chapter

3

Biomolecules

Biochemistry aims to explain biological form and function in chemical terms. 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 chemical structure or catalytic activity. As we begin our study of biomolecules and their interactions in living cells, some basic questions come naturally to mind. What kinds of molecules are present in living organisms, and in what proportions? What are the structures of these molecules, and what forces stabilize their structures? What are their chemical properties and reactivities in isolation? How do they interact with each other? How and where did the biomolecules of the first living cells originate?

In this chapter, we review some of the chemical principles that govern the properties of biological molecules: the covalent bonding of carbon with itself and with other elements, the functional groups that occur in common biological molecules, the three-dimensional structure and stereochemistry of carbon compounds, the effects of chemical structure on reactivity, and the common classes of chemical reactions that occur in living organisms. Next, we discuss the monomeric subunits from which macromolecules are constructed and the energetics of their polymerization. Finally, we consider the origin of these monomeric subunits from simple compounds in the earth's atmosphere during prebiological times—that is, prebiotic evolution.

The chemical composition of living material, such as this jellyfish, differs from that of its physical environment, which for this organism is salt water.

Chemical Composition and Bonding

By the late eighteenth century, it had become clear to chemists 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.

Only about 30 of the more than 90 naturally occurring chemical elements are essential to living organisms. Most of the elements in living matter have relatively low atomic numbers; only five have atomic numbers above that of selenium, 34 (Fig. 3–1). The four most abundant elements in living organisms, in terms of the percentage of the total number of atoms, are hydrogen, oxygen, nitrogen, and carbon, which together make up over 99% of the mass of most cells. They are the lightest elements capable of forming one, two, three, and four bonds, respectively (Fig. 3–2). In general, the lightest elements form the strongest bonds. The trace elements (Fig. 3–1) represent a miniscule fraction of the weight of the human body, but all

figure 3–1

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.

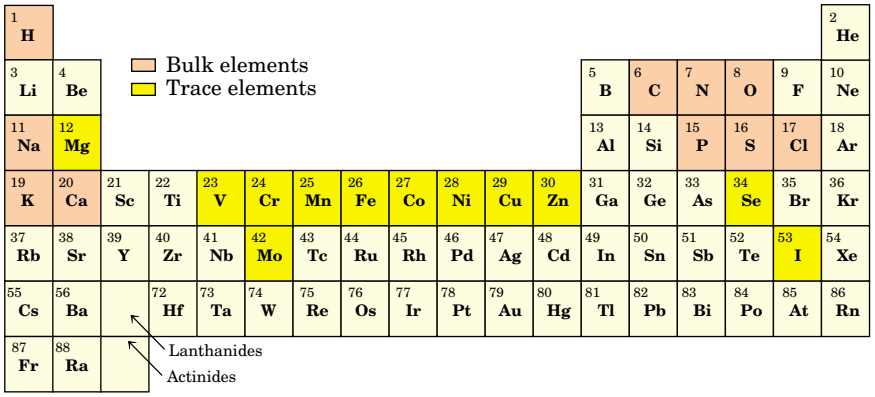
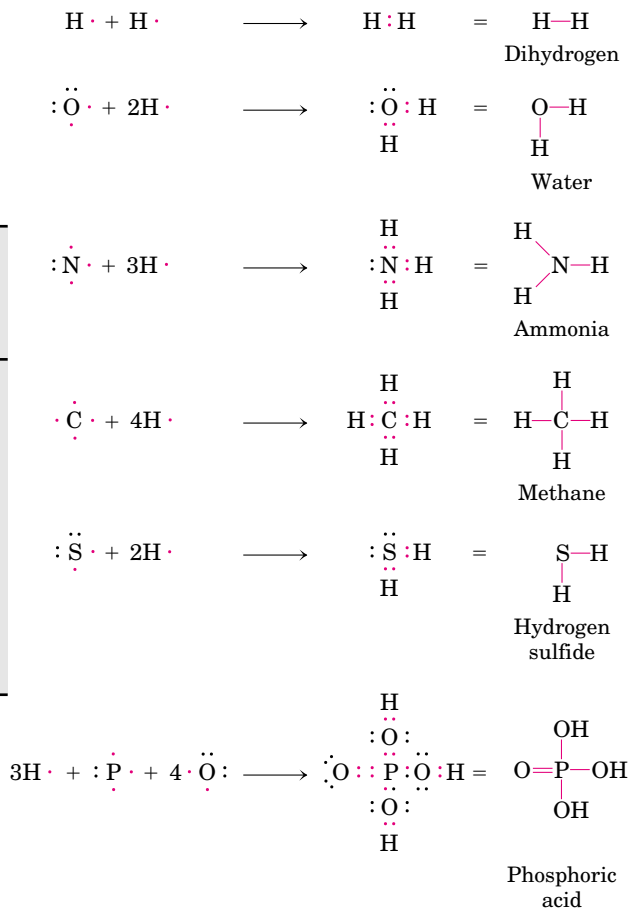


figure 3–2

Covalent bonding. Two atoms with unpaired electrons in their outer shells can form covalent bonds with each other by sharing electron pairs. Atoms participating in covalent bonding tend to fill their outer electron shells.

Atom	Number of unpaired electrons (in red)	Number of electrons in complete outer shell
H·	1	2
:Ö·	2	8
:N·	3	8
·C·	4	8
:S·	2	8
:P·	3	8



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

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. 3–3). Of greatest significance in biology is the ability of carbon atoms to share electron pairs with each other to form very stable carbon–carbon single bonds. Each carbon atom can form single bonds with one, two, three, or four other carbon atoms. Two carbon atoms also can share two (or three) electron pairs, thus forming double (or triple) bonds (Fig. 3–3).

The four single covalent bonds that can be formed by a carbon atom are arranged tetrahedrally, with an angle of about 109.5° between any two bonds (Fig. 3–4) 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. Molecules with covalently bonded carbon backbones are called **organic compounds;** they occur in limitless variety. Most biomolecules are organic compounds; we can therefore infer 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.

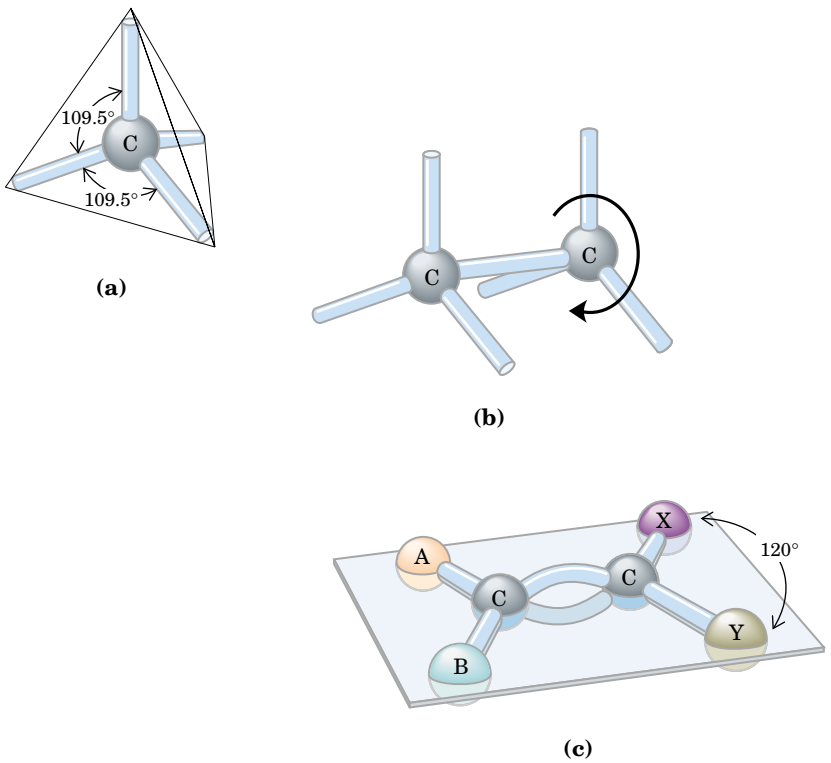


figure 3–3
Versatility of carbon in forming covalent single, double, and triple bonds (in red), particularly between carbon atoms. Triple bonds occur only rarely in biomolecules.

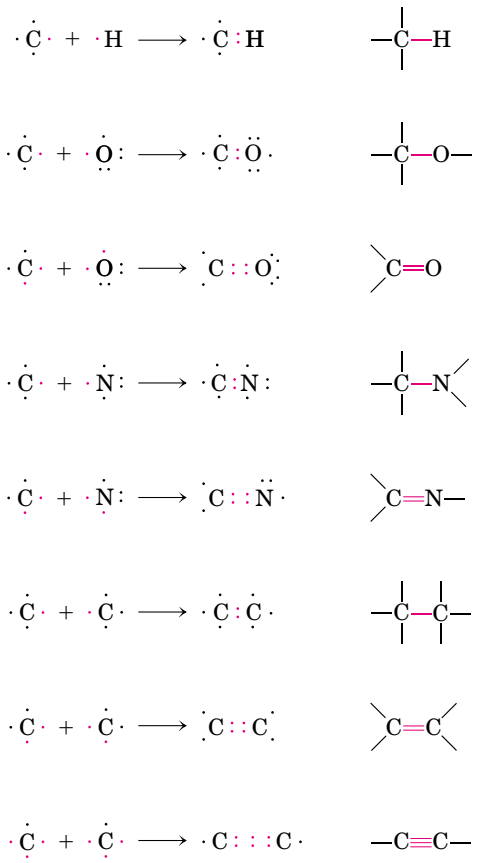


figure 3–4
Geometry of carbon bonding. (a) Carbon atoms have a characteristic tetrahedral arrangement of their four single bonds, which are about 0.154 nm long and at an angle of 109.5° to each other. **(b)** Carbon–carbon single bonds have freedom of rotation, as shown for the compound ethane (CH_3—CH_3). **(c)** Double bonds are shorter and do not allow free rotation. The single bonds on each doubly bonded carbon make an angle of 120° with each other. The two doubly bonded carbons and the atoms designated A, B, X, and Y all lie in the same rigid plane.

