ATP $\longrightarrow$ glucose 6-phosphate + ADP

Because more energy is released in reaction 2 than is consumed in reaction 1, the free-energy change for reaction 3, ΔG_3, is negative, and the synthesis of glucose 6-phosphate can therefore occur by reaction 3.

The coupling of exergonic and endergonic reactions through a shared intermediate is absolutely central to the energy exchanges in living systems. As we shall see, the breakdown of ATP (reaction 2 in Fig. 1–26b) is the exergonic reaction that drives many endergonic processes in cells. In fact, ATP is the major carrier of chemical energy in all cells.

K_{eq} and $\Delta G°$ Are Measures of a Reaction's Tendency to Proceed Spontaneously

The tendency of a chemical reaction to go to completion can be expressed as an equilibrium constant. For the reaction

$$aA + bB \longrightarrow cC + dD$$

the equilibrium constant, K_{eq}, is given by

$$K_{eq} = \frac{[C_{eq}]^c[D_{eq}]^d}{[A_{eq}]^a[B_{eq}]^b}$$

where $[A_{eq}]$ is the concentration of A, $[B_{eq}]$ the concentration of B, and so on, *when the system has reached equilibrium.* A large value of K_{eq} means the reaction tends to proceed until the reactants have been almost completely converted into the products.

Gibbs showed that ΔG for any chemical reaction is a function of the **standard free-energy change, $\Delta G°$**— a constant that is characteristic of each specific reaction—and a term that expresses the initial concentrations of reactants and products:

$$\Delta G = \Delta G° + RT \ln \frac{[C_i]^c[D_i]^d}{[A_i]^a[B_i]^b} \qquad (1–1)$$

where $[A_i]$ is the initial concentration of A, and so forth; R is the gas constant; and T is the absolute temperature.

When a reaction has reached equilibrium, no driving force remains and it can do no work: $\Delta G = 0$. For this special case, $[A_i] = [A_{eq}]$, and so on, for all reactants and products, and

$$\frac{[C_i]^c[D_i]^d}{[A_i]^a[B_i]^b} = \frac{[C_{eq}]^c[D_{eq}]^d}{[A_{eq}]^a[B_{eq}]^b} = K_{eq}$$

Substituting 0 for ΔG and K_{eq} for $[C_i]^c[D_i]^d/[A_i]^a[B_i]^b$ in Equation 1–1, we obtain the relationship

$$\Delta G° = -RT \ln K_{eq}$$

from which we see that $\Delta G°$ is simply a second way (besides K_{eq}) of expressing the driving force on a reaction. Because K_{eq} is experimentally measurable, we have a way of determining $\Delta G°$, the thermodynamic constant characteristic of each reaction.

The units of $\Delta G°$ and ΔG are joules per mole (or calories per mole). When $K_{eq} >> 1$, $\Delta G°$ is large and negative; when $K_{eq} << 1$, $\Delta G°$ is large and positive. From a table of experimentally determined values of either K_{eq} or $\Delta G°$, we can see at a glance which reactions tend to go to completion and which do not.

One caution about the interpretation of $\Delta G°$: *thermodynamic* constants such as this show where the final equilibrium for a reaction lies but tell us nothing about how fast that equilibrium will be achieved. The rates of reactions are governed by the parameters of *kinetics,* a topic we consider in detail in Chapter 6.

Enzymes Promote Sequences of Chemical Reactions

All biological macromolecules are much less thermodynamically stable than their monomeric subunits, yet they are *kinetically stable:* their *uncatalyzed* breakdown occurs so slowly (over years rather than seconds) that, on a time scale that matters for the organism, these molecules are stable. Virtually every chemical reaction in a cell occurs at a significant rate only because of the presence of **enzymes**—biocatalysts that, like all other catalysts, greatly enhance the rate of specific chemical reactions without being consumed in the process.

The path from reactant(s) to product(s) almost invariably involves an energy barrier, called the activation barrier (Fig. 1–27), that must be surmounted for any reaction to proceed. The breaking of existing bonds and formation of new ones generally requires, first, the distortion of the existing bonds, creating a **transition state** of higher free energy than either reactant or product. The highest point in the reaction coordinate diagram represents the transition state, and the difference in energy between the reactant in its ground state and in its transition state is the **activation energy, $\Delta G^‡$.** An enzyme catalyzes a reaction by providing a more comfortable fit for the transition state: a surface that complements the transition state in stereochemistry, polarity, and charge. The binding of enzyme to the transition state is exergonic, and the energy released by this binding reduces the activation energy for the reaction and greatly increases the reaction rate.

A further contribution to catalysis occurs when two or more reactants bind to the enzyme's surface close to each other and with stereospecific orientations that fa-

vor the reaction. This increases by orders of magnitude the probability of productive collisions between reactants. As a result of these factors and several others, discussed in Chapter 6, enzyme-catalyzed reactions commonly proceed at rates greater than 10^{12} times faster than the uncatalyzed reactions.

Cellular catalysts are, with a few exceptions, proteins. (In some cases, RNA molecules have catalytic roles, as discussed in Chapters 26 and 27.) Again with a few exceptions, each enzyme catalyzes a specific reaction, and each reaction in a cell is catalyzed by a different enzyme. Thousands of different enzymes are therefore required by each cell. The multiplicity of enzymes, their specificity (the ability to discriminate between reactants), and their susceptibility to regulation give cells the capacity to lower activation barriers selectively. This selectivity is crucial for the effective regulation of cellular processes. By allowing specific reactions to proceed at significant rates at particular times, enzymes determine how matter and energy are channeled into cellular activities.

The thousands of enzyme-catalyzed chemical reactions in cells are functionally organized into many sequences of consecutive reactions, called **pathways,** in which the product of one reaction becomes the reactant in the next. Some pathways degrade organic nutrients into simple end products in order to extract chemical energy and convert it into a form useful to the cell; together these degradative, free-energy-yielding reactions are designated **catabolism.** Other pathways start with small precursor molecules and convert them to progressively larger and more complex molecules, including proteins and nucleic acids. Such synthetic pathways,

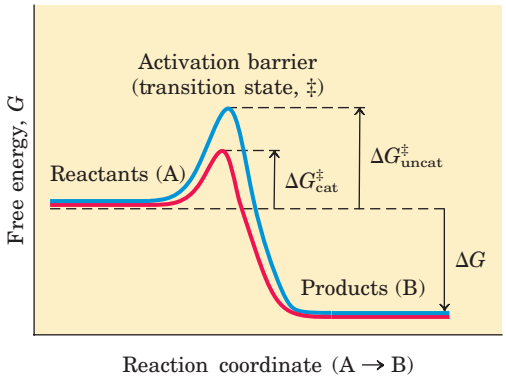


FIGURE 1–27 Energy changes during a chemical reaction. An activation barrier, representing the transition state, must be overcome in the conversion of reactants (A) into products (B), even though the products are more stable than the reactants, as indicated by a large, negative free-energy change (ΔG). The energy required to overcome the activation barrier is the activation energy ($\Delta G^{\ddagger}$). Enzymes catalyze reactions by lowering the activation barrier. They bind the transition-state intermediates tightly, and the binding energy of this interaction effectively reduces the activation energy from $\Delta G^{\ddagger}_{uncat}$ to $\Delta G^{\ddagger}_{cat}$. (Note that activation energy is *not* related to free-energy change, ΔG.)

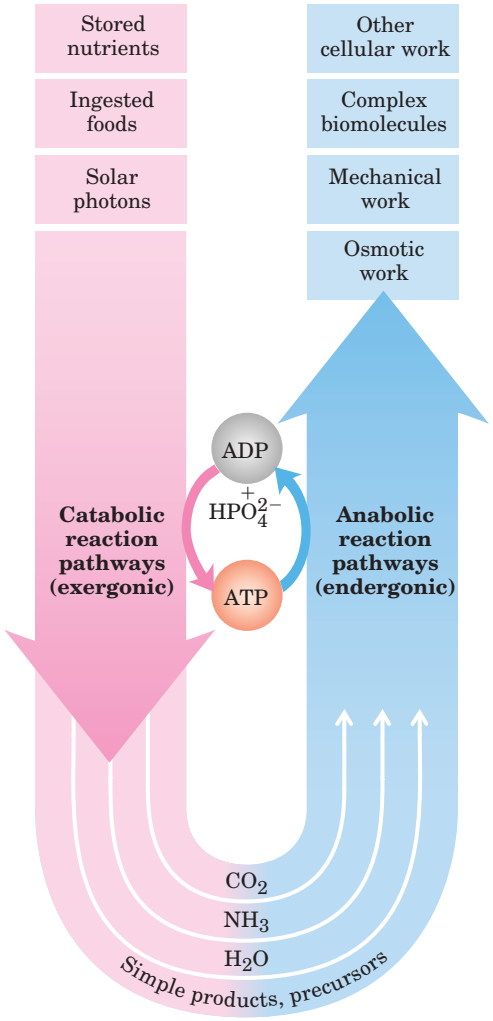


FIGURE 1–28 The central role of ATP in metabolism. ATP is the shared chemical intermediate linking energy-releasing to energy-requiring cell processes. Its role in the cell is analogous to that of money in an economy: it is "earned/produced" in exergonic reactions and "spent/consumed" in endergonic ones.

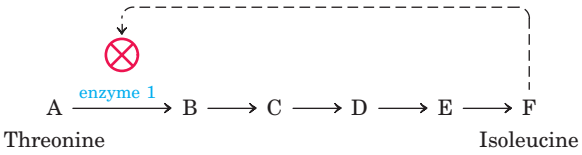
which invariably require the input of energy, are collectively designated **anabolism.** The overall network of enzyme-catalyzed pathways constitutes cellular **metabolism.** ATP is the major connecting link (the shared intermediate) between the catabolic and anabolic components of this network (shown schematically in Fig. 1–28). The pathways of enzyme-catalyzed reactions that act on the main constituents of cells—proteins, fats, sugars, and nucleic acids—are virtually identical in all living organisms.

Metabolism Is Regulated to Achieve Balance and Economy

Not only do living cells simultaneously synthesize thousands of different kinds of carbohydrate, fat, protein, and nucleic acid molecules and their simpler subunits, but they do so in the precise proportions required by

the cell under any given circumstance. For example, during rapid cell growth the precursors of proteins and nucleic acids must be made in large quantities, whereas in nongrowing cells the requirement for these precursors is much reduced. Key enzymes in each metabolic pathway are regulated so that each type of precursor molecule is produced in a quantity appropriate to the current requirements of the cell.

Consider the pathway in *E. coli* that leads to the synthesis of the amino acid isoleucine, a constituent of proteins. The pathway has five steps catalyzed by five different enzymes (A through F represent the intermediates in the pathway):



If a cell begins to produce more isoleucine than is needed for protein synthesis, the unused isoleucine accumulates and the increased concentration inhibits the catalytic activity of the first enzyme in the pathway, immediately slowing the production of isoleucine. Such **feedback inhibition** keeps the production and utilization of each metabolic intermediate in balance.

Although the concept of discrete pathways is an important tool for organizing our understanding of metabolism, it is an oversimplification. There are thousands of metabolic intermediates in a cell, many of which are part of more than one pathway. Metabolism would be better represented as a meshwork of interconnected and interdependent pathways. A change in the concentration of any one metabolite would have an impact on other pathways, starting a ripple effect that would influence the flow of materials through other sectors of the cellular economy. The task of understanding these complex interactions among intermediates and pathways in quantitative terms is daunting, but new approaches, discussed in Chapter 15, have begun to offer important insights into the overall regulation of metabolism.