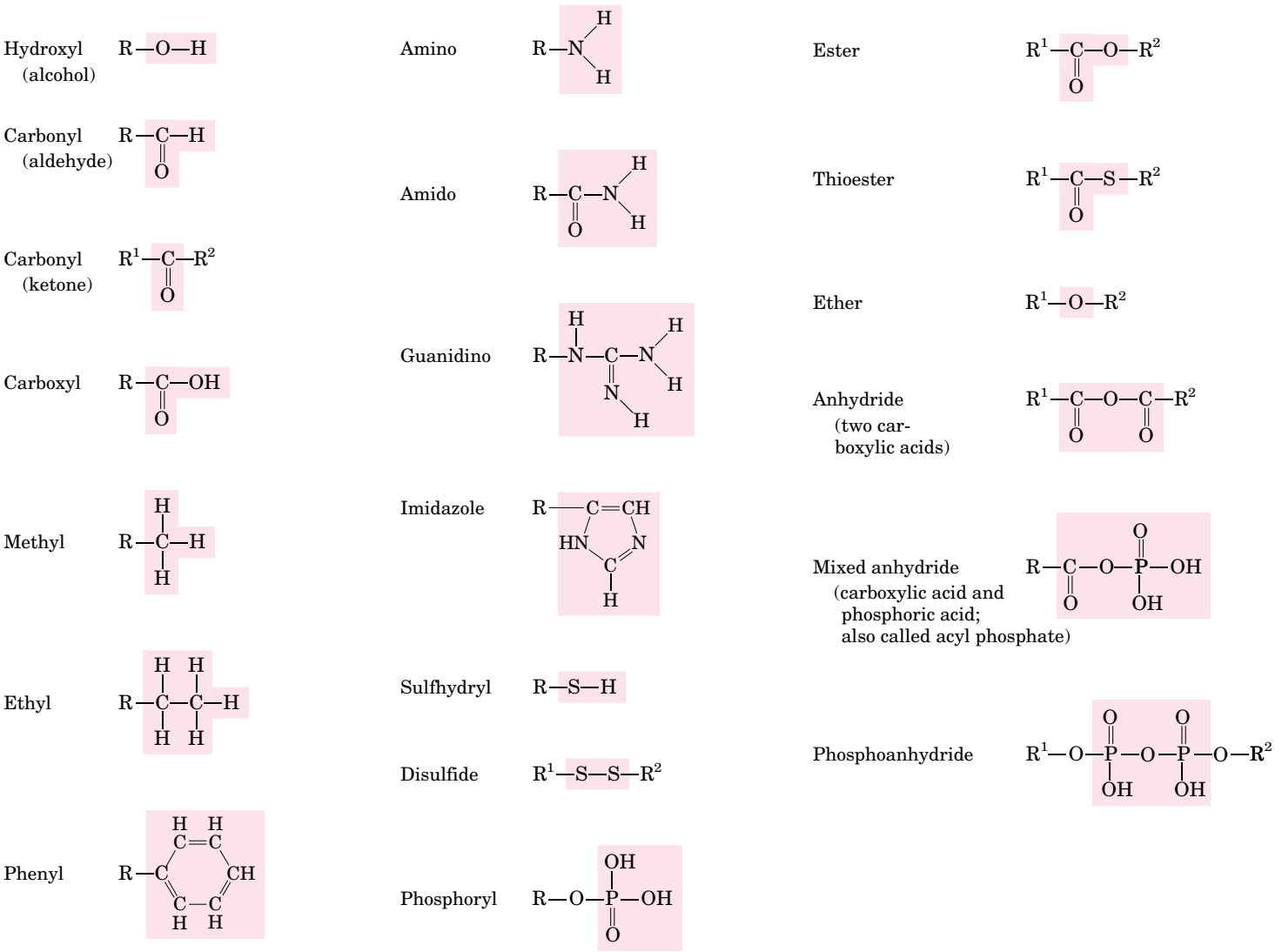
Functional Groups Determine Chemical Properties

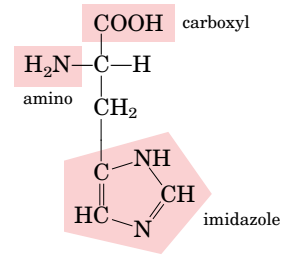
Most biomolecules can be regarded as derivatives of hydrocarbons, compounds with a covalently linked carbon backbone to which only hydrogen atoms are bonded. The backbones of hydrocarbons are very stable. The hydrogen atoms may be replaced by a variety of functional groups to yield different families of organic compounds. Typical of these are alcohols, which have one or more hydroxyl groups; amines, which have amino groups; aldehydes and ketones, which have carbonyl groups; and carboxylic acids, which have carboxyl groups (Fig. 3–5).

Many biomolecules are polyfunctional, containing two or more different kinds of functional groups (Fig. 3–6), each with its own chemical characteristics and reactions. The chemical "personality" of a compound such as epinephrine or acetyl-coenzyme A is determined by the chemistry of its functional groups and their disposition in three-dimensional space.

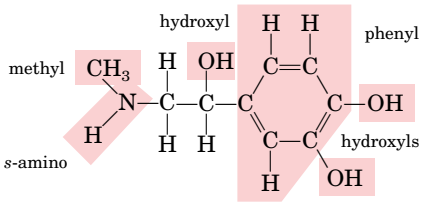
figure 3–5

Some common functional groups of biomolecules. All groups are shown in their uncharged (nonionized) form. 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.

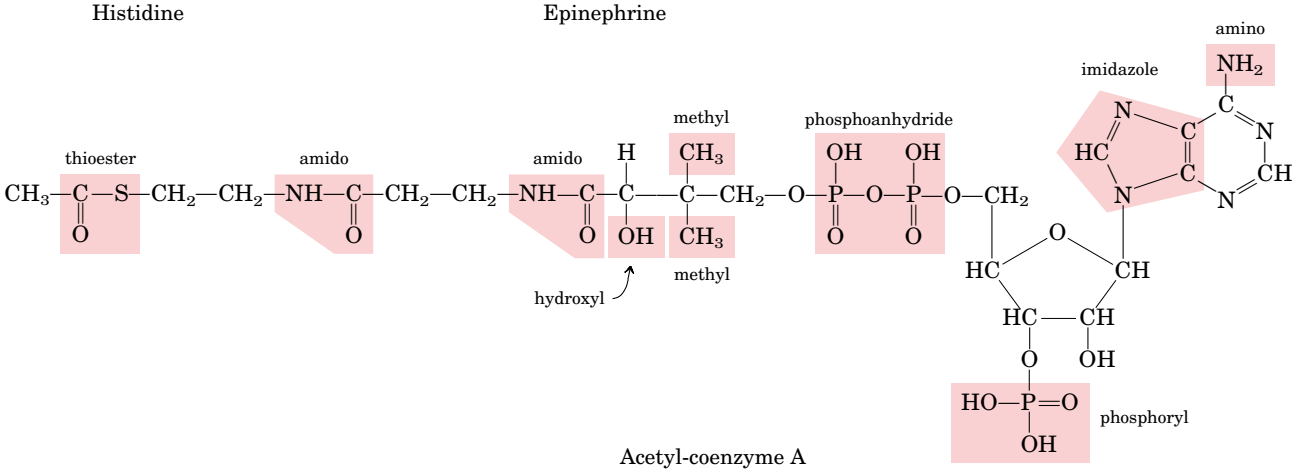




Histidine



Epinephrine



Acetyl-coenzyme A

figure 3–6

Common functional groups in biomolecules. Histidine is one of the amino acids found in proteins; epinephrine is a hormone; acetyl-coenzyme A is a carrier of acetyl groups in some enzymatic reactions. Note that the amino group in epinephrine is a secondary (*s*) amino group, with one of its amino hydrogens replaced by another group.

Three-Dimensional Structure: Configuration and Conformation

Although the covalent bonds and functional groups of a biomolecule are central to its function, the arrangement of the molecule's constituent atoms in three-dimensional space—its stereochemistry—is also crucially important. Compounds of carbon commonly exist as **stereoisomers,** different molecules in which the order of bonding is the same, but the spatial relationship among the atoms is different. Molecular interactions between biomolecules are invariably stereospecific; that is, they require specific stereochemistry in the interacting molecules.

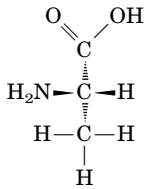
Figure 3–7 shows three ways to illustrate the stereochemical configuration of simple molecules. The perspective diagram specifies configuration unambiguously, but bond angles and center-to-center bond lengths are better represented with ball-and-stick models. In space-filling models, the radius of each atom is proportional to its van der Waals radius (Table 3–1), and the contours of the molecule represent the outer limits of the region from which atoms of other molecules are excluded.

table 3–1

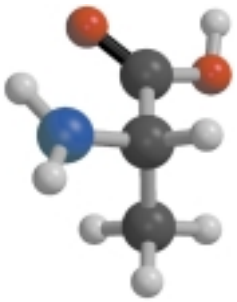
Van der Waals Radii and Covalent (Single-Bond) Radii of Some Elements*

Element	Van der Waals radius (nm)	Covalent radius for single bond (nm)
H	0.1	0.030
O	0.14	0.074
N	0.15	0.073
C	0.17	0.077
S	0.18	0.103
P	0.19	0.110
I	0.22	0.133

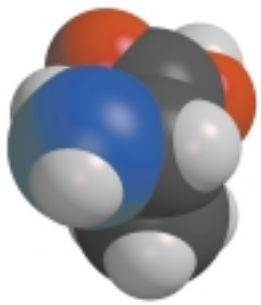
*Van der Waals radii describe the space-filling dimensions of atoms. When two atoms are joined covalently, the atomic radii at the point of bonding are less than the van der Waals radii, because the joined atoms are pulled together by the shared electron pair. The distance between nuclei in a van der Waals interaction or in a covalent bond is about equal to the sum of the van der Waals radii or the covalent radii, respectively, for the two atoms. Thus the length of a carbon–carbon single bond is about 0.077 nm + 0.077 nm = 0.154 nm.



(a)



(b)



(c)

figure 3–7

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 (see Table 3–1).

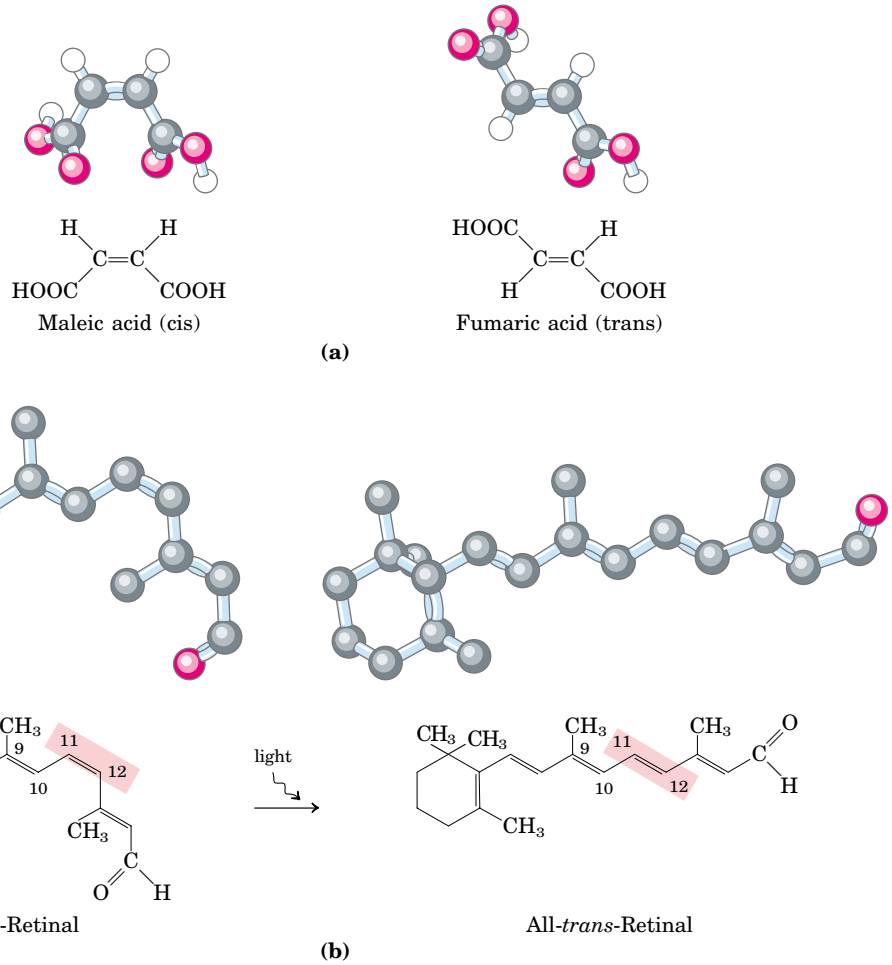
The Configuration of a Molecule Is Changed Only by Breaking a Bond

Configuration denotes the fixed spatial arrangement of atoms in an organic molecule that 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 3–8a 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. 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 these compounds have distinct biological roles despite their similar chemistry.

figure 3–8

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.



Four different substituents bonded to a tetrahedral carbon atom may be arranged two different ways in space (i.e., have two configurations; Fig. 3–9), 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 asym-

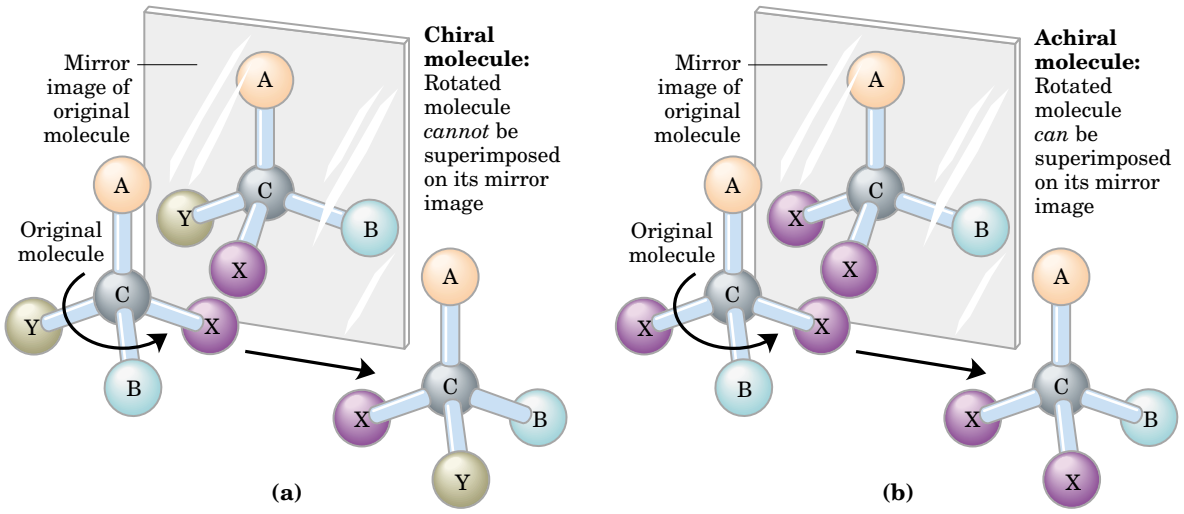


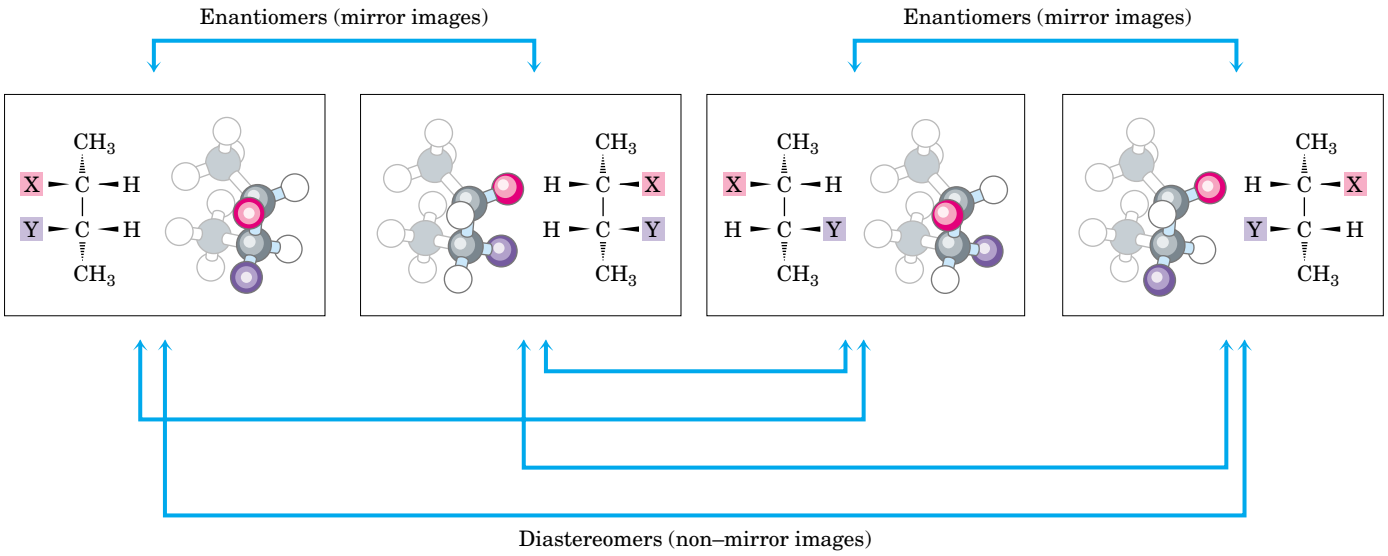
figure 3–9

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). Such a carbon atom is asymmetric and 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.

metric 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 only two stereoisomers, but 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. 3–9). Pairs of stereoisomers that are not mirror images of each other are called **diastereomers** (Fig. 3–10).

figure 3–10

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, and thus enantiomers. Other pairs are not mirror images; these are diastereomers. (Adapted from Carroll, F. (1998) *Perspectives on Structure and Mechanism in Organic Chemistry,* p. 63, Brooks/Cole Publishing Co., Pacific Grove, CA.)

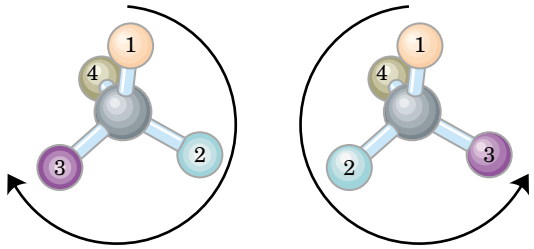


As Louis Pasteur observed (Box 3–1), enantiomers have nearly identical chemical properties but differ in a characteristic physical property, their interaction with plane-polarized light. In separate solutions, two enantiomers rotate the plane of plane-polarized light in opposite directions, but equimolar solutions of the two enantiomers (**racemic mixtures,** in the terminology of Pasteur) show no optical rotation. Compounds without chiral centers do not rotate the plane of plane-polarized light.

Biological interactions (between enzyme and substrate, receptor and hormone, or antibody and antigen, for example) are stereospecific: the "fit" in such interactions must be stereochemically correct. We must therefore name and represent the structure of a biomolecule so as to make its stereochemistry unambiguous. For compounds with more than one chiral center, the RS system of nomenclature is often more useful than the D and L system described in Chapter 5. In the RS system, each group attached to a chiral carbon is assigned a *priority.* The priorities of some common substituents are

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

The chiral atom is viewed with the group of lowest priority (4) 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 is (*S*) (Latin *sinister*, "left").



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

In this way each chiral carbon is designated as 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.

Molecular Conformation Is Changed by Rotation about Single Bonds

Molecular **conformation** refers to the spatial arrangement of substituent groups that, without breaking any bonds, are free to assume different positions in space because of the freedom of bond rotation. In the simple hydrocarbon ethane, for example, there is nearly complete freedom of rotation around the C—C bond. Many different, interconvertible conformations of the ethane molecule are therefore possible, depending on the degree of rotation (Fig. 3–11). 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. It is not possible to isolate either of these conformational forms, because 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.

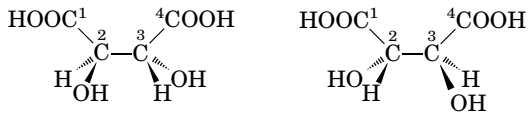
box 3–1 Louis Pasteur and Optical Activity: *In Vino, Veritas*

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 ("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.*

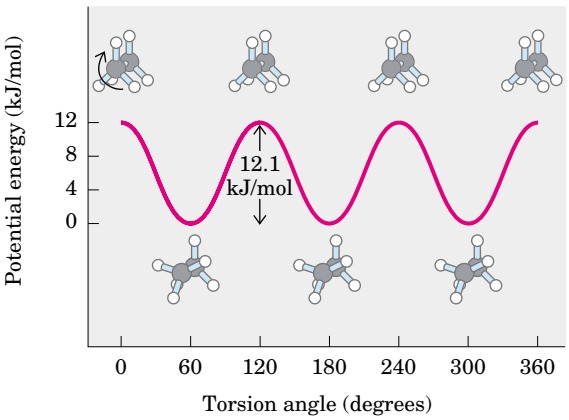
Now we do know. X-ray crystallographic studies in 1951 confirmed that the levorotatory and dextrorotatory forms of tartaric acid are mirror images of each other at the molecular level, and established the absolute configuration of each (Fig. 1). The same approach has been used to demonstrate that although 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 5).



(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 described 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.



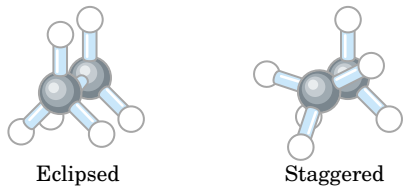


figure 3–11
Many conformations of ethane are possible because of freedom of rotation around the C—C bond. When the front carbon atom (as viewed by the reader) with its three attached hydrogens is rotated relative to the rear carbon atom, the potential energy of the molecule rises in the fully eclipsed conformation (torsion angle 0°, 120°, etc.), then falls 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.

Configuration and Conformation Define Biomolecular Structures

The three-dimensional structure of biomolecules large and small—the combination of configuration and conformation—is of the utmost importance in their biological interactions. For example, in the binding of a substrate (reactant) to the catalytic site of an enzyme, the two molecules must complement each other closely for effective catalysis. Such complementarity is also required in the binding of a hormone molecule to its receptor on a cell surface, or in the recognition and binding of an antigen by a specific antibody (Fig. 3–12).

The study of biomolecular stereochemistry with precise physical methods is an important part of modern research on cell structure and biochemical function. To date, the most productive method for structural investigations has been x-ray crystallography. For a compound that can be crystallized, the diffraction of x rays by crystals can be used to determine with great precision the position of every atom in the molecule relative to every other atom. The structures of most small biomolecules (those with fewer than about 50 atoms) and of hundreds of proteins and small nucleic acids have been determined by this means.

X-ray crystallography yields a static picture of the molecule within the confines of the crystal. However, biomolecules almost never exist as crystals within cells; rather, they are dissolved in the cytosol or associated with some other component(s) of the cell. Molecules have more freedom of intramolecular motion in solution than in a crystal. In a large molecule such as a protein, the small variations allowed in the conformations of its monomeric subunits add up to extensive flexibility; a protein may have several alternative stable conformations. Techniques such as nuclear magnetic resonance (NMR) spectroscopy complement x-ray crystallography by providing information about the three-dimensional structures of biomolecules in solution, including the flexibility of protein segments that might be held rigid in the crystal. With advances in the technology for producing very strong magnetic fields, NMR spectroscopy is increasingly used to determine the structure of larger molecules, including proteins (Fig. 3–13). The techniques of x-ray crystallography and NMR spectroscopy are described further in Box 6–3.

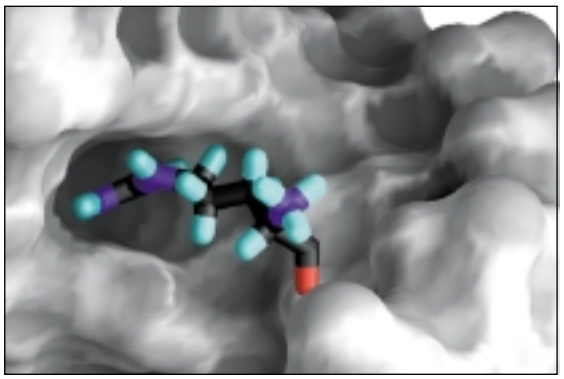


figure 3–12

Complementary fit of a macromolecule and a small molecule. Shown here are a segment of RNA from the regulatory region TAR of the HIV genome (gray) and argininamide (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 of the atoms in the molecule or by the "solvent exclusion volume," beyond which a water molecule cannot penetrate.