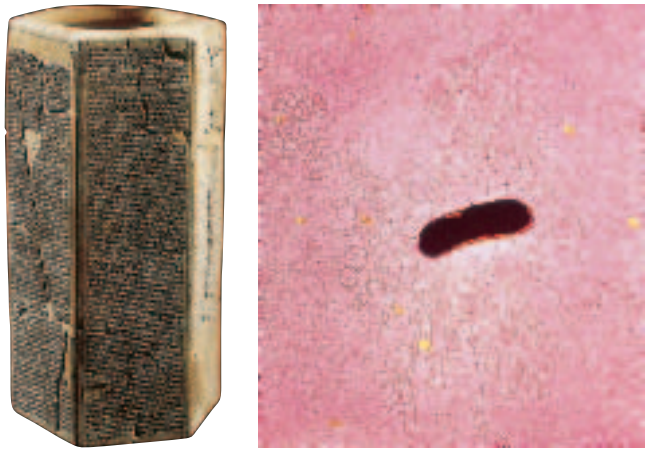
Cells also regulate the synthesis of their own catalysts, the enzymes, in response to increased or diminished need for a metabolic product; this is the substance of Chapter 28. The expression of genes (the translation of information in DNA to active protein in the cell) and synthesis of enzymes are other layers of metabolic control in the cell. All layers must be taken into account when describing the overall control of cellular metabolism. Assembling the complete picture of how the cell regulates itself is a huge job that has only just begun.

SUMMARY 1.3 Physical Foundations

- Living cells are open systems, exchanging matter and energy with their surroundings, extracting and channeling energy to maintain themselves in a dynamic steady state distant from equilibrium. Energy is obtained from sunlight or fuels by converting the energy from electron flow into the chemical bonds of ATP.
- The tendency for a chemical reaction to proceed toward equilibrium can be expressed as the free-energy change, ΔG, which has two components: enthalpy change, ΔH, and entropy change, ΔS. These variables are related by the equation $\Delta G = \Delta H - T\,\Delta S$.
- When ΔG of a reaction is negative, the reaction is exergonic and tends to go toward completion; when ΔG is positive, the reaction is endergonic and tends to go in the reverse direction. When two reactions can be summed to yield a third reaction, the ΔG for this overall reaction is the sum of the ΔGs of the two separate reactions. This provides a way to couple reactions.
- The conversion of ATP to P_i and ADP is highly exergonic (large negative ΔG), and many endergonic cellular reactions are driven by coupling them, through a common intermediate, to this reaction.
- The standard free-energy change for a reaction, $\Delta G°$, is a physical constant that is related to the equilibrium constant by the equation $\Delta G° = -RT \ln K_{eq}$.
- Most exergonic cellular reactions proceed at useful rates only because enzymes are present to catalyze them. Enzymes act in part by stabilizing the transition state, reducing the activation energy, $\Delta G^{‡}$, and increasing the reaction rate by many orders of magnitude. The catalytic activity of enzymes in cells is regulated.
- Metabolism is the sum of many interconnected reaction sequences that interconvert cellular metabolites. Each sequence is regulated so as to provide what the cell needs at a given time and to expend energy only when necessary.

1.4 Genetic Foundations

Perhaps the most remarkable property of living cells and organisms is their ability to reproduce themselves for countless generations with nearly perfect fidelity. This continuity of inherited traits implies constancy, over millions of years, in the structure of the molecules that contain the genetic information. Very few historical records of civilization, even those etched in copper or carved in stone (Fig. 1–29), have survived for a thousand years. But there is good evidence that the genetic instructions in living organisms have remained nearly unchanged over very much longer periods; many bacteria have nearly



(a) (b)

FIGURE 1–29 Two ancient scripts. (a) The Prism of Sennacherib, inscribed in about 700 B.C.E., describes in characters of the Assyrian language some historical events during the reign of King Sennacherib. The Prism contains about 20,000 characters, weighs about 50 kg, and has survived almost intact for about 2,700 years. **(b)** The single DNA molecule of the bacterium *E. coli,* seen leaking out of a disrupted cell, is hundreds of times longer than the cell itself and contains all the encoded information necessary to specify the cell's structure and functions. The bacterial DNA contains about 10 million characters (nucleotides), weighs less than 10^{-10} g, and has undergone only relatively minor changes during the past several million years. (The yellow spots and dark specks in this colorized electron micrograph are artifacts of the preparation.)

the same size, shape, and internal structure and contain the same kinds of precursor molecules and enzymes as bacteria that lived nearly four billion years ago.

Among the seminal discoveries in biology in the twentieth century were the chemical nature and the three-dimensional structure of the genetic material, **deoxyribonucleic acid, DNA.** The sequence of the monomeric subunits, the nucleotides (strictly, deoxyribonucleotides, as discussed below), in this linear polymer encodes the instructions for forming all other cellular components and provides a template for the production of identical DNA molecules to be distributed to progeny when a cell divides. The continued existence of a biological species requires its genetic information to be maintained in a stable form, expressed accurately in the form of gene products, and reproduced with a minimum of errors. Effective storage, expression, and reproduction of the genetic message defines individual species, distinguishes them from one another, and assures their continuity over successive generations.

Genetic Continuity Is Vested in Single DNA Molecules

DNA is a long, thin organic polymer, the rare molecule that is constructed on the atomic scale in one dimension (width) and the human scale in another (length: a molecule of DNA can be many centimeters long). A human sperm or egg, carrying the accumulated hereditary information of billions of years of evolution, transmits this inheritance in the form of DNA molecules, in which the linear sequence of covalently linked nucleotide subunits encodes the genetic message.

Usually when we describe the properties of a chemical species, we describe the average behavior of a very large number of identical molecules. While it is difficult to predict the behavior of any single molecule in a collection of, say, a picomole (about 6×10^{11} molecules) of a compound, the *average* behavior of the molecules is predictable because so many molecules enter into the average. Cellular DNA is a remarkable exception. The DNA that is the entire genetic material of *E. coli* is a *single molecule* containing 4.64 million nucleotide pairs. That single molecule must be replicated perfectly in every detail if an *E. coli* cell is to give rise to identical progeny by cell division; there is no room for averaging in this process! The same is true for all cells. A human sperm brings to the egg that it fertilizes just one molecule of DNA in each of its 23 different chromosomes, to combine with just one DNA molecule in each corresponding chromosome in the egg. The result of this union is very highly predictable: an embryo with all of its 35,000 genes, constructed of 3 billion nucleotide pairs, intact. An amazing chemical feat!

The Structure of DNA Allows for Its Replication and Repair with Near-Perfect Fidelity

The capacity of living cells to preserve their genetic material and to duplicate it for the next generation results from the structural complementarity between the two halves of the DNA molecule (Fig. 1–30). The basic unit of DNA is a linear polymer of four different monomeric subunits, **deoxyribonucleotides,** arranged in a precise linear sequence. It is this linear sequence that encodes the genetic information. Two of these polymeric strands are twisted about each other to form the DNA double helix, in which each deoxyribonucleotide in one strand pairs specifically with a complementary deoxyribonucleotide in the opposite strand. Before a cell divides, the two DNA strands separate and each serves as a template for the synthesis of a new complementary strand, generating two identical double-helical molecules, one for each daughter cell. If one strand is damaged, continuity of information is assured by the information present in the other strand, which acts as a template for repair of the damage.

The Linear Sequence in DNA Encodes Proteins with Three-Dimensional Structures

The information in DNA is encoded in its linear (one-dimensional) sequence of deoxyribonucleotide subunits, but the expression of this information results in

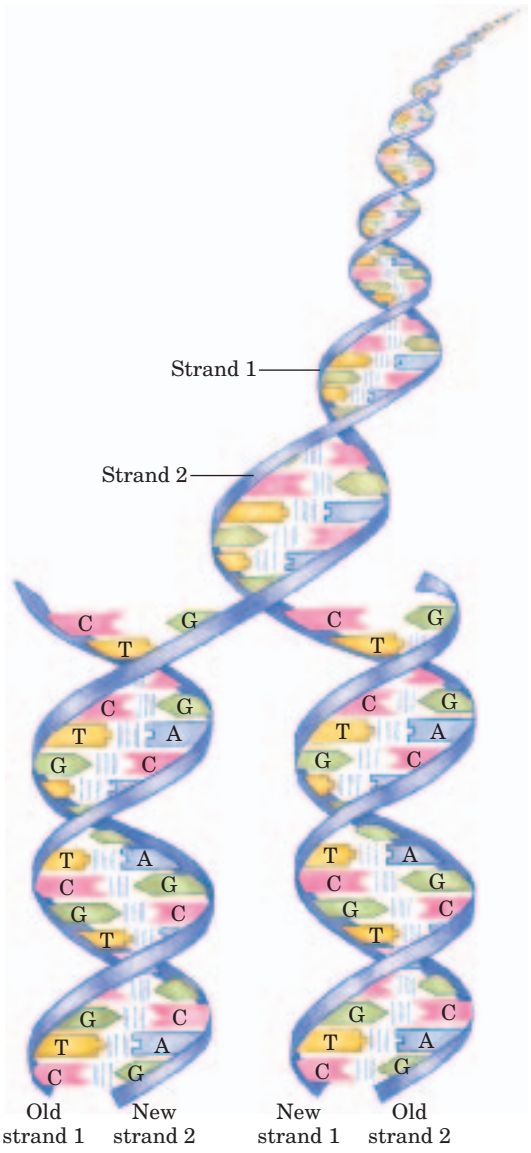


FIGURE 1–30 Complementarity between the two strands of DNA. DNA is a linear polymer of covalently joined deoxyribonucleotides, of four types: deoxyadenylate (A), deoxyguanylate (G), deoxycytidylate (C), and deoxythymidylate (T). Each nucleotide, with its unique three-dimensional structure, can associate very specifically but noncovalently with one other nucleotide in the complementary chain: A always associates with T, and G with C. Thus, in the double-stranded DNA molecule, the entire sequence of nucleotides in one strand is *complementary* to the sequence in the other. The two strands, held together by hydrogen bonds (represented here by vertical blue lines) between each pair of complementary nucleotides, twist about each other to form the DNA double helix. In DNA replication, the two strands separate and two new strands are synthesized, each with a sequence complementary to one of the original strands. The result is two double-helical molecules, each identical to the original DNA.

a three-dimensional cell. This change from one to three dimensions occurs in two phases. A linear sequence of deoxyribonucleotides in DNA codes (through an intermediary, RNA) for the production of a protein with a corresponding linear sequence of amino acids (Fig. 1–31). The protein folds into a particular three-dimensional shape, determined by its amino acid sequence and stabilized primarily by noncovalent interactions. Although the final shape of the folded protein is dictated by its amino acid sequence, the folding process is aided by "molecular chaperones," which catalyze the process by discouraging incorrect folding. The precise three-dimensional structure, or **native conformation,** of the protein is crucial to its function.

Once in its native conformation, a protein may associate noncovalently with other proteins, or with nucleic acids or lipids, to form supramolecular complexes such as chromosomes, ribosomes, and membranes. The individual molecules of these complexes have specific, high-affinity binding sites for each other, and within the cell they spontaneously form functional complexes.