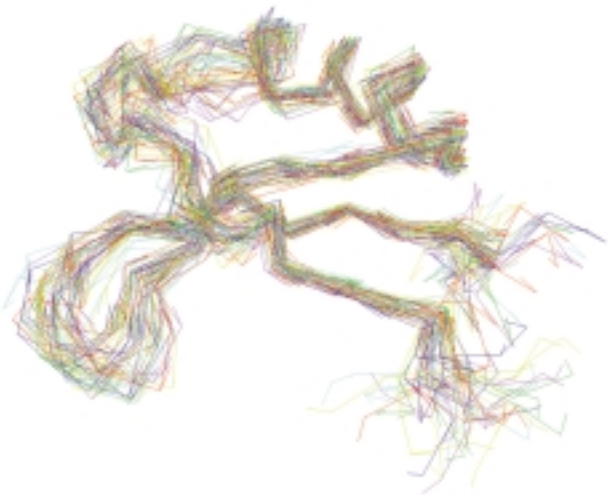


figure 3–13

Structure of the protein brazzein as determined by NMR spectroscopy. Brazzein, isolated from the fruit of *Pentadiplandra,* a plant of western Africa, is gram-for-gram 2,000 times sweeter than sucrose. In this representation, only the backbone of the protein is shown, without the individual amino acid side chains. NMR spectroscopy shows that in solution the protein can assume up to 43 very similar conformations, each shown here in a different color. This kind of conformational variability is not seen by x-ray crystallography, which requires that a protein be essentially frozen into one conformation in its crystalline form.

Interactions between Biomolecules Are Stereospecific

In living organisms, chiral molecules are usually present in only one of their chiral forms. For example, the amino acids occur in proteins only as the L isomers. Glucose, the monomeric subunit of starch, occurs biologically in only one of its chiral forms, the D isomer. (The conventions for naming stereoisomers of the amino acids are described in Chapter 5; those for sugars, in Chapter 9.) In contrast, when a compound with an asymmetric carbon atom is chemically synthesized in the laboratory, the nonbiological reactions usually produce all possible chiral forms in an equimolar mixture that does not rotate polarized light (a racemic mixture). The chiral forms in such a mixture can be separated only by painstaking physical methods (recall that Pasteur separated crystals with forceps). Chiral compounds in living cells are produced in only one chiral form because the enzymes that synthesize them are also chiral molecules.

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 3–14. Specific sensory receptors (for smell and taste) easily distinguish the members of each pair of diastereomers.

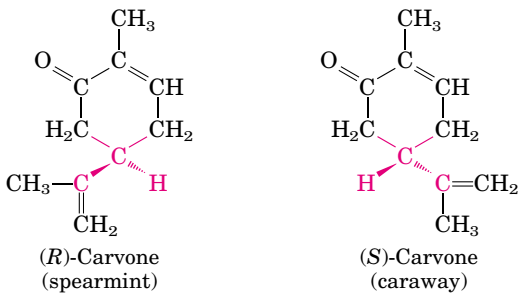


figure 3–14
Stereoisomers distinguishable by smell and taste in humans. **(a)** Two stereoisomers of carvone: (*R*)-carvone (isolated from spearmint oil) has the characteristic fragrance of spearmint; (*S*)-carvone (from caraway seed oil) smells like caraway. **(b)** Aspartame, the artificial sweetener 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.

Chemical Reactivity

The mechanisms of biochemical reactions are not fundamentally different from those of other chemical reactions. They may be understood and predicted from the nature of the functional groups of the reactants. Functional groups alter the electron distribution and the geometry of neighboring atoms and thus affect the chemical reactivity of the entire molecule. Although a large number of different chemical reactions occur in a typical cell, these reactions are of only a few general types. We will briefly review the basic facts about chemical bonding and reactivity, and then summarize five reaction types commonly encountered in biochemistry. Later we will consider specific reaction types in more detail.

Bond Strength Is Related to the Properties of the Bonded Atoms

In chemical reactions, bonds are broken and new ones are formed. The strength of a chemical bond depends on the relative electronegativities—the relative affinities for electrons—of the bonding elements (Table 3–2), the distance of the bonding electrons from each nucleus, and the nuclear charge of each atom. The number of electrons shared also influences bond strength; double bonds are stronger than single bonds, and triple bonds are stronger yet (Table 3–3). The strength of a bond is expressed as bond energy, in joules. (Biochemists have often used the calorie as a unit of energy—bond energy and free energy, for example, are often given in kcal/mol. The joule is the unit of energy in the International System of Units and has replaced the calorie in most biochemical usage; we use it throughout this book. For conversions, 1 cal equals 4.184 J.) Bond energy can be thought of as either the amount of energy required to break a bond or the amount of energy gained by the surroundings when the bond forms. One way to put energy into a system is to heat it, which gives the molecules more kinetic energy, raising the fraction of molecules with kinetic energies

table 3–2

The Electronegativities of Some Elements

Element	Electronegativity*
F	4.0
O	3.5
Cl	3.0
N	3.0
Br	2.8
S	2.5
C	2.5
I	2.5
Se	2.4
P	2.1
H	2.1
Cu	1.9
Fe	1.8
Co	1.8
Ni	1.8
Mo	1.8
Zn	1.6
Mn	1.5
Mg	1.2
Ca	1.0
Li	1.0
Na	0.9
K	0.8

*The higher the number, the more electronegative (the greater the electron affinity of) the element.

table 3–3

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	461	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.

high enough to react. When molecular motion is sufficiently violent, intramolecular vibrations and intermolecular collisions sometimes break chemical bonds and allow new ones to form.

When bonds are broken and formed in a chemical reaction, the difference between the energy extracted from the surroundings to break bonds and the energy released to the surroundings during the formation of new bonds can be approximated as the **enthalpy change, ΔH,** for the reaction. If heat energy is absorbed by the system as the change occurs (i.e., if the reaction is endothermic), then H has, by definition, a positive value; if heat is given off, the reaction is exothermic and H is negative. In short, the change in enthalpy for a chemical reaction reflects the kinds and numbers of bonds that are made and broken. As we shall see later in this chapter, the enthalpy change is one of three factors that determine the free-energy change for a reaction; the other two are the temperature and the change in entropy.

Five General Types of Chemical Transformations Occur in Cells

Most cells have the capacity to carry out thousands of specific, enzyme-catalyzed reactions: for example, transformation of a simple nutrient such as glucose into amino acids, nucleotides, or lipids; extraction of energy from fuels by oxidation; or polymerization of monomeric subunits into macromolecules. Fortunately for the student of biochemistry, there are patterns within this multitude of reactions; you do not need to learn all of these reactions to comprehend the molecular logic of life. Most of the reactions in living cells fall into one of five general categories: (1) oxidation-reductions, (2) cleavage and formation of carbon–carbon bonds, (3) internal rearrangements, (4) group transfers, and (5) condensation reactions in which monomeric subunits are joined, with the elimination of a molecule of water. Reactions within each general category usually occur by similar mechanisms.

All Oxidation-Reduction Reactions Involve Electron Transfer

When the two atoms sharing electrons in a covalent bond have the same electronegativity, as in the case of two carbon atoms, the bond is nonpolar. When two elements that differ in electronegativity form a covalent bond (e.g., C and O), that bond is polarized; the shared electrons are more likely to be in the region of the more electronegative atom (O in this case) than of the less electronegative atom (C in this case; see Table 3–2). In the extreme case of two elements of very different electronegativity, such as Na and Cl, one atom gives up its electron(s) to the other atom, resulting in the formation of ions and ionic interactions such as those in solid NaCl.

In carbon–hydrogen bonds, the more electronegative C "owns" the two electrons shared with H, but in carbon–oxygen bonds, electron sharing is unequal in favor of oxygen. Thus, in going from —CH_3 (an alkane) to —CH_2OH (an alcohol), the carbon atom has effectively lost electrons—which is, by definition, oxidation. As Figure 3–15 shows, carbon atoms encountered in biochemistry can exist in five oxidation states, depending on the elements with which carbon shares electrons.

In many biological oxidations, a compound loses two electrons and two hydrogen ions (i.e., two hydrogen atoms); these reactions are commonly called **dehydrogenations** and the enzymes that catalyze them are called **dehydrogenases** (Fig. 3–16). In some, but not all, biological oxidations, a carbon atom becomes covalently bonded to an oxygen atom. The enzymes that catalyze these oxidations are generally called **oxidases** or, if the oxygen atom is derived directly from molecular oxygen (O_2), **oxygenases.**

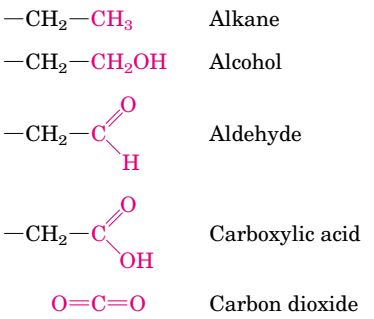


figure 3–15

The oxidation states of carbon in biomolecules. Each compound is formed by oxidation of the red carbon in the compound listed above it. Carbon dioxide is the most highly oxidized form of carbon found in living systems.

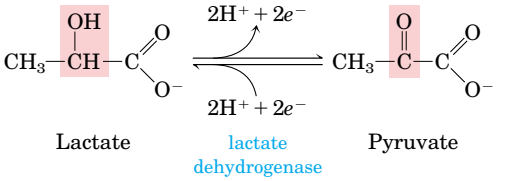


figure 3–16

An oxidation-reduction reaction. Shown here is the oxidation of lactate to pyruvate. In this dehydrogenation, two electrons and two hydrogen ions (the equivalent of two hydrogen atoms) are removed from the C-2 of lactate, a ketone. In cells the reaction is catalyzed by lactate dehydrogenase and the electrons are transferred to a cofactor called nicotinamide adenine dinucleotide. This reaction is fully reversible; pyruvate can be reduced by electrons from the cofactor. We will discuss the factors that determine the direction of a reaction in Chapter 14.

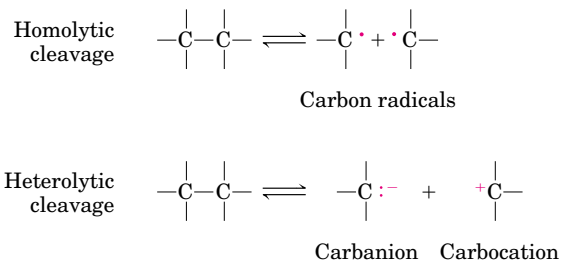
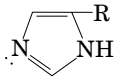


figure 3–17

Two mechanisms for cleavage of a C—C bond. In homolytic cleavages, each carbon atom keeps one of the bonding electrons, resulting in two carbon radicals (i.e., carbons having unpaired electrons). In heterolytic cleavages, one of the two carbon atoms keeps both bonding electrons, producing a carbanion; the other becomes a carbocation.

table 3–4

Some Functional Groups Active as Nucleophiles within Cells*

Water	$H\ddot{O}H$
Hydroxide ion	$H\ddot{O}:^-$
Hydroxyl (alcohol)	$R\ddot{O}H$
Alkoxyl	$R\ddot{O}:^-$
Sulfhydryl	RSH
Sulfide	$R\ddot{S}^-$
Amino	$R\ddot{N}H_2$
Carboxylate	$R-C(=O)-O^-$
Imidazole	
Inorganic orthophosphate	$^-O-P(=O)(-O^-)-OH$

*Listed in order of decreasing strength. Weaker nucleophiles make better leaving groups.