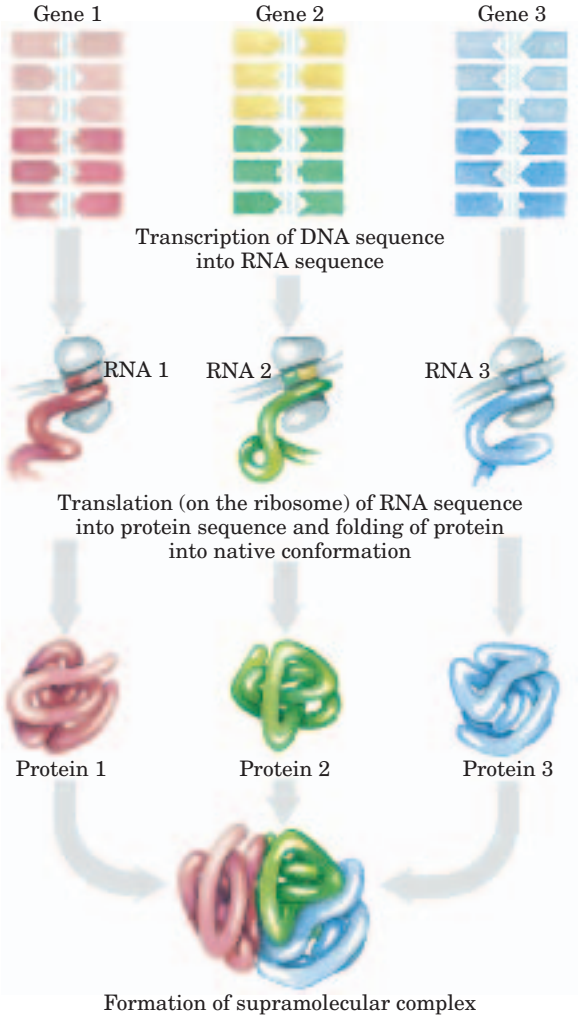


FIGURE 1–31 DNA to RNA to protein. Linear sequences of deoxyribonucleotides in DNA, arranged into units known as genes, are transcribed into ribonucleic acid (RNA) molecules with complementary ribonucleotide sequences. The RNA sequences are then translated into linear protein chains, which fold into their native three-dimensional shapes, often aided by molecular chaperones. Individual proteins commonly associate with other proteins to form supramolecular complexes, stabilized by numerous weak interactions.

Although protein sequences carry all necessary information for the folding into their native conformation, this correct folding requires the right environment—pH, ionic strength, metal ion concentrations, and so forth. Self-assembly therefore requires both information (provided by the DNA sequence) and environment (the interior of a living cell), and in this sense the DNA sequence alone is not enough to dictate the formation of a cell. As Rudolph Virchow, the nineteenth-century Prussian pathologist and researcher, concluded, "Omnis cellula e cellula": every cell comes from another cell.

SUMMARY 1.4 Genetic Foundations

- Genetic information is encoded in the linear sequence of four deoxyribonucleotides in DNA.
- The double-helical DNA molecule contains an internal template for its own replication and repair.
- The linear sequence of amino acids in a protein, which is encoded in the DNA of the gene for that protein, produces a protein's unique three-dimensional structure.
- Individual macromolecules with specific affinity for other macromolecules self-assemble into supramolecular complexes.

1.5 Evolutionary Foundations

Nothing in biology makes sense except in the light of evolution.

—*Theodosius Dobzhansky,* The American Biology Teacher, *March 1973*

Great progress in biochemistry and molecular biology during the decades since Dobzhansky made this striking generalization has amply confirmed its validity. The remarkable similarity of metabolic pathways and gene sequences in organisms across the phyla argues strongly that all modern organisms share a common evolutionary progenitor and were derived from it by a series of small changes (mutations), each of which conferred a selective advantage to some organism in some ecological niche.

Changes in the Hereditary Instructions Allow Evolution

Despite the near-perfect fidelity of genetic replication, infrequent, unrepaired mistakes in the DNA replication process lead to changes in the nucleotide sequence of DNA, producing a genetic **mutation** (Fig. 1–32) and changing the instructions for some cellular component. Incorrectly repaired damage to one of the DNA strands has the same effect. Mutations in the DNA handed down

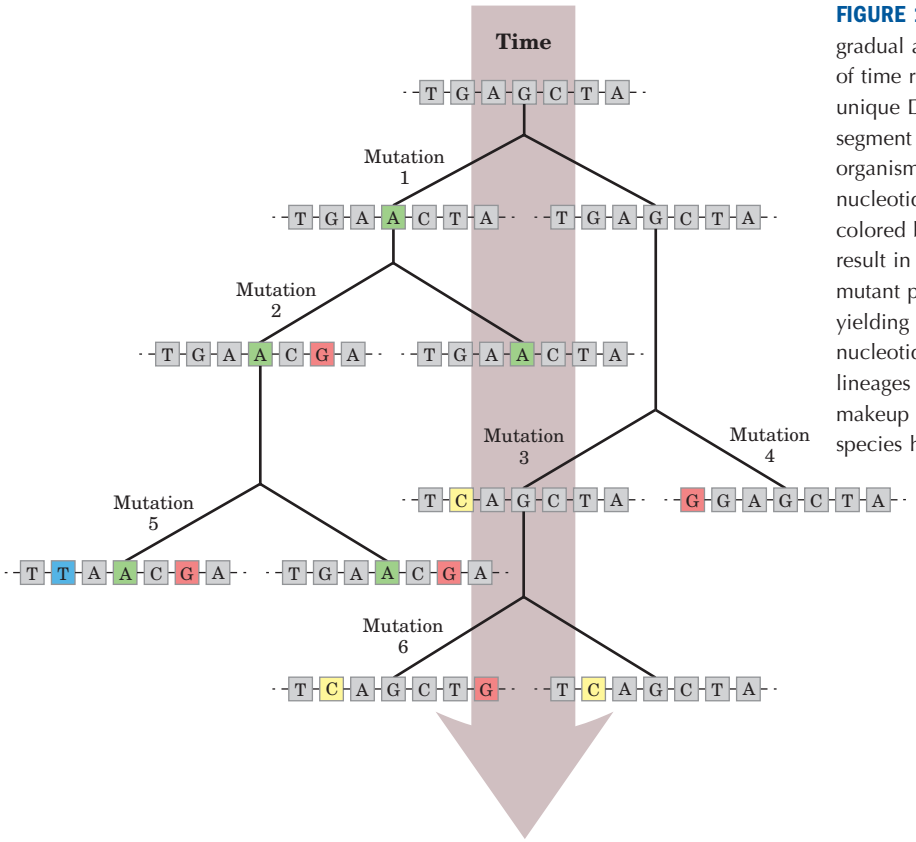


FIGURE 1–32 Role of mutation in evolution. The gradual accumulation of mutations over long periods of time results in new biological species, each with a unique DNA sequence. At the top is shown a short segment of a gene in a hypothetical progenitor organism. With the passage of time, changes in nucleotide sequence (mutations, indicated here by colored boxes), *occurring one nucleotide at a time,* result in progeny with different DNA sequences. These mutant progeny also undergo occasional mutations, yielding their own progeny that differ by two or more nucleotides from the progenitor sequence. When two lineages have diverged so much in their genetic makeup that they can no longer interbreed, a new species has been created.

to offspring—that is, mutations that are carried in the reproductive cells—may be harmful or even lethal to the organism; they may, for example, cause the synthesis of a defective enzyme that is not able to catalyze an essential metabolic reaction. Occasionally, however, a mutation better equips an organism or cell to survive in its environment. The mutant enzyme might have acquired a slightly different specificity, for example, so that it is now able to use some compound that the cell was previously unable to metabolize. If a population of cells were to find itself in an environment where that compound was the only or the most abundant available source of fuel, the mutant cell would have a selective advantage over the other, unmutated **(wild-type)** cells in the population. The mutant cell and its progeny would survive and prosper in the new environment, whereas wild-type cells would starve and be eliminated. This is what Darwin meant by "survival of the fittest under selective pressure."

Occasionally, a whole gene is duplicated. The second copy is superfluous, and mutations in this gene will not be deleterious; it becomes a means by which the cell may evolve: by producing a new gene with a new function while retaining the original gene and gene function. Seen in this light, the DNA molecules of modern organisms are historical documents, records of the long journey from the earliest cells to modern organisms. The historical accounts in DNA are not complete; in the course of evolution, many mutations must have been erased or written over. But DNA molecules are the best source of biological history that we have.

Several billion years of adaptive selection have refined cellular systems to take maximum advantage of the chemical and physical properties of the molecular raw materials for carrying out the basic energy-transforming and self-replicating activities of a living cell. Chance genetic variations in individuals in a population, combined with natural selection (survival and reproduction of the fittest individuals in a challenging or changing environment), have resulted in the evolution of an enormous variety of organisms, each adapted to life in its particular ecological niche.

Biomolecules First Arose by Chemical Evolution

In our account thus far we have passed over the first chapter of the story of evolution: the appearance of the first living cell. Apart from their occurrence in living organisms, organic compounds, including the basic biomolecules such as amino acids and carbohydrates, are found in only trace amounts in the earth's crust, the sea, and the atmosphere. How did the first living organisms acquire their characteristic organic building blocks? In 1922, the biochemist Aleksandr I. Oparin proposed a theory for the origin of life early in the history of Earth, postulating that the atmosphere was very different from that of today. Rich in methane, ammonia, and water, and essentially devoid of oxygen, it was a reducing atmosphere, in contrast to the oxidizing environment of our era. In Oparin's theory, electrical energy from lightning discharges or heat energy from volcanoes caused ammonia, methane, water vapor, and other components of the primitive atmosphere to react, forming simple organic compounds. These compounds then dissolved in the ancient seas, which over many millennia became enriched with a large variety of simple organic substances. In this warm solution (the "primordial soup"), some organic molecules had a greater tendency than others to associate into larger complexes. Over millions of years, these in turn assembled spontaneously to form membranes and catalysts (enzymes), which came together to become precursors of the earliest cells. Oparin's views remained speculative for many years and appeared untestable—until a surprising experiment was conducted using simple equipment on a desktop.

Chemical Evolution Can Be Simulated in the Laboratory

The classic experiment on the abiotic (nonbiological) origin of organic biomolecules was carried out in 1953 by Stanley Miller in the laboratory of Harold Urey. Miller subjected gaseous mixtures of NH_3, CH_4, H_2O, and H_2 to electrical sparks produced across a pair of electrodes (to simulate lightning) for periods of a week or more, then analyzed the contents of the closed reaction vessel (Fig. 1–33). The gas phase of the resulting mixture contained CO and CO_2, as well as the starting materials. The water phase contained a variety of organic compounds, including some amino acids, hydroxy acids, aldehydes, and hydrogen cyanide (HCN). This experiment established the possibility of abiotic production of biomolecules in relatively short times under relatively mild conditions.

More refined laboratory experiments have provided good evidence that many of the chemical components of living cells, including polypeptides and RNA-like molecules, can form under these conditions. Polymers of RNA can act as catalysts in biologically significant reactions (as we discuss in Chapters 26 and 27), and RNA probably played a crucial role in prebiotic evolution, both as catalyst and as information repository.

RNA or Related Precursors May Have Been the First Genes and Catalysts

In modern organisms, nucleic acids encode the genetic information that specifies the structure of enzymes, and enzymes catalyze the replication and repair of nucleic acids. The mutual dependence of these two classes of biomolecules brings up the perplexing question: which came first, DNA or protein?