Every oxidation must be accompanied by a reduction, in which an electron acceptor acquires the electrons removed by oxidation. Oxidation reactions generally release energy (think of camp fires, in which various compounds in wood are oxidized by oxygen molecules in the air). Most living cells obtain the energy needed for cellular work by oxidizing metabolic fuels such as carbohydrate or fat; photosynthetic organisms can also trap and use the energy of sunlight. The catabolic (energy-yielding) pathways described in Chapters 15 to 19 are oxidative reaction sequences that result in the transfer of electrons from fuel molecules through a series of electron carriers and finally to oxygen. The high affinity of O_2 for electrons makes the overall electron-transfer process highly exergonic, providing the energy that drives ATP synthesis—the central goal of catabolism.

Carbon–Carbon Bonds Are Cleaved and Formed by Nucleophilic Substitution Reactions

A covalent bond can be broken in two general ways (Fig. 3–17). In **homolytic** cleavage, each atom leaves the bond as a radical, carrying one of the two electrons (now unpaired) that held the bonded atoms together. Homolytic reactions occur only rarely in living organisms (but see Fig. 22–39 for an example). More common are **heterolytic** cleavages in which one atom keeps both bonding electrons (forming an anion), leaving the other atom one electron short (a cation). When a second electron-rich group replaces the departing anion, a **nucleophilic substitution** occurs (Fig. 3–18). Many biochemical reactions involve interactions between **nucleophiles,** functional groups rich in electrons and capable of donating them, and **electrophiles,** electron-deficient functional groups that seek electrons. Nucleophiles combine with, and give up electrons to, electrophiles. Functional groups containing oxygen, nitrogen, and sulfur are important biological nucleophiles (Table 3–4). Positively charged hydrogen atoms (hydrogen ions, or protons) and positively charged metals (cations) frequently act as electrophiles in cells. A carbon atom can act as either a nucleophile or an electrophile, depending on which bonds and functional groups surround it.

There are two general mechanisms by which one nucleophile can replace another in the formation of carbon–carbon bonds. In the first (Fig. 3–19a), the leaving group (the nucleophile W; see Fig. 3–18) departs with its electrons, leaving the former partner as a relatively unstable carbocation (positively charged carbon, an electrophile), before the substituting group (Z, a nucleophile) comes on the scene. This mechanism is called an **SN1 reaction,** (SN1 indicating substitution nucleophilic, unimolecular). In the second type of nucleophilic substitution, an attacking nucleophile (Z) arrives before the leaving group (W) departs, and a pentacovalent intermediate forms transiently (Fig. 3–19b). This is an **SN2 reaction** (substitution nucleophilic, bimolecular). As Figure 3–19 suggests, SN2 reactions typically result in an inversion of configuration around the attacked carbon upon departure of the leaving group, whereas SN1 reactions usually result in either retention of the original configuration or racemization. In general, weaker nucleophiles make better leaving groups and stronger nucleophiles are better attacking species.

The aldol condensation catalyzed by aldolase (see Fig. 15–4) is an example of a nucleophilic substitution employed to form carbon–carbon bonds in cells. These reactions are reversible; aldolase can join two three-carbon moieties to form a six-carbon sugar, or it can split the six-carbon sugar to yield two three-carbon moieties.

figure 3–18

A nucleophilic substitution reaction. An electron-rich nucleophile (Z) attacks an electron-poor center (a carbon atom, for example) and displaces a nucleophilic group (W), which is called the leaving group.

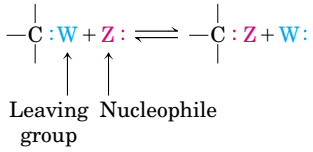
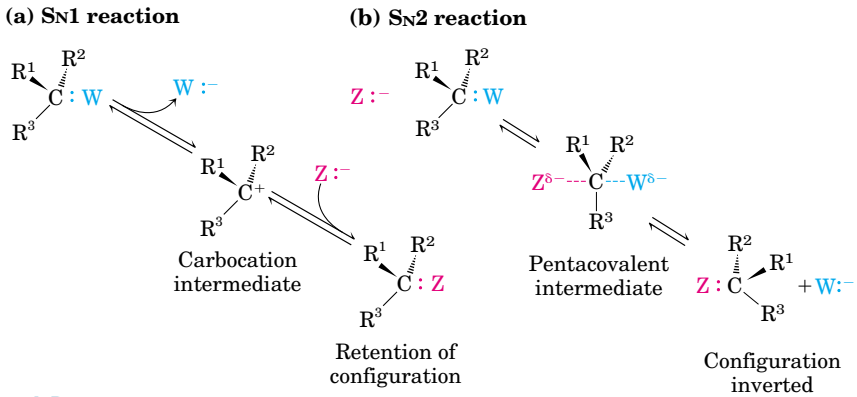


figure 3–19

Two classes of nucleophilic substitution reactions. **(a)** SN1: The leaving group (W) departs with a bonding electron, leaving a carbocation, before the attacking nucleophile (Z) arrives. **(b)** SN2: The attacking nucleophile (Z) approaches one side of the electrophilic carbon while the leaving group (W) remains bonded to the other side, resulting in a transient pentacovalent intermediate. The departure of W leaves the substituted compound with a completely inverted configuration at the reacting carbon atom.



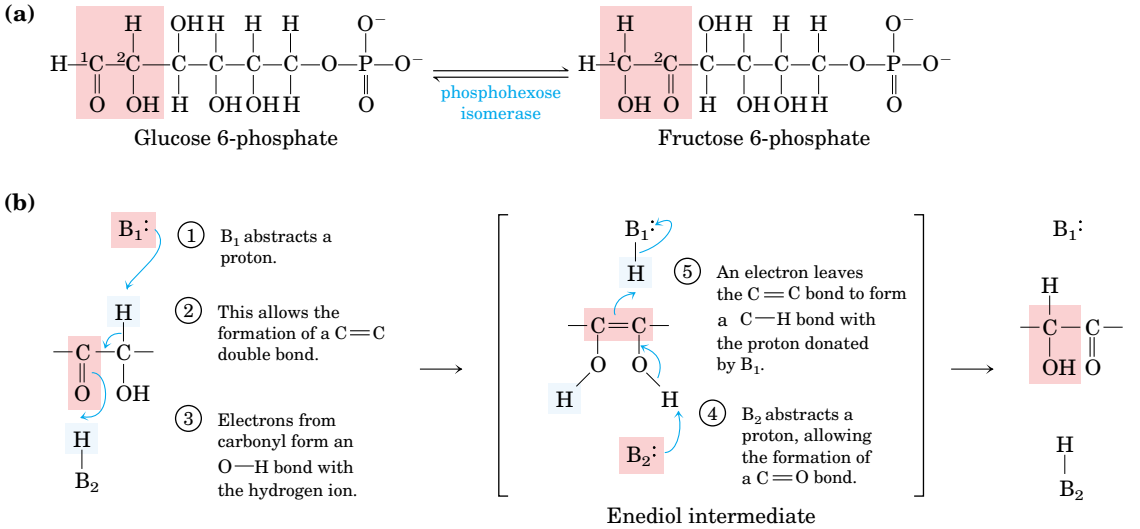
Electron Transfers within a Molecule Produce Internal Rearrangements

Another common cellular reaction type is intramolecular rearrangement, in which redistribution of electrons results in isomerization, transposition of double bonds, and cis-trans rearrangements of double bonds. An example of isomerization is the formation of fructose 6-phosphate from glucose 6-phosphate during sugar metabolism (Chapter 15). In this reaction (Fig. 3–20a), C-1 is reduced (from aldehyde to alcohol) and C-2 is oxidized (from alcohol to ketone). In Figure 3–20b, which shows the details of the electron movements that result in isomerization, we have employed the convention of "electron-pushing" diagrams, which we will use to indicate reaction mechanisms throughout the book. Curved blue arrows show the movement of electrons as the reaction proceeds.

A simple transposition of one C=C bond occurs during metabolism of the common fatty acid oleic acid (see Fig. 17–9), and we will see spectacular examples of double-bond repositioning in the synthesis of cholesterol (see Fig. 21–35).

figure 3–20

An isomerization reaction. **(a)** The conversion of glucose 6-phosphate to fructose 6-phosphate, a reaction of sugar metabolism catalyzed by phosphohexose isomerase. **(b)** This reaction proceeds through an enediol intermediate. The curved blue arrows represent the movement of bonding electrons from nucleophile (red) to electrophile (blue). B_1 and B_2 are basic groups on the enzyme; they are capable of donating and accepting hydrogen ions (protons) as the reaction progresses.



Group Transfer Reactions Activate Metabolic Intermediates

A general theme in metabolism is the attachment of a good leaving group to a metabolic intermediate to "activate" the intermediate for subsequent reaction. Among the better leaving groups in nucleophilic substitution reactions (Table 3–4) are inorganic orthophosphate (the ionized form of H_3PO_4 at neutral pH, a mixture of $H_2PO_4^-$ and HPO_4^{2-}, commonly abbreviated P_i) and inorganic pyrophosphate ($P_2O_7^{6-}$, abbreviated PP_i). Esters and anhydrides of phosphoric acid play central roles in cellular chemistry. Nucleophilic substitutions in which the phosphoryl group ($—PO_3^{2-}$) serves as a leaving group occur in hundreds of metabolic reactions; nucleophilic substitution is made more favorable by the attachment of a phosphoryl group to an otherwise poor leaving group such as —OH.

Phosphorus can form five covalent bonds. The conventional representation of P_i (Fig. 3–21a) with three P—O bonds and one P═O bond is not an accurate picture. In P_i, four equivalent P—O bonds share some double-bond character, and the anion has a tetrahedral structure (Fig. 3–21b). As oxygen is more electronegative than phosphorus, the sharing of electrons is unequal. The central phosphorus bears a positive charge and can therefore act as an electrophile. In a very large number of metabolic reactions, a phosphoryl group ($—PO_3^{2-}$) is transferred from ATP to an alcohol (forming a phosphate ester) (Fig. 3–21c) or to a carboxylic acid (forming a mixed anhydride; see Fig. 3–5). When a nucleophile attacks the electrophilic phosphorus atom in ATP, a relatively stable pentacovalent structure is formed as a reaction intermediate (Fig. 3–21d). With departure of the leaving group (ADP), the transfer of a phosphoryl group is complete. The large family of enzymes that catalyze phosphoryl group transfers with ATP as donor are called kinases (Greek *kinein*, "to move"). Hexokinase, for example, "moves" a phosphoryl group from ATP to glucose.

Phosphoryl groups are not the only activators of this type. Thioalcohols (thiols), in which the oxygen atom of an alcohol is replaced with a sulfur atom, are also good leaving groups. Thiols activate carboxylic acids by forming thioesters (thiol esters) with them (Fig. 3–5). We will encounter a number of cases, including the reactions catalyzed by the fatty acyl transferases in lipid synthesis (see Fig. 21–2), in which nucleophilic substitution at the carbonyl carbon of a thioester results in transfer of the acyl group to another moiety.