The answer may be: neither. The discovery that RNA molecules can act as catalysts in their own forma-

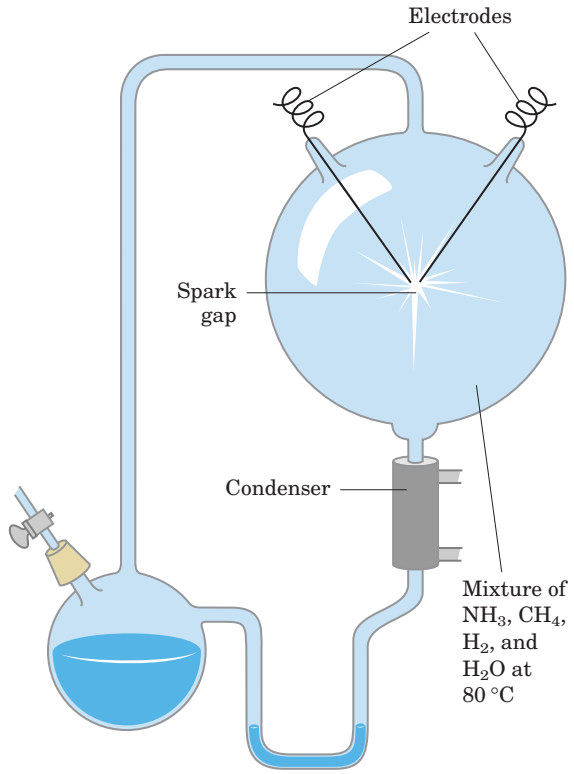


FIGURE 1–33 Abiotic production of biomolecules. Spark-discharge apparatus of the type used by Miller and Urey in experiments demonstrating abiotic formation of organic compounds under primitive atmospheric conditions. After subjection of the gaseous contents of the system to electrical sparks, products were collected by condensation. Biomolecules such as amino acids were among the products.

tion suggests that RNA or a similar molecule may have been the first gene *and* the first catalyst. According to this scenario (Fig. 1–34), one of the earliest stages of biological evolution was the chance formation, in the primordial soup, of an RNA molecule that could catalyze the formation of other RNA molecules of the same sequence—a self-replicating, self-perpetuating RNA. The concentration of a self-replicating RNA molecule would increase exponentially, as one molecule formed two, two formed four, and so on. The fidelity of self-replication was presumably less than perfect, so the process would generate variants of the RNA, some of which might be even better able to self-replicate. In the competition for nucleotides, the most efficient of the self-replicating sequences would win, and less efficient replicators would fade from the population.

The division of function between DNA (genetic information storage) and protein (catalysis) was, according to the "RNA world" hypothesis, a later development. New variants of self-replicating RNA molecules developed, with the additional ability to catalyze the condensation of amino acids into peptides. Occasionally, the peptide(s) thus formed would reinforce the self-replicating ability of the RNA, and the pair—RNA molecule and helping peptide—could undergo further modifications in sequence, generating even more efficient self-replicating systems. The recent, remarkable discovery that, in the protein-synthesizing machinery of modern cells (ribosomes), RNA molecules, not proteins, catalyze the formation of peptide bonds is certainly consistent with the RNA world hypothesis.

Some time after the evolution of this primitive protein-synthesizing system, there was a further development: DNA molecules with sequences complementary to the self-replicating RNA molecules took over the function of conserving the "genetic" information, and RNA molecules evolved to play roles in protein synthesis. (We explain in Chapter 8 why DNA is a more stable molecule than RNA and thus a better repository of inheritable information.) Proteins proved to be versatile catalysts and, over time, took over that function. Lipidlike compounds in the primordial soup formed relatively impermeable layers around self-replicating collections of molecules. The concentration of proteins and nucleic acids within these lipid enclosures favored the molecular interactions required in self-replication.

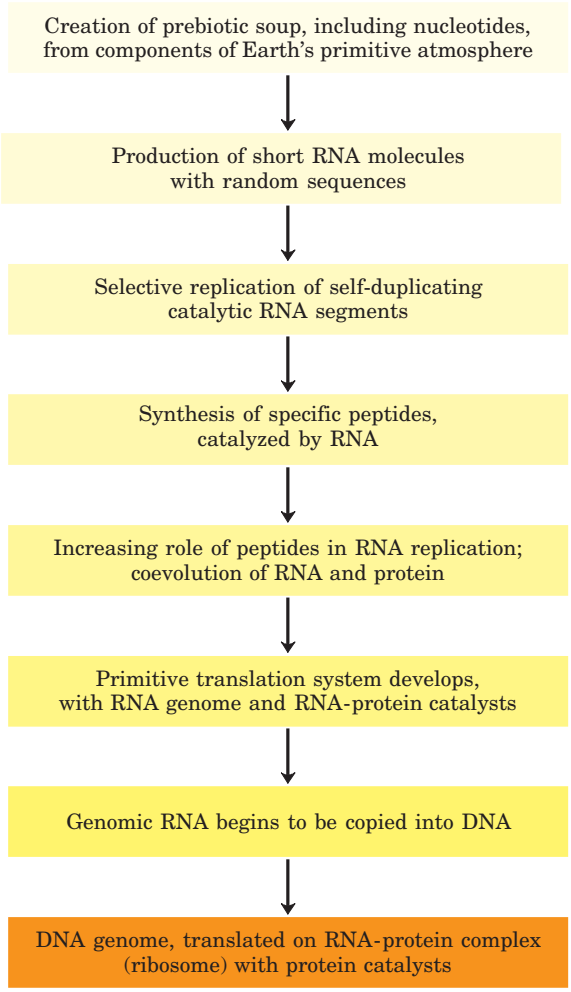


FIGURE 1–34 A possible "RNA world" scenario.

Biological Evolution Began More Than Three and a Half Billion Years Ago

Earth was formed about 4.5 billion years ago, and the first evidence of life dates to more than 3.5 billion years ago. In 1996, scientists working in Greenland found not fossil remains but chemical evidence of life from as far back as 3.85 billion years ago, forms of carbon embedded in rock that appear to have a distinctly biological origin. Somewhere on Earth during its first billion years there arose the first simple organism, capable of replicating its own structure from a template (RNA?) that was the first genetic material. Because the terrestrial atmosphere at the dawn of life was nearly devoid of oxygen, and because there were few microorganisms to scavenge organic compounds formed by natural processes, these compounds were relatively stable. Given this stability and eons of time, the improbable became inevitable: the organic compounds were incorporated into evolving cells to produce increasingly effective self-reproducing catalysts. The process of biological evolution had begun.

The First Cell Was Probably a Chemoheterotroph

The earliest cells that arose in the rich mixture of organic compounds, the primordial soup of prebiotic times, were almost certainly chemoheterotrophs (Fig. 1–5). The organic compounds they required were originally synthesized from components of the early atmosphere—CO, CO_2, N_2, CH_4, and such—by the nonbiological actions of volcanic heat and lightning. Early heterotrophs gradually acquired the ability to derive energy from compounds in their environment and to use that energy to synthesize more of their own precursor molecules, thereby becoming less dependent on outside sources. A very significant evolutionary event was the development of pigments capable of capturing the energy of light from the sun, which could be used to reduce, or "fix," CO_2 to form more complex, organic compounds. The original electron donor for these **photosynthetic** processes was probably H_2S, yielding elemental sulfur or sulfate (SO_4^{2-}) as the by-product, but later cells developed the enzymatic capacity to use H_2O as the electron donor in photosynthetic reactions, eliminating O_2 as waste. Cyanobacteria are the modern descendants of these early photosynthetic oxygen-producers.

Because the atmosphere of Earth in the earliest stages of biological evolution was nearly devoid of oxygen, the earliest cells were anaerobic. Under these conditions, chemoheterotrophs could oxidize organic compounds to CO_2 by passing electrons not to O_2 but to acceptors such as SO_4^{2-}, yielding H_2S as the product. With the rise of O_2-producing photosynthetic bacteria, the atmosphere became progressively richer in oxygen—a powerful oxidant and deadly poison to anaerobes. Responding to the evolutionary pressure of the "oxygen holocaust," some lineages of microorganisms gave rise to aerobes that obtained energy by passing electrons from fuel molecules to oxygen. Because the transfer of electrons from organic molecules to O_2 releases a great deal of energy, aerobic organisms had an energetic advantage over their anaerobic counterparts when both competed in an environment containing oxygen. This advantage translated into the predominance of aerobic organisms in O_2-rich environments.

Modern bacteria inhabit almost every ecological niche in the biosphere, and there are bacteria capable of using virtually every type of organic compound as a source of carbon and energy. Photosynthetic bacteria in both fresh and marine waters trap solar energy and use it to generate carbohydrates and all other cell constituents, which are in turn used as food by other forms of life. The process of evolution continues—and in rapidly reproducing bacterial cells, on a time scale that allows us to witness it in the laboratory.

Eukaryotic Cells Evolved from Prokaryotes in Several Stages

Starting about 1.5 billion years ago, the fossil record begins to show evidence of larger and more complex organisms, probably the earliest eukaryotic cells (Fig. 1–35).

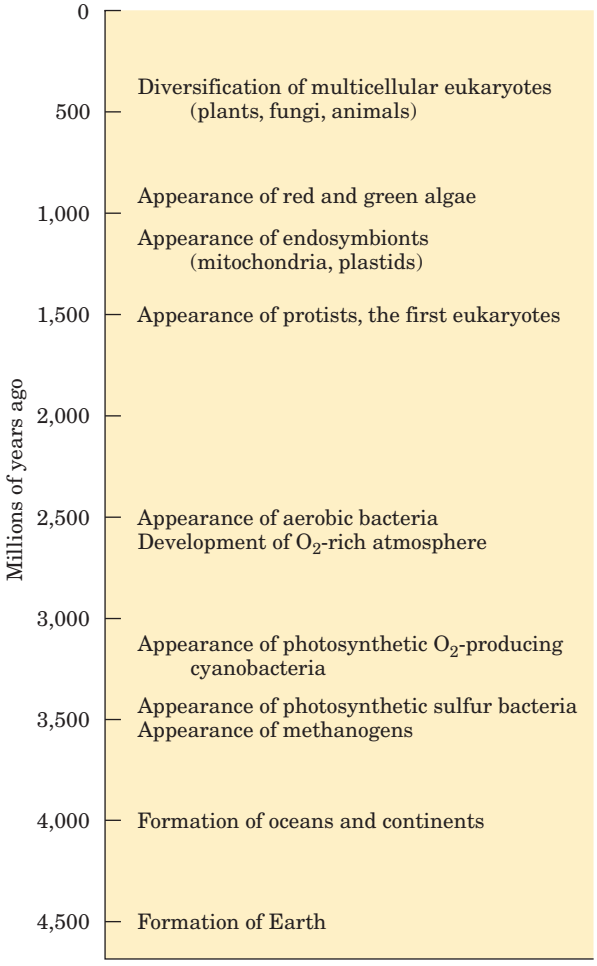


FIGURE 1–35 **Landmarks in the evolution of life on Earth.**

Details of the evolutionary path from prokaryotes to eukaryotes cannot be deduced from the fossil record alone, but morphological and biochemical comparisons of modern organisms have suggested a sequence of events consistent with the fossil evidence.

Three major changes must have occurred as prokaryotes gave rise to eukaryotes. First, as cells acquired more DNA, the mechanisms required to fold it compactly into discrete complexes with specific proteins and to divide it equally between daughter cells at cell division became more elaborate. For this, specialized proteins were required to stabilize folded DNA and to pull the resulting DNA-protein complexes (*chromosomes*) apart during cell division. Second, as cells became larger, a system of intracellular membranes developed, including a double membrane surrounding the DNA. This membrane segregated the nuclear process of RNA synthesis on a DNA template from the cytoplasmic process of protein synthesis on ribosomes. Finally, early eukaryotic cells, which were incapable of photosynthesis or aerobic metabolism, enveloped aerobic bacteria or photosynthetic bacteria to form **endosymbiotic** associations that became permanent (Fig. 1–36). Some aerobic bacteria evolved into the mitochondria of modern eukaryotes, and some photosynthetic cyanobacteria became the plastids, such as the chloroplasts of green algae, the likely ancestors of modern plant cells. Prokaryotic and eukaryotic cells are compared in Table 1–3.

At some later stage of evolution, unicellular organisms found it advantageous to cluster together, thereby acquiring greater motility, efficiency, or reproductive success than their free-living single-celled competitors. Further evolution of such clustered organisms led to permanent associations among individual cells and eventually to specialization within the colony—to cellular differentiation.

The advantages of cellular specialization led to the evolution of ever more complex and highly differentiated organisms, in which some cells carried out the sensory functions, others the digestive, photosynthetic, or reproductive functions, and so forth. Many modern multicellular organisms contain hundreds of different cell types, each specialized for some function that supports the entire organism. Fundamental mechanisms that evolved early have been further refined and embellished through evolution. The same basic structures and mechanisms that underlie the beating motion of cilia in *Paramecium* and of flagella in *Chlamydomonas* are employed by the highly differentiated vertebrate sperm cell.

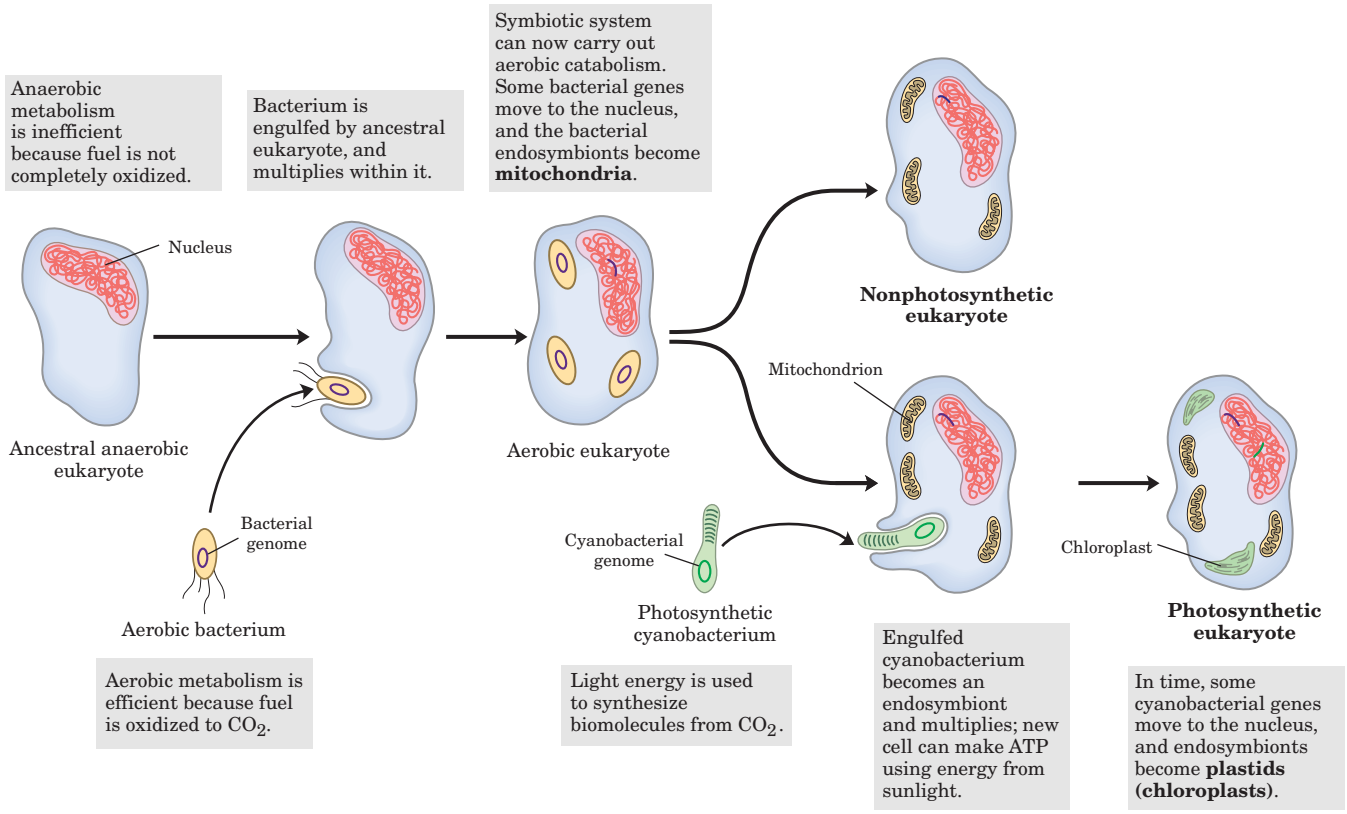


FIGURE 1–36 Evolution of eukaryotes through endosymbiosis. The earliest eukaryote, an anaerobe, acquired endosymbiotic purple bacteria (yellow), which carried with them their capacity for aerobic catabolism and became, over time, mitochondria. When photosynthetic cyanobacteria (green) subsequently became endosymbionts of some aerobic eukaryotes, these cells became the photosynthetic precursors of modern green algae and plants.

TABLE 1–3 Comparison of Prokaryotic and Eukaryotic Cells

Characteristic	*Prokaryotic cell*	*Eukaryotic cell*
Size	Generally small (1–10 μm)	Generally large (5–100 μm)
Genome	DNA with nonhistone protein; genome in nucleoid, not surrounded by membrane	DNA complexed with histone and nonhistone proteins in chromosomes; chromosomes in nucleus with membranous envelope
Cell division	Fission or budding; no mitosis	Mitosis, including mitotic spindle; centrioles in many species
Membrane-bounded organelles	Absent	Mitochondria, chloroplasts (in plants, some algae), endoplasmic reticulum, Golgi complexes, lysosomes (in animals), etc.
Nutrition	Absorption; some photosynthesis	Absorption, ingestion; photosynthesis in some species
Energy metabolism	No mitochondria; oxidative enzymes bound to plasma membrane; great variation in metabolic pattern	Oxidative enzymes packaged in mitochondria; more unified pattern of oxidative metabolism
Cytoskeleton	None	Complex, with microtubules, intermediate filaments, actin filaments
Intracellular movement	None	Cytoplasmic streaming, endocytosis, phagocytosis, mitosis, vesicle transport

Source: Modified from Hickman, C.P., Roberts, L.S., & Hickman, F.M. (1990) *Biology of Animals,* 5th edn, p. 30, Mosby-Yearbook, Inc., St. Louis, MO.

Molecular Anatomy Reveals Evolutionary Relationships

The eighteenth-century naturalist Carolus Linnaeus recognized the anatomic similarities and differences among living organisms and used them to provide a framework for assessing the relatedness of species. Charles Darwin, in the nineteenth century, gave us a unifying hypothesis to explain the phylogeny of modern organisms—the origin of different species from a common ancestor. Biochemical research in the twentieth century revealed the molecular anatomy of cells of different species—the monomeric subunit sequences and the three-dimensional structures of individual nucleic acids and proteins. Biochemists now have an enormously rich and increasing treasury of evidence that can be used to analyze evolutionary relationships and to refine evolutionary theory.

The sequence of the **genome** (the complete genetic endowment of an organism) has been entirely determined for numerous eubacteria and for several archaebacteria; for the eukaryotic microorganisms *Saccharomyces cerevisiae* and *Plasmodium* sp.; for the plants *Arabidopsis thaliana* and rice; and for the multicellular animals *Caenorhabditis elegans* (a roundworm), *Drosophila melanogaster* (the fruit fly), mice, rats, and *Homo sapiens* (you) (Table 1–4). More sequences are being added to this list regularly. With such sequences in hand, detailed and quantitative comparisons among species can provide deep insight into the evolutionary process. Thus far, the molecular phylogeny derived from gene sequences is consistent with, but in many cases more precise than, the classical phylogeny based on macroscopic structures. Although organisms have continuously diverged at the level of gross anatomy, at the molecular level the basic unity of life is readily apparent; molecular structures and mechanisms are remarkably similar from the simplest to the most complex organisms. These similarities are most easily seen at the level of sequences, either the DNA sequences that encode proteins or the protein sequences themselves.

When two genes share readily detectable sequence similarities (nucleotide sequence in DNA or amino acid sequence in the proteins they encode), their sequences

Carolus Linnaeus, 1701–1778

Charles Darwin, 1809–1882

TABLE 1–4 Some Organisms Whose Genomes Have Been Completely Sequenced

Organism	*Genome size (millions of nucleotide pairs)*	*Biological interest*
Mycoplasma pneumoniae	0.8	Causes pneumonia
Treponema pallidum	1.1	Causes syphilis
Borrelia burgdorferi	1.3	Causes Lyme disease
Helicobacter pylori	1.7	Causes gastric ulcers
Methanococcus jannaschii	1.7	Grows at 85 °C!
Haemophilus influenzae	1.8	Causes bacterial influenza
Methanobacterium thermo-autotrophicum	1.8	Member of the Archaea
Archaeoglobus fulgidus	2.2	High-temperature methanogen
Synechocystis sp.	3.6	Cyanobacterium
Bacillus subtilis	4.2	Common soil bacterium
Escherichia coli	4.6	Some strains cause toxic shock syndrome
Saccharomyces cerevisiae	12.1	Unicellular eukaryote
Plasmodium falciparum	23	Causes human malaria
Caenorhabditis elegans	97.1	Multicellular roundworm
Anopheles gambiae	278	Malaria vector
Mus musculus domesticus	2.5×10^3	Laboratory mouse
Homo sapiens	2.9×10^3	Human

are said to be homologous and the proteins they encode are **homologs.** If two homologous genes occur in the *same* species, they are said to be paralogous and their protein products are **paralogs.** Paralogous genes are presumed to have been derived by gene duplication followed by gradual changes in the sequences of both copies (Fig. 1–37). Typically, paralogous proteins are similar not only in sequence but also in three-dimensional structure, although they commonly have acquired different functions during their evolution.

Two homologous genes (or proteins) found in *different* species are said to be orthologous, and their protein products are **orthologs.** Orthologs are commonly found to have the same function in both organisms, and when a newly sequenced gene in one species is found to be strongly orthologous with a gene in another, this gene is presumed to encode a protein with the same function in both species. By this means, the function of gene products can be deduced from the genomic sequence, without any biochemical characterization of the gene product. An **annotated genome** includes, in addition to the DNA sequence itself, a description of the likely function of each gene product, deduced from comparisons with other genomic sequences and established protein functions. In principle, by identifying the pathways (sets of enzymes) encoded in a genome, we can deduce from the genomic sequence alone the organism's metabolic capabilities.