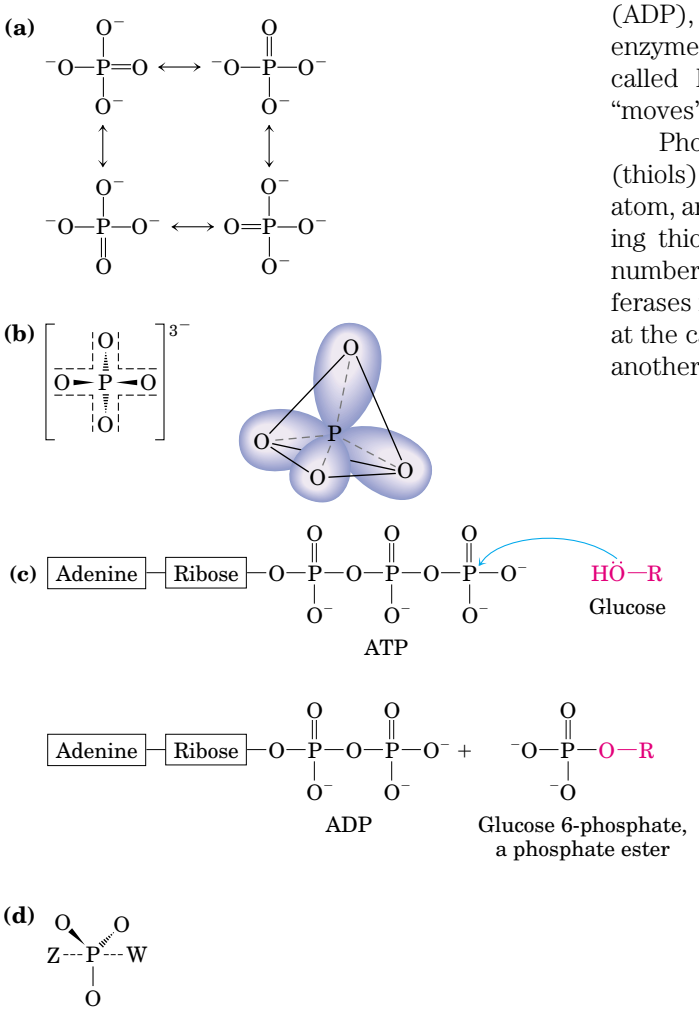


figure 3–21

Alternative ways of showing the structure of inorganic orthophosphate. (a) In one (inadequate) representation, three oxygens are single-bonded to phosphorus, and the fourth is double-bonded, allowing the four different resonance structures shown. **(b)** The four resonance structures can be represented more accurately by showing all four P—O bonds with some double-bond character; the hybrid orbitals so represented are arranged in a tetrahedron with P at its center. **(c)** When a nucleophile Z (in this case, the —OH on C-6 of glucose) attacks ATP, it displaces ADP (W). In this S_N2 reaction, a pentacovalent intermediate **(d)** forms transiently.

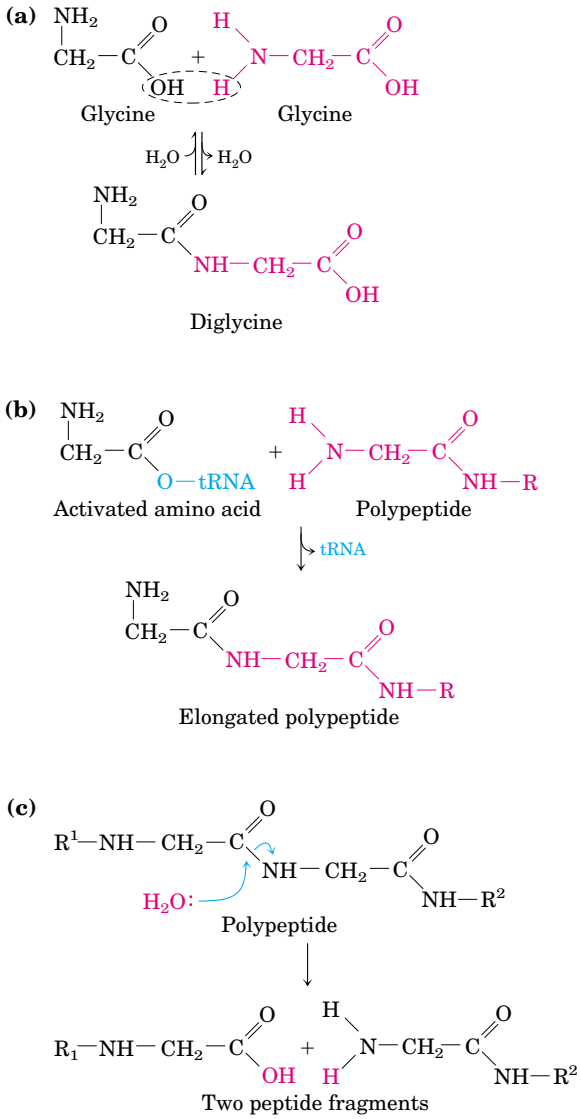


figure 3–22

Condensation and hydrolysis. Shown here are formation and hydrolysis of a peptide bond. **(a)** Removal of the elements of water from two molecules of the amino acid glycine produces a peptide bond, but because —OH is not a good leaving group, this reaction is unfavorable. **(b)** In cells, amino acids are activated prior to polymerization by attachment of a better leaving group than —OH, a short RNA (transfer RNA or tRNA) that forms an oxygen ester with the α-carboxyl group. **(c)** The hydrolysis of a peptide bond (shown here in a polypeptide) is essentially the reverse of the reaction in **(a)**: H_2O makes a nucleophilic attack on the carbonyl carbon, displacing the nitrogen of the α-amino group.

Biopolymers Are Formed by Condensations

The monomeric subunits that make up proteins, nucleic acids, and polysaccharides are joined by nucleophilic displacement reactions that replace a good leaving group. For example, the joining of two amino acid molecules to form a dipeptide could occur by the simple mechanism in Figure 3–22a. However, —OH is a poor leaving group, and the reaction by this mechanism is not efficient. Cells solve this problem by first attaching a better leaving group, a small RNA molecule (transfer RNA, about 75 nucleotides long), in ester linkage to the α-carboxyl group of the amino acid. This activates the carboxyl group for condensation with the α-amino group of another amino acid (Fig. 3–22b). Similar strategies are employed in the biosynthesis of nucleic acids and polysaccharides.

Macromolecules can be broken down by hydrolysis reactions, in which H_2O is the attacking nucleophile, displacing a monomeric subunit or a smaller polymer fragment (Fig. 3–22c). Enzymes that catalyze hydrolysis of biopolymers (hydrolases) are essential in the digestive process and serve also to regulate the level of such critical macromolecules as messenger RNA.

Macromolecules and Their Monomeric Subunits

Many biological molecules are macromolecules, polymers of high molecular weight assembled from relatively simple precursors. Polysaccharides, proteins, and nucleic acids are produced by the polymerization of relatively small compounds with molecular weights of 500 or less. The total 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, which are constructed of about 70 different proteins and several different RNA molecules.

Macromolecules Are the Major Constituents of Cells

Table 3–5 shows the major classes of biomolecules in the bacterium *Escherichia coli*. Water is the most abundant compound in *E. coli* and in virtually all other cells and organisms. Nearly all of the solid matter in cells is organic and is present in four forms: proteins, nucleic acids, polysaccharides, and lipids. Inorganic salts and mineral elements constitute a very small fraction of the total dry weight.

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 per-

table 3–5

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

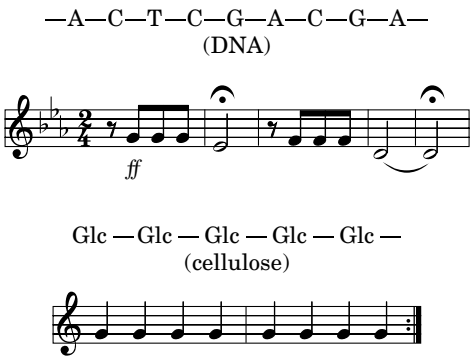


figure 3–23

Informational and structural macromolecules. A, T, C, and G represent the four subunits of DNA, and glucose (Glc) is the repeating subunit of cellulose. The variety of sequences that can be made from the four subunits of DNA is limitless, as is the number of melodies possible with a few musical notes. Cellulose, a polymer of one subunit type, is information-poor and monotonous.

haps 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 roles in macromolecular complexes. The **polysaccharides,** polymers of simple sugars such as glucose, have two major functions: as energy-yielding fuel stores and as extracellular structural elements. 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. Proteins, nucleotides, polysaccharides, and lipids are synthesized in condensation reactions. In the first three categories, the number of monomeric subunits in the polymer is very large. Proteins have molecular weights in the range of 5,000 to over 1 million; nucleic acids have molecular weights ranging up to several billion; and polysaccharides, such as starch, have molecular weights into the millions. Individual lipid molecules are much smaller (M_r 750 to 1,500) and are not classified as macromolecules. However, when large numbers of lipid molecules associate noncovalently, very large structures result. Cellular membranes are built of enormous aggregates containing millions of lipid molecules.

Macromolecules Are Composed of Monomeric Subunits

Although living organisms contain a very large number of different proteins and different nucleic acids, a fundamental simplicity underlies their structure (Chapter 1). The simple monomeric units from which all proteins and all nucleic acids are constructed are few in number and identical in all living species. Proteins and nucleic acids are **informational macromolecules:** each protein and each nucleic acid has a characteristic information-rich subunit sequence (Fig. 3–23).

Polysaccharides with only a single kind of subunit, or with two different alternating units, are not informational molecules in the same sense as are proteins and nucleic acids. However, short sugar polymers made up of six or more different kinds of sugars connected in branched chains have the structural and stereochemical variety to carry information recognizable by other macromolecules.

Monomeric Subunits Have Simple Structures

Figure 3–24 shows the structures of some of the monomeric units of the large biomolecules, arranged in families. Twenty different amino acids are found in proteins; all have an amino group (an imino group in the case of proline) and a carboxyl group attached to the same carbon atom, designated the α carbon. These α-amino acids differ from each other only in their side chains (Fig. 3–24a).

The recurring structural units of all nucleic acids are eight different nucleotides; four kinds of nucleotides are the structural units of DNA, and four others are the units of RNA (Fig. 3–24b). Each nucleotide is made up of three components: a nitrogenous organic base and a five-carbon sugar (which combined are called a nucleoside), to which is attached a phosphate group. The eight different nucleotides of DNA and RNA are built from five different organic bases and two different sugars.

Lipids also are constructed from relatively few kinds of compounds. Most lipid molecules contain one or more long-chain fatty acids, of which palmitic acid and oleic acid are parent compounds (Fig. 3–24c). Many lipids also contain an alcohol, such as glycerol, and some contain phosphate.

The most abundant polysaccharides in nature, starch and cellulose, consist of repeating units of D-glucose (Fig. 3–24d). Other polysaccharides are composed of a variety of sugar molecules derived from glucose.

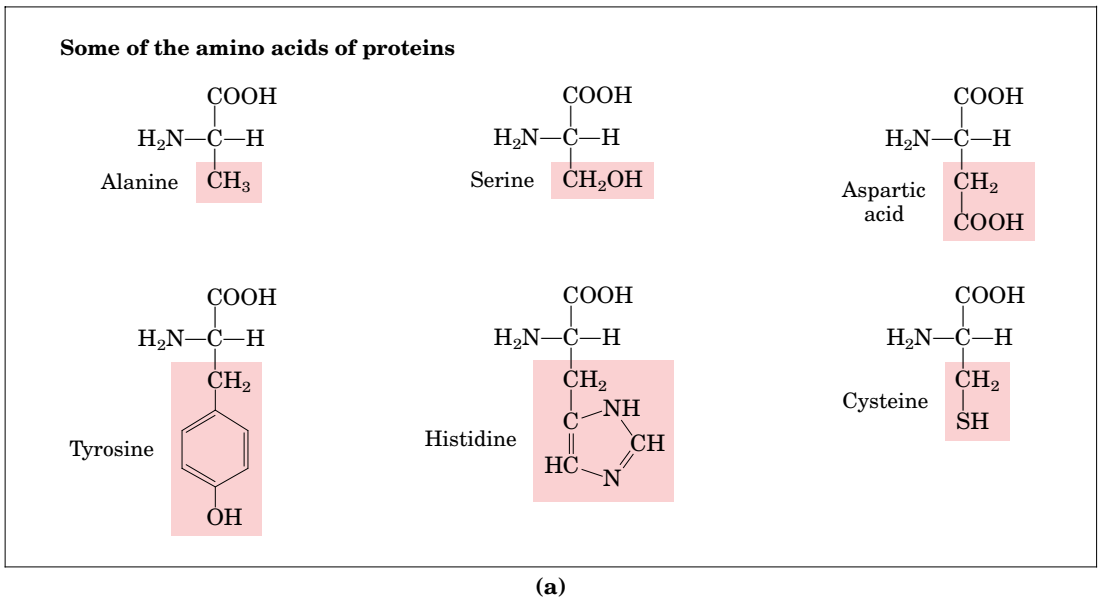


figure 3–24

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 red); **(b)** the five nitrogenous bases, two five-carbon sugars, and phosphoric acid from which all nucleic acids are built; **(c)** five components of many 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. All compounds are shown in their nonionized form.