The sequence differences between homologous genes may be taken as a rough measure of the degree to which the two species have diverged during evolution—of how long ago their common evolutionary precursor gave rise to two lines with different evolutionary fates. The larger the number of sequence differences, the earlier the divergence in evolutionary history. One can construct a phylogeny (family tree) in which the evolutionary distance between any two species is represented by their proximity on the tree (Fig. 1–4 is an example).

As evolution advances, new structures, processes, or regulatory mechanisms are acquired, reflections of the changing genomes of the evolving organisms. The genome of a simple eukaryote such as yeast should have genes related to formation of the nuclear membrane, genes not present in prokaryotes. The genome of an insect should contain genes that encode proteins involved in specifying the characteristic insect segmented body plan, genes not present in yeast. The genomes of all vertebrate animals should share genes that specify the development of a spinal column, and those of mammals should have unique genes necessary for the development of the placenta, a characteristic of mammals—and so on. Comparisons of the whole genomes of species in each phylum may lead to the identification of genes critical to fundamental evolutionary changes in body plan and development.

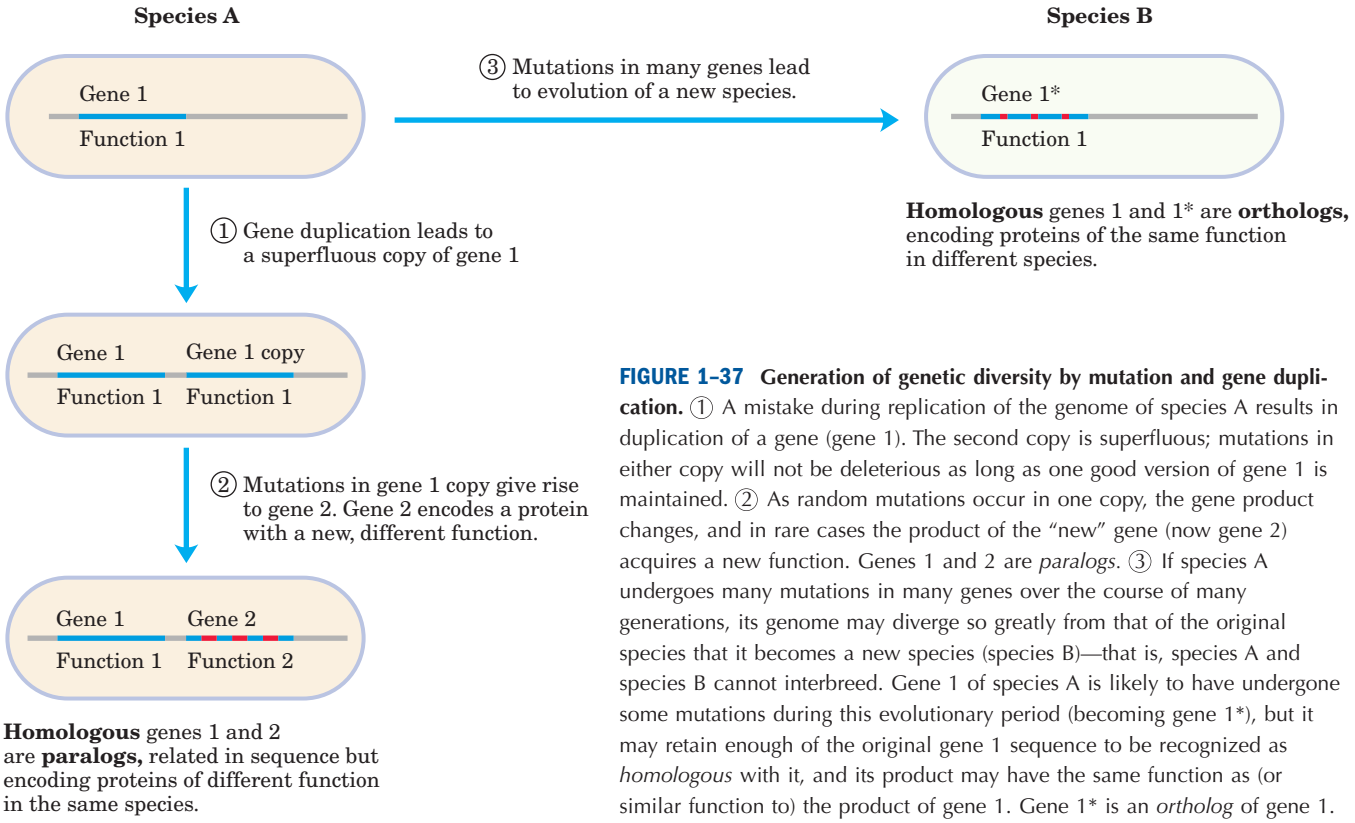


FIGURE 1–37 Generation of genetic diversity by mutation and gene duplication. ① A mistake during replication of the genome of species A results in duplication of a gene (gene 1). The second copy is superfluous; mutations in either copy will not be deleterious as long as one good version of gene 1 is maintained. ② As random mutations occur in one copy, the gene product changes, and in rare cases the product of the "new" gene (now gene 2) acquires a new function. Genes 1 and 2 are *paralogs*. ③ If species A undergoes many mutations in many genes over the course of many generations, its genome may diverge so greatly from that of the original species that it becomes a new species (species B)—that is, species A and species B cannot interbreed. Gene 1 of species A is likely to have undergone some mutations during this evolutionary period (becoming gene 1*), but it may retain enough of the original gene 1 sequence to be recognized as *homologous* with it, and its product may have the same function as (or similar function to) the product of gene 1. Gene 1* is an *ortholog* of gene 1.

Functional Genomics Shows the Allocations of Genes to Specific Cellular Processes

When the sequence of a genome is fully determined and each gene is annotated (that is, assigned a function), molecular geneticists can group genes according to the processes (DNA synthesis, protein synthesis, generation of ATP, and so forth) in which they function and thus find what fraction of the genome is allocated to each of a cell's activities. The largest category of genes in *E. coli*, *A. thaliana*, and *H. sapiens* consists of genes of as yet unknown function, which make up more than 40% of the genes in each species. The transporters that move ions and small molecules across plasma membranes take up a significant proportion of the genes in all three species, more in the bacterium and plant than in the mammal (10% of the 4,269 genes of *E. coli*, ~8% of the 25,706 genes of *A. thaliana*, and ~4% of the ~35,000 genes of *H. sapiens*). Genes that encode the proteins and RNA required for protein synthesis make up 3% to 4% of the *E. coli* genome, but in the more complex cells of *A. thaliana*, more genes are needed for targeting proteins to their final location in the cell than are needed to synthesize those proteins (about 6% and 2%, respectively). In general, the more complex the organism, the greater the proportion of its genome that encodes genes involved in the *regulation* of cellular processes and the smaller the proportion dedicated to the basic processes themselves, such as ATP generation and protein synthesis.

Genomic Comparisons Will Have Increasing Importance in Human Biology and Medicine

The genomes of chimpanzees and humans are 99.9% identical, yet the differences between the two species are vast. The relatively few differences in genetic endowment must explain the possession of language by humans, the extraordinary athleticism of chimpanzees, and myriad other differences. Genomic comparison will allow researchers to identify candidate genes linked to divergences in the developmental programs of humans and the other primates and to the emergence of complex functions such as language. The picture will become clearer only as more primate genomes become available for comparison with the human genome.

Similarly, the differences in genetic endowment among humans are vanishingly small compared with the differences between humans and chimpanzees, yet these differences account for the variety among us—including differences in health and in susceptibility to chronic diseases. We have much to learn about the variability in sequence among humans, and during the next decade the availability of genomic information will almost certainly transform medical diagnosis and treatment. We may expect that for some genetic diseases, palliatives will be replaced by cures; and that for disease susceptibilities associated with particular genetic markers, forewarning and perhaps increased preventive measures will prevail. Today's "medical history" may be replaced by a "medical forecast." ■

SUMMARY 1.5 Evolutionary Foundations

- Occasional inheritable mutations yield an organism that is better suited for survival in an ecological niche and progeny that are preferentially selected. This process of mutation and selection is the basis for the Darwinian evolution that led from the first cell to all the organisms that now exist, and it explains the fundamental similarity of all living organisms.
- Life originated about 3.5 billion years ago, most likely with the formation of a membrane-enclosed compartment containing a self-replicating RNA molecule. The components for the first cell were produced by the action of lightning and high temperature on simple atmospheric molecules such as CO_2 and NH_3.
- The catalytic and genetic roles of the early RNA genome were separated over time, with DNA becoming the genomic material and proteins the major catalytic species.
- Eukaryotic cells acquired the capacity for photosynthesis and for oxidative phosphorylation from endosymbiotic bacteria. In multicellular organisms, differentiated cell types specialize in one or more of the functions essential to the organism's survival.
- Knowledge of the complete genomic nucleotide sequences of organisms from different branches of the phylogenetic tree provides insights into the evolution and function of extant organisms and offers great opportunities in human medicine.

Key Terms

All terms are defined in the glossary.

metabolite 3
nucleus 3
genome 3
eukaryote 4
prokaryote 4
archaebacteria 4
eubacteria 4
cytoskeleton 9
stereoisomers 16
configuration 16
chiral center 17
conformation 19
entropy, *S* 23
enthalpy, *H* 23
free-energy change, ΔG 23
endergonic reaction 23
exergonic reaction 23
equilibrium 24
standard free-energy change, ΔG° 26
activation energy, $\Delta G^\ddagger$ 26
catabolism 27
anabolism 27
metabolism 27
mutation 31

Further Reading

General

Fruton, J.S. (1999) *Proteins, Enzymes, Genes: The Interplay of Chemistry and Biochemistry,* Yale University Press, New Haven.
A distinguished historian of biochemistry traces the development of this science and discusses its impact on medicine, pharmacy, and agriculture.

Harold, F.M. (2001) *The Way of the Cell: Molecules, Organisms, and the Order of Life,* Oxford University Press, Oxford.

Judson, H.F. (1996) *The Eighth Day of Creation: The Makers of the Revolution in Biology,* expanded edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
A highly readable and authoritative account of the rise of biochemistry and molecular biology in the twentieth century.

Kornberg, A. (1987) The two cultures: chemistry and biology. *Biochemistry* **26,** 6888–6891.
The importance of applying chemical tools to biological problems, described by an eminent practitioner.

Monod, J. (1971) *Chance and Necessity,* Alfred A. Knopf, Inc., New York. [Paperback edition, Vintage Books, 1972.] Originally published (1970) as *Le hasard et la nécessité,* Editions du Seuil, Paris.
An exploration of the philosophical implications of biological knowledge.

Pace, N.R. (2001) The universal nature of biochemistry. *Proc. Natl. Acad. Sci. USA* **98,** 805–808.
A short discussion of the minimal definition of life, on Earth and elsewhere.

Schrödinger, E. (1944) *What Is Life?* Cambridge University Press, New York. [Reprinted (1956) in *What Is Life? and Other Scientific Essays,* Doubleday Anchor Books, Garden City, NY.]
A thought-provoking look at life, written by a prominent physical chemist.

Cellular Foundations

Alberts, B., Johnson, A., Bray, D., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002) *Molecular Biology of the Cell,* 4th edn, Garland Publishing, Inc., New York.
A superb textbook on cell structure and function, covering the topics considered in this chapter, and a useful reference for many of the following chapters.

Becker, W.M., Kleinsmith, L.J., & Hardin, J. (2000) *The World of the Cell,* 5th edn, The Benjamin/Cummings Publishing Company, Redwood City, CA.
An excellent introductory textbook of cell biology.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C.A., Krieger, M., Scott, M.R., Zipursky, S.L., & Darnell, J. (2003) *Molecular Cell Biology,* 5th edn, W. H. Freeman and Company, New York.

Like the book by Alberts and coauthors, a superb text useful for this and later chapters.

Purves, W.K., Sadava, D., Orians, G.H., & Heller, H.C. (2001) *Life: The Science of Biology,* 6th edn, W. H. Freeman and Company, New York.

Chemical Foundations

Barta, N.S. & Stille, J.R. (1994) Grasping the concepts of stereochemistry. *J. Chem. Educ.* **71,** 20–23.

A clear description of the RS system for naming stereoisomers, with practical suggestions for determining and remembering the "handedness" of isomers.

Brewster, J.H. (1986) Stereochemistry and the origins of life. *J. Chem. Educ.* **63,** 667–670.

An interesting and lucid discussion of the ways in which evolution could have selected only one of two stereoisomers for the construction of proteins and other molecules.

Kotz, J.C. & Treichel, P., Jr. (1998) *Chemistry and Chemical Reactivity,* Saunders College Publishing, Fort Worth, TX.

An excellent, comprehensive introduction to chemistry.

Vollhardt, K.P.C. & Shore, N.E. (2002) *Organic Chemistry: Structure and Function,* W. H. Freeman and Company, New York.

Up-to-date discussions of stereochemistry, functional groups, reactivity, and the chemistry of the principal classes of biomolecules.

Physical Foundations

Atkins, P. W. & de Paula, J. (2001) *Physical Chemistry,* 7th edn, W. H. Freeman and Company, New York.

Atkins, P.W. & Jones, L. (1999) *Chemical Principles: The Quest for Insight,* W. H. Freeman and Company, New York.

Blum, H.F. (1968) *Time's Arrow and Evolution,* 3rd edn, Princeton University Press, Princeton.

An excellent discussion of the way the second law of thermodynamics has influenced biological evolution.

Genetic Foundations

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* **287,** 2185–2195.

Determination of the entire genome sequence of the fruit fly.

***Arabidopsis* Genome Initiative.** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408,** 796–815.

***C. elegans* Sequencing Consortium.** (1998) Genome sequence of the nematode *C. elegans:* a platform for investigating biology. *Science* **282,** 2012–2018.

Griffiths, A.J.F., Gelbart, W.M., Lewinton, R.C., & Miller, J.H. (2002) *Modern Genetic Analysis: Integrating Genes and Genomes,* W. H. Freeman and Company, New York.

International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921.

Jacob, F. (1973) *The Logic of Life: A History of Heredity,* Pantheon Books, Inc., New York. Originally published (1970) as *La logique du vivant: une histoire de l'hérédité,* Editions Gallimard, Paris.

A fascinating historical and philosophical account of the route by which we came to the present molecular understanding of life.

Pierce, B. (2002) *Genetics: A Conceptual Approach,* W. H. Freeman and Company, New York.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001) The sequence of the human genome. *Science* **291,** 1304–1351.

Evolutionary Foundations

Brow, J.R. & Doolittle, W.F. (1997) Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev*. **61,** 456–502.

A very thorough discussion of the arguments for placing the Archaea on the phylogenetic branch that led to multicellular organisms.

Darwin, C. (1964) *On the Origin of Species: A Facsimile of the First Edition (published in 1859),* Harvard University Press, Cambridge.

One of the most influential scientific works ever published.

de Duve, C. (1995) The beginnings of life on earth. *Am. Sci.* **83,** 428–437.

One scenario for the succession of chemical steps that led to the first living organism.

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

Dyer, B.D. & Obar, R.A. (1994) *Tracing the History of Eukaryotic Cells: The Enigmatic Smile,* Columbia University Press, New York.

Evolution of Catalytic Function. (1987) *Cold Spring Harb. Symp. Quant. Biol.* **52.**

A collection of almost 100 articles on all aspects of prebiotic and early biological evolution; probably the single best source on molecular evolution.

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

Discussion of the endosymbiotic hypothesis in the light of modern endosymbiotic anaerobic organisms.

Gesteland, R.F. & Atkins, J.F. (eds) (1993) *The RNA World,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

A collection of stimulating reviews on a wide range of topics related to the RNA world scenario.

Hall, B.G. (1982) Evolution on a Petri dish: the evolved β-galactosidase system as a model for studying acquisitive evolution in the laboratory. *Evolutionary Biol*. **15,** 85–150.

Knoll, A.H. (1991) End of the Proterozoic eon. *Sci. Am.* **265** (October), 64–73.

Discussion of the evidence that an increase in atmospheric oxygen led to the development of multicellular organisms, including large animals.

Lazcano, A. & Miller, S.L. (1996) The origin and early evolution of life: prebiotic chemistry, the pre-RNA world, and time. *Cell* **85,** 793–798.

Brief review of developments in studies of the origin of life: primitive atmospheres, submarine vents, autotrophic versus heterotrophic origin, the RNA and pre-RNA worlds, and the time required for life to arise.

Margulis, L. (1996) Archaeal-eubacterial mergers in the origin of Eukarya: phylogenetic classification of life. *Proc. Natl. Acad. Sci. USA* **93,** 1071–1076.

The arguments for dividing all living creatures into five kingdoms: Monera, Protoctista, Fungi, Animalia, Plantae. (Compare the Woese et al. paper below.)

Margulis, L., Gould, S.J., Schwartz, K.V., & Margulis, A.R. (1998) *Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth,* 3rd edn, W. H. Freeman and Company, New York.
Description of all major groups of organisms, beautifully illustrated with electron micrographs and drawings.

Mayr, E. (1997) *This Is Biology: The Science of the Living World,* Belknap Press, Cambridge, MA.
A history of the development of science, with special emphasis on Darwinian evolution, by an eminent Darwin scholar.

Miller, S.L. (1987) Which organic compounds could have occurred on the prebiotic earth? *Cold Spring Harb. Symp. Quant. Biol.* **52,** 17–27.
Summary of laboratory experiments on chemical evolution, by the person who did the original Miller-Urey experiment.

Morowitz, H.J. (1992) *Beginnings of Cellular Life: Metabolism Recapitulates Biogenesis,* Yale University Press, New Haven.

Schopf, J.W. (1992) *Major Events in the History of Life,* Jones and Bartlett Publishers, Boston.

Smith, J.M. & Szathmáry, E. (1995) *The Major Transitions in Evolution,* W. H. Freeman and Company, New York.

Woese, C.R., Kandler, O., & Wheelis, M.L. (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA* **87,** 4576–4579.
The arguments for dividing all living creatures into three kingdoms. (Compare the Margulis (1996) paper above.)

Problems

Some problems related to the contents of the chapter follow. (In solving end-of-chapter problems, you may wish to refer to the tables on the inside of the back cover.) Each problem has a title for easy reference and discussion.

1. The Size of Cells and Their Components

(a) If you were to magnify a cell 10,000 fold (typical of the magnification achieved using an electron microscope), how big would it appear? Assume you are viewing a "typical" eukaryotic cell with a cellular diameter of 50 μm.

(b) If this cell were a muscle cell (myocyte), how many molecules of actin could it hold? (Assume the cell is spherical and no other cellular components are present; actin molecules are spherical, with a diameter of 3.6 nm. The volume of a sphere is 4/3 πr^3.)

(c) If this were a liver cell (hepatocyte) of the same dimensions, how many mitochondria could it hold? (Assume the cell is spherical; no other cellular components are present; and the mitochondria are spherical, with a diameter of 1.5 μm.)

(d) Glucose is the major energy-yielding nutrient for most cells. Assuming a cellular concentration of 1 mM, calculate how many molecules of glucose would be present in our hypothetical (and spherical) eukaryotic cell. (Avogadro's number, the number of molecules in 1 mol of a nonionized substance, is 6.02×10^{23}.)

(e) Hexokinase is an important enzyme in the metabolism of glucose. If the concentration of hexokinase in our eukaryotic cell is 20 μM, how many glucose molecules are present per hexokinase molecule?

2. Components of *E. coli* *E. coli* cells are rod-shaped, about 2 μm long and 0.8 μm in diameter. The volume of a cylinder is $\pi r^2 h$, where h is the height of the cylinder.

(a) If the average density of *E. coli* (mostly water) is 1.1×10^3 g/L, what is the mass of a single cell?

(b) *E. coli* has a protective cell envelope 10 nm thick. What percentage of the total volume of the bacterium does the cell envelope occupy?

(c) *E. coli* is capable of growing and multiplying rapidly because it contains some 15,000 spherical ribosomes (diameter 18 nm), which carry out protein synthesis. What percentage of the cell volume do the ribosomes occupy?

3. Genetic Information in *E. coli* DNA The genetic information contained in DNA consists of a linear sequence of coding units, known as codons. Each codon is a specific sequence of three deoxyribonucleotides (three deoxyribonucleotide pairs in double-stranded DNA), and each codon codes for a single amino acid unit in a protein. The molecular weight of an *E. coli* DNA molecule is about 3.1×10^9 g/mol. The average molecular weight of a nucleotide pair is 660 g/mol, and each nucleotide pair contributes 0.34 nm to the length of DNA.

(a) Calculate the length of an *E. coli* DNA molecule. Compare the length of the DNA molecule with the cell dimensions (see Problem 2). How does the DNA molecule fit into the cell?

(b) Assume that the average protein in *E. coli* consists of a chain of 400 amino acids. What is the maximum number of proteins that can be coded by an *E. coli* DNA molecule?

4. The High Rate of Bacterial Metabolism Bacterial cells have a much higher rate of metabolism than animal cells. Under ideal conditions some bacteria double in size and divide every 20 min, whereas most animal cells under rapid growth conditions require 24 hours. The high rate of bacterial metabolism requires a high ratio of surface area to cell volume.

(a) Why does surface-to-volume ratio affect the maximum rate of metabolism?

(b) Calculate the surface-to-volume ratio for the spherical bacterium *Neisseria gonorrhoeae* (diameter 0.5 μm), responsible for the disease gonorrhea. Compare it with the surface-to-volume ratio for a globular amoeba, a large eukaryotic cell (diameter 150 μm). The surface area of a sphere is $4\pi r^2$.

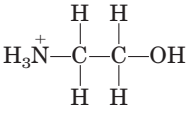
5. Fast Axonal Transport Neurons have long thin processes called axons, structures specialized for conducting signals throughout the organism's nervous system. Some axonal processes can be as long as 2 m—for example, the axons that originate in your spinal cord and terminate in the muscles of your toes. Small membrane-enclosed vesicles carrying materials essential to axonal function move along microtubules of the cytoskeleton, from the cell body to the tips of the axons.

(a) If the average velocity of a vesicle is 1 μm/s, how long does it take a vesicle to move from a cell body in the spinal cord to the axonal tip in the toes?

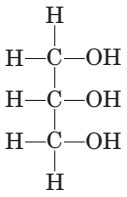
(b) Movement of large molecules by diffusion occurs relatively slowly in cells. (For example, hemoglobin diffuses at a rate of approximately 5 μm/s.) However, the diffusion of sucrose in an aqueous solution occurs at a rate approaching that of fast cellular transport mechanisms (about 4 μm/s). What are some advantages to a cell or an organism of fast, directed transport mechanisms, compared with diffusion alone?