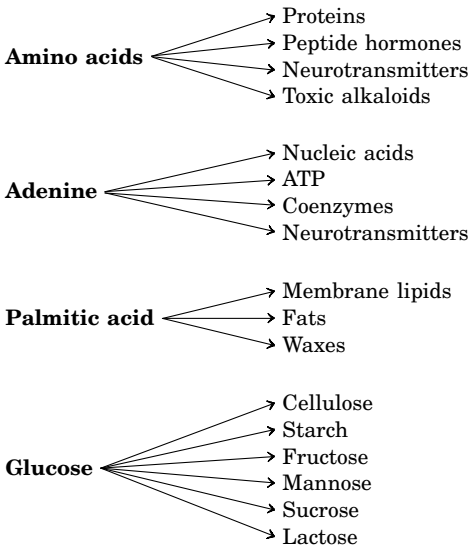
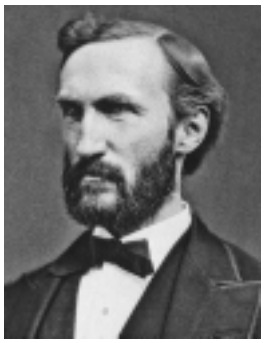


figure 3–25
Each compound in Figure 3–24 is a precursor of many other kinds of biomolecules.



J. Willard Gibbs
1839–1903

Thus, only three dozen organic compounds are the parents of most biomolecules. Each of the compounds in Figure 3–24 has multiple functions in living organisms (Fig. 3–25). Amino acids are not only the monomeric subunits of proteins; some also act as neurotransmitters and as precursors of hormones and toxins. The nitrogenous base adenine serves both as a subunit of nucleic acids and ATP and as a neurotransmitter. Fatty acids serve as components of membrane lipids and fuel-storage fats, and also as precursors of a group of potent signaling molecules, the eicosanoids. D-Glucose is the monomeric subunit of starch and cellulose and is also the precursor of other sugars such as D-mannose and sucrose.

Subunit Condensation Creates Order and Requires Energy

It is extremely improbable that amino acids in a mixture would spontaneously condense into a protein molecule with a unique sequence. This would represent increased order in a population of molecules; but according to the second law of thermodynamics (Chapter 14), the tendency in nature is toward ever-greater disorder in the universe. To bring about the synthesis of macromolecules from their monomeric units, free energy must be supplied to the system, in this case the cell (Chapter 1).

The randomness of the components of a chemical system is expressed as **entropy, *S*.** Any change in randomness of the system is expressed as entropy change, ΔS, which 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 (p. 65); 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 ΔH, reflecting the kinds and numbers of chemical bonds and noncovalent interactions broken and formed, and ΔS, the change in the system's randomness:

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

Recall from Chapter 1 that a process tends to occur spontaneously only if ΔG is negative. Yet cells depend on many molecules, such as proteins and nucleic acids, for which the free energy of formation is positive: they are less stable and more highly ordered than a mixture of their monomeric components. To overcome the free-energy deficit of thermodynamically unfavorable (endergonic) reactions, cells couple these reactions 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 connecting phosphate groups (represented as Ⓟ) in ATP:

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 strategy, cells are able to synthesize and maintain the information-rich polymers (DNA, RNA, and protein) essential to life.

Cells Have a Structural Hierarchy

The monomeric units in Figure 3–24 are much smaller than biological macromolecules. A molecule of alanine is less than 0.5 nm long. Hemoglobin, the oxygen-carrying protein of erythrocytes, consists of nearly 600 amino

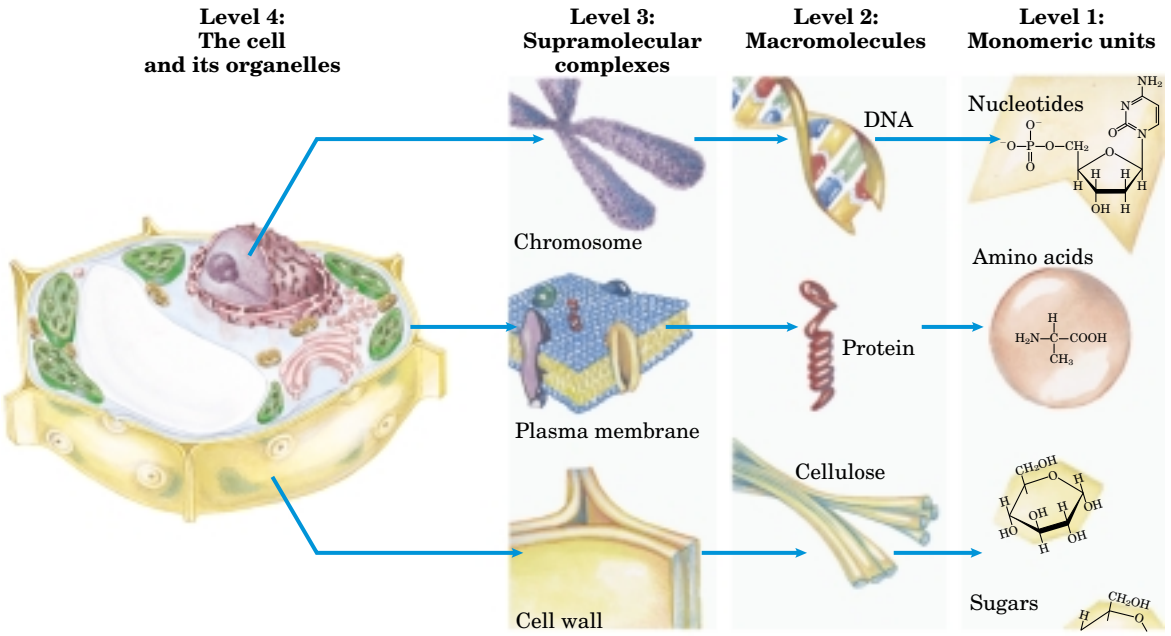
acid units covalently linked into four long chains, which are folded into globular shapes and associated in a tetrameric structure with a diameter of 5.5 nm. Protein molecules in turn are small compared with ribosomes (about 20 nm in diameter), which are in turn much smaller than organelles such as mitochondria, typically 1,000 nm in diameter. It is a long jump from simple biomolecules to cellular structures that can be seen with the light microscope. Figure 3–26 illustrates the structural hierarchy in cellular organization.

In proteins, nucleic acids, and polysaccharides, the individual monomeric subunits are joined by covalent bonds. In supramolecular complexes, however, macromolecules are held together by noncovalent interactions—much weaker, individually, than covalent bonds. Among these 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 3–3). The nature of these noncovalent interactions is described in the next chapter. The large numbers of weak interactions between macromolecules in supramolecular complexes stabilize these noncovalent assemblies, producing their unique "native" structures.

Although the monomeric subunits of macromolecules are much smaller than cells and organelles, they influence the shape and function of these much larger structures. In sickle-cell anemia, a hereditary human disorder, the hemoglobin molecule is defective: valine occurs in two of the four polypeptide chains (the two β chains) of hemoglobin at a position normally occupied by glutamic acid. This single difference in the sequence of the 146 amino acids of the β chain affects only a tiny portion of the hemoglobin molecule, yet, as explained in Chapter 7, it causes the molecules to form large aggregates within the erythrocytes, which become deformed (sickled) and functionally abnormal.

figure 3–26

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—DNA and many different proteins. Each type of macromolecule is constructed from simple subunits—DNA from deoxyribonucleotides, for example.



Prebiotic Evolution

The finding that all biological macromolecules in all organisms are made from the same three dozen subunits has provided strong evidence that modern organisms are descended from a single primordial cell line whose fundamental chemistry would be recognizable even today. Furthermore, several billion years of adaptive selection have refined cellular systems to take maximum advantage of the chemical and physical properties of these molecular raw materials for carrying out the basic energy-transforming and self-replicating features of a living cell.

Biomolecules First Arose by Chemical Evolution

We come now to a puzzle. Apart from their occurrence in living organisms, organic compounds, including the basic biomolecules such as amino acids and carbohydrates, occur 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 the earth, postulating that the atmosphere then 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 millenia 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. For many years, Oparin's views remained speculative and appeared untestable, until a surprising experiment was conducted using simple equipment on a desktop.

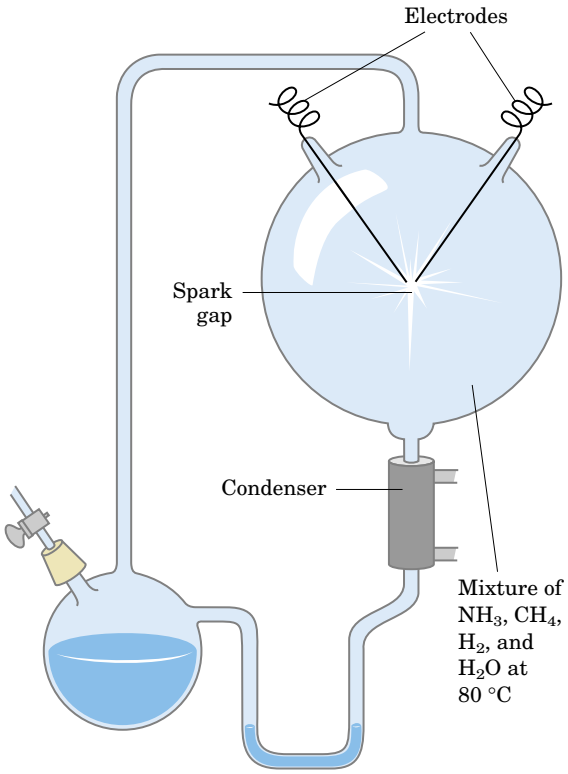


figure 3–27
Spark-discharge apparatus of the type used by Miller and Urey in experiments demonstrating abiotic formation of organic compounds under primitive atmospheric conditions. After subjection of the gaseous contents of the system to electrical sparks, products were collected by condensation. Biomolecules such as amino acids were among the products (see Table 3–6).

Chemical Evolution Can Be Simulated in the Laboratory

A 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. 3–27). 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.

Several developments have allowed more refined studies of the type pioneered by Miller and Urey and have yielded strong evidence that a wide variety of biomolecules, including polypeptides and short polymers of nucleosides, could have been produced spontaneously from simple starting materials probably present on the earth at the time life arose.