6. Vitamin C: Is the Synthetic Vitamin as Good as the Natural One? A claim put forth by some purveyors of health foods is that vitamins obtained from natural sources are more healthful than those obtained by chemical synthesis. For example, pure L-ascorbic acid (vitamin C) extracted from rose hips is better than pure L-ascorbic acid manufactured in a chemical plant. Are the vitamins from the two sources different? Can the body distinguish a vitamin's source?

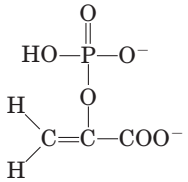
7. Identification of Functional Groups Figures 1–15 and 1–16 show some common functional groups of biomolecules. Because the properties and biological activities of biomolecules are largely determined by their functional groups, it is important to be able to identify them. In each of the compounds below, circle and identify by name each functional group.



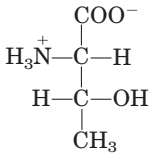
Ethanolamine
(a)



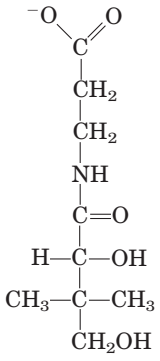
Glycerol
(b)



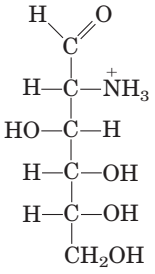
Phosphoenolpyruvate, an intermediate in glucose metabolism
(c)



Threonine, an amino acid
(d)

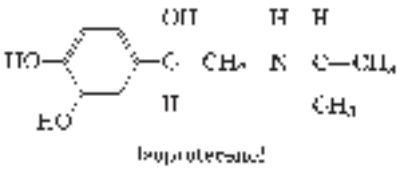


Pantothenate, a vitamin
(e)



D-Glucosamine
(f)

8. Drug Activity and Stereochemistry The quantitative differences in biological activity between the two enantiomers of a compound are sometimes quite large. For example, the D isomer of the drug isoproterenol, used to treat mild asthma, is 50 to 80 times more effective as a bronchodilator than the L isomer. Identify the chiral center in isoproterenol. Why do the two enantiomers have such radically different bioactivity?

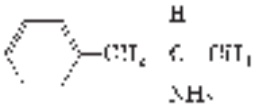


Isoproterenol

9. Separating Biomolecules In studying a particular biomolecule (a protein, nucleic acid, carbohydrate, or lipid) in the laboratory, the biochemist first needs to separate it from other biomolecules in the sample—that is, to *purify* it. Specific purification techniques are described later in the text. However, by looking at the monomeric subunits of a biomolecule, you should have some ideas about the characteristics of the molecule that would allow you to separate it from other molecules. For example, how would you separate (a) amino acids from fatty acids and (b) nucleotides from glucose?

10. Silicon-Based Life? Silicon is in the same group of the periodic table as carbon and, like carbon, can form up to four single bonds. Many science fiction stories have been based on the premise of silicon-based life. Is this realistic? What characteristics of silicon make it *less* well adapted than carbon as the central organizing element for life? To answer this question, consider what you have learned about carbon's bonding versatility, and refer to a beginning inorganic chemistry textbook for silicon's bonding properties.

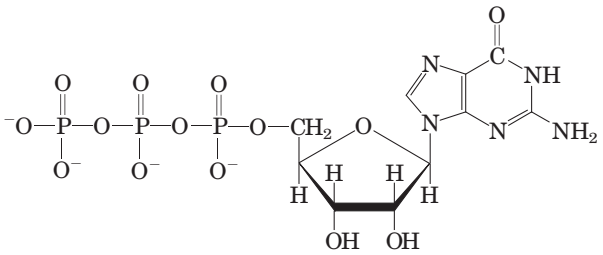
11. Drug Action and Shape of Molecules Some years ago two drug companies marketed a drug under the trade names Dexedrine and Benzedrine. The structure of the drug is shown below.



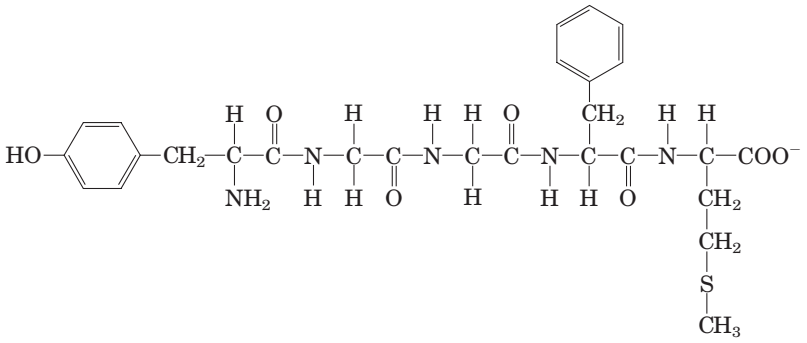
The physical properties (C, H, and N analysis, melting point, solubility, etc.) of Dexedrine and Benzedrine were identical. The recommended oral dosage of Dexedrine (which is still available) was 5 mg/day, but the recommended dosage of Benzedrine (no longer available) was twice that. Apparently it required considerably more Benzedrine than Dexedrine to yield the same physiological response. Explain this apparent contradiction.

12. Components of Complex Biomolecules Figure 1–10 shows the major components of complex biomolecules. For each of the three important biomolecules below (shown in their ionized forms at physiological pH), identify the constituents.

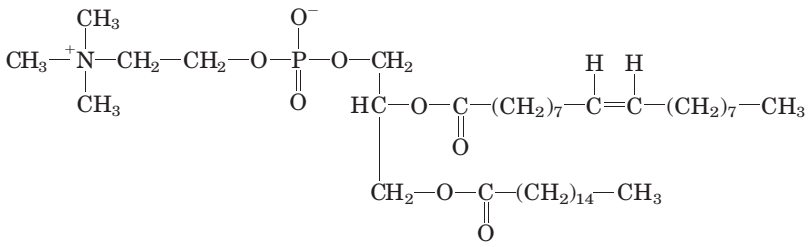
(a) Guanosine triphosphate (GTP), an energy-rich nucleotide that serves as a precursor to RNA:



(b) Phosphatidylcholine, a component of many membranes:



(c) Methionine enkephalin, the brain's own opiate:



13. Determination of the Structure of a Biomolecule An unknown substance, X, was isolated from rabbit muscle. Its structure was determined from the following observations and experiments. Qualitative analysis showed that X was composed entirely of C, H, and O. A weighed sample of X was completely oxidized, and the H_2O and CO_2 produced were measured; this quantitative analysis revealed that X contained 40.00% C, 6.71% H, and 53.29% O by weight. The molecular mass of X, determined by mass spectrometry, was 90.00 u (atomic mass units; see Box 1–1). Infrared spectroscopy showed that X contained one double bond. X dissolved readily in water to give an acidic solution; the solution demonstrated optical activity when tested in a polarimeter.

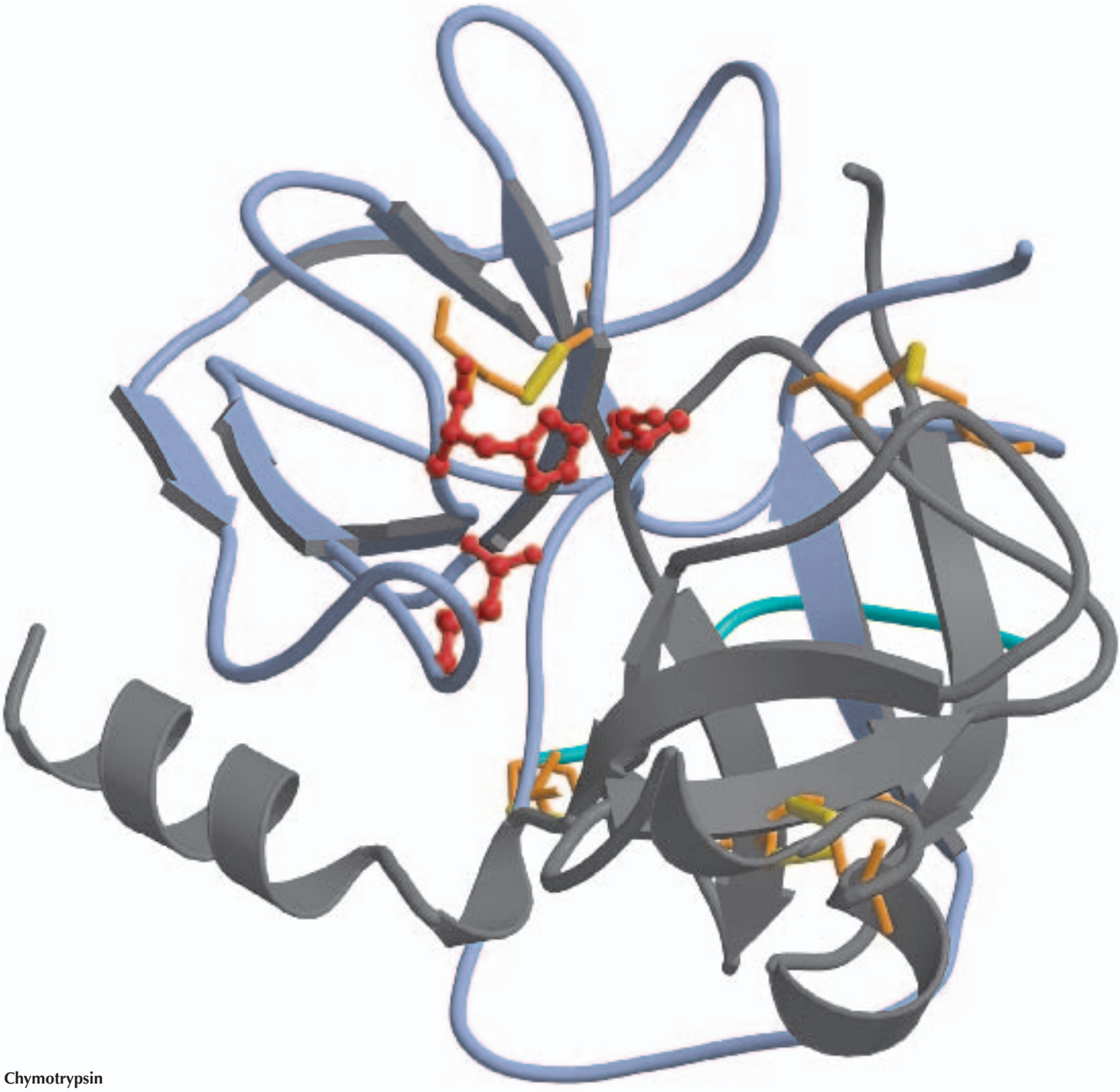
(a) Determine the empirical and molecular formula of X.

(b) Draw the possible structures of X that fit the molecular formula and contain one double bond. Consider *only* linear or branched structures and disregard cyclic structures. Note that oxygen makes very poor bonds to itself.

(c) What is the structural significance of the observed optical activity? Which structures in (b) are consistent with the observation?

(d) What is the structural significance of the observation that a solution of X was acidic? Which structures in (b) are consistent with the observation?

(e) What is the structure of X? Is more than one structure consistent with all the data?



Chymotrypsin