Modern extensions of the Miller experiments have employed "atmospheres" that include CO_2 and CO, with N_2 and H_2, and much-improved technology for identifying small quantities of products. The formation of hundreds of organic compounds has been demonstrated (Table 3–6). These compounds include common amino acids, a variety of mono-, di-, and tri-

table 3–6

Some Products Formed under Prebiotic Conditions

Carboxylic acids	Nucleic acid bases	Amino acids	Sugars
Formic acid	Adenine	Glycine	Straight and branched pentoses and hexoses
Acetic acid	Guanine	Alanine	
Propionic acid	Xanthine	α-Aminobutyric acid	
Straight and branched fatty acids (C_4–C_{10})	Hypoxanthine	Valine	
Glycolic acid	Cytosine	Leucine	
Lactic acid	Uracil	Isoleucine	
Succinic acid		Proline	
		Aspartic acid	
		Glutamic acid	
		Serine	
		Threonine	

Source: From Miller, S.L. (1987) Which organic compounds could have occurred on the prebiotic earth? *Cold Spring Harb. Symp. Quant. Biol.* **52,** 17–27.

carboxylic acids, fatty acids, adenine, and formaldehyde. Under certain conditions, formaldehyde polymerizes to form sugars containing three, four, five, and six carbons. In addition to the many small molecules that form in these experiments, polymers of nucleotides (nucleic acids) and of amino acids (proteins) also form. Some products of the self-condensation of HCN are effective promoters of such polymerization reactions, and inorganic ions present in the earth's crust (Cu^{2+}, Ni^{2+}, and Zn^{2+}) also enhance the rate of polymerization. The sources of energy that are effective in bringing about the formation of these compounds include heat, visible and ultraviolet (UV) light, x rays, gamma radiation, ultrasound and shock waves, and bombardment with alpha and beta particles.

In short, laboratory experiments on the spontaneous formation of biomolecules under prebiotic conditions have provided good evidence that many of the chemical components of living cells, including polypeptides and RNA-like molecules, can form under these conditions. Short polymers of RNA can act as catalysts in biologically significant reactions (Chapter 26), and RNA probably played a crucial role in prebiotic evolution, both as catalyst and as information repository.

If life evolved on Earth by this chemical evolution process, it is likely that life arose also on suitable planets of other solar systems. Many prebiotic compounds such as HCN, formic acid, and cyanoacetylene have been found in comets, in the atmospheres of Jupiter, Saturn, and Titan (a moon of Saturn), and in the dust clouds of interstellar space. Analysis of the Murchison meteorite, which fell to Earth in 1969, revealed the presence of amino acids, hydroxy acids, purines, and pyrimidines. It is therefore conceivable that the organic precursors for the evolution of life on Earth originated elsewhere in the solar system.

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 have the ability to 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?

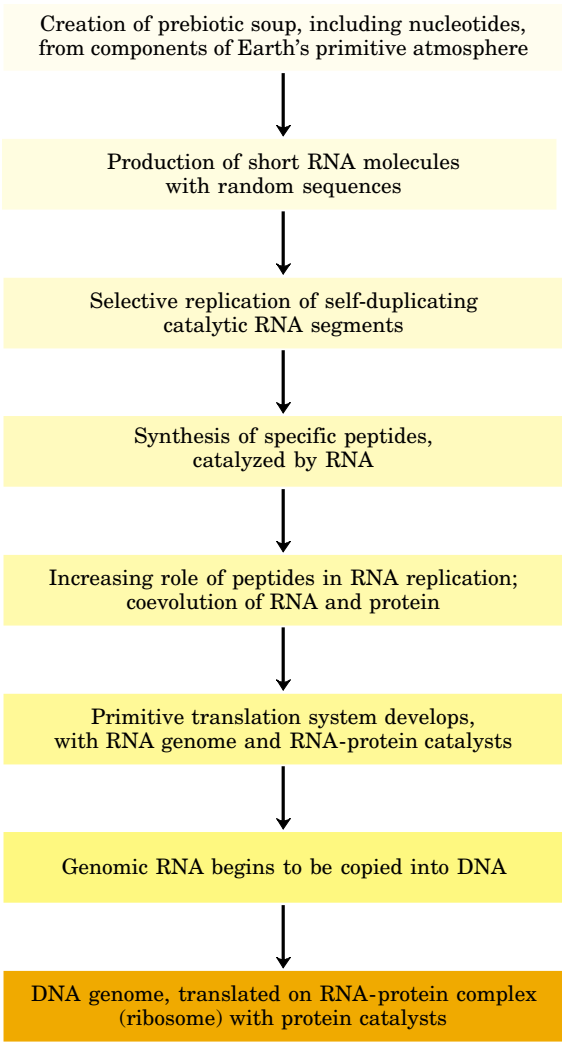


figure 3–28
One possible "RNA world" scenario, showing the transition from the prebiotic RNA world (shades of yellow) to the biotic DNA world (orange).

The answer may be: neither. The discovery that RNA molecules can act as catalysts in their own formation suggests that RNA or a similar molecule may have been the first gene *and* the first catalyst. According to this scenario (Fig. 3–28), one of the earliest stages of biological evolution was the chance formation, in the primordial soup, of an RNA molecule that had the ability to 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. 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. 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.

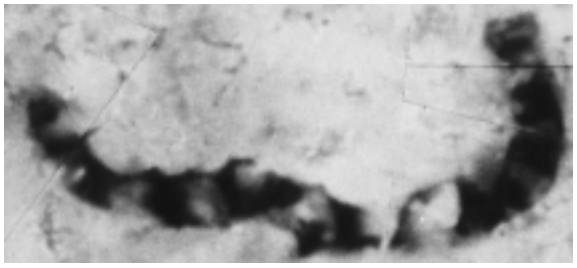
This "RNA world" hypothesis is plausible but by no means universally accepted. The hypothesis does make testable predictions, and to the extent that experimental tests are possible within finite times, the hypothesis will be tested and refined.

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

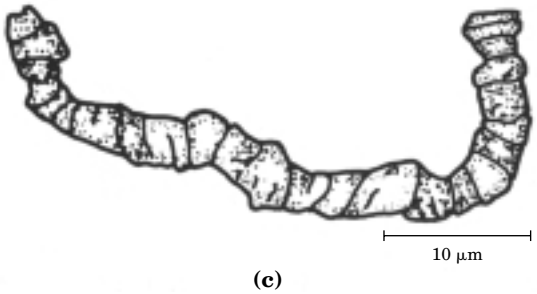
Earth was formed about 4.5 billion years ago, and the first evidence of life dates to more than 3.5 billion years ago. An international group of scientists showed in 1980 that certain ancient rock formations (stromatolites) in Western Australia contained fossils of primitive microorganisms (Fig. 3–29). 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 more and more effective self-reproducing catalysts. The process of biological evolution had begun.

Organisms developed mechanisms to harness the energy of sunlight through photosynthesis, to make sugars and other organic molecules from carbon dioxide, and to incorporate molecular nitrogen from the atmosphere into nitrogenous biomolecules such as amino acids. By developing their own capacities to synthesize biomolecules, cells became independent of the random processes by which such compounds had first appeared on Earth. As evolution proceeded, organisms began to interact and to derive mutual benefits from each other's products, forming increasingly complex ecological systems.

(a)



(b)



(c)

figure 3–29

Ancient reefs in Australia contain fossil evidence of microbial life in the sea of 3.5 billion years ago. Bits of sand and limestone became trapped in the sticky extracellular coats of cyanobacteria, gradually building up these stromatolites found in Hamelin Bay, Western Australia **(a).** Microscopic examination of thin sections of such ancient rock reveals microfossils of filamentous bacteria **(b),** interpreted as shown in the drawing **(c).**

summary

Most of the dry weight of living organisms consists of organic compounds, molecules containing covalently bonded carbon backbones to which other carbon, hydrogen, oxygen, or nitrogen atoms may be attached. The different kinds of functional groups attached to the backbone determine the chemical properties of the molecule. Organic biomolecules also have characteristic shapes (configurations and conformations) in three dimensions. Many biomolecules occur in asymmetric or chiral forms called enantiomers, stereoisomers that are nonsuperimposable mirror images of each other. Usually, only one of a pair of enantiomers has biological activity.

The strength of covalent chemical bonds, measured in joules, depends on the electronegativities and sizes of the atoms that share electrons. The enthalpy change (ΔH) for a chemical reaction reflects the numbers and kinds of bonds made and broken. For endothermic reactions, ΔH is positive; for exothermic reactions, negative. The many different chemical reactions that occur within a cell fall into five general categories: oxidation-reduction reactions, breakage or formation of carbon–carbon bonds, rearrangements of the bonds around carbon atoms, group transfers, and condensations.

Most of the organic matter in living cells consists of macromolecules: nucleic acids, proteins, and polysaccharides. Lipid molecules, another important component of cells, are small molecules that form large aggregates. Biological macromolecules are composed of small, covalently linked monomeric subunits of relatively few kinds. Proteins are polymers of 20 different kinds of amino acids, nucleic acids are polymers of different nucleotide units (four kinds in DNA, four in RNA), and polysaccharides are polymers of recurring sugar units. Nucleic acids and proteins are informational macromolecules; the characteristic sequences of their subunits constitute the genetic individuality of a species. Simple polysaccharides act as structural components; some complex polysaccharides are informational macromolecules, their specific sequence and stereochemistry allowing them to be recognized with high specificity by other molecules.

There is a structural hierarchy in the molecular organization of cells. Cells contain organelles, such as nuclei, mitochondria, and chloroplasts, which in turn contain supramolecular complexes, such as membranes and ribosomes, and these consist in turn of clusters of macromolecules bound together by many relatively weak, noncovalent forces. Macromolecules consist of covalently linked subunits. The formation of macromolecules from simple subunits creates order (decreases entropy); this synthesis requires energy and therefore must be coupled to exergonic reactions.

The small biomolecules such as amino acids and sugars probably first arose spontaneously from atmospheric gases and water under the influence of electrical energy (lightning) during the early history of the earth. The same process can be simulated in the laboratory. The monomeric components of cellular macromolecules appear to have been selected early in biological evolution. These subunit molecules are relatively few in number, yet very versatile; evolution has combined small biomolecules to yield macromolecules of immense diversity. The first macromolecules may have been RNA-like molecules capable of catalyzing their own replication. Later in evolution, DNA took over the function of storing genetic information, proteins became the cellular catalysts, and RNA mediated between DNA and protein, allowing the expression of genetic information as proteins.

further reading

General

Frausto da Silva, J.J.R. & Williams, R.J.P. (1994) *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life,* Clarendon Press, Oxford.

An excellent, highly readable text on the role of inorganic elements in biochemistry. Clear diagrams, good references.

Frieden, E. (1972) The chemical elements of life. *Sci. Am.* **227** (July), 52–61.

The Molecules of Life. (1985) *Sci. Am.* **253** (October).

An entire issue devoted to the structure and function of biomolecules. It includes articles on DNA, RNA, and proteins, and their subunits.

Chemistry and Stereochemistry

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.

Hegstrom, R.A. & Kondepudi, D.K. (1990) The handedness of the universe. *Sci. Am.* **262** (January), 108–115.

Stereochemistry and the asymmetry of biomolecules, viewed in the context of the universe.

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

An excellent, comprehensive introduction to chemistry.

Loudon, M. (1995) *Organic Chemistry,* 3rd edn, The Benjamin/Cummings Publishing Company, Menlo Park, CA.

This and the following two books provide details on stereochemistry and the chemical reactivity of functional groups. All are excellent textbooks.

Morrison, R.T. & Boyd, R.N. (1999) *Organic Chemistry,* 7th edn, Allyn & Bacon, Inc., Boston, MA.

Streitwieser, A., Jr., Heathcock, C.H., & Kosower, E.M. (1998) *Introduction to Organic Chemistry,* 4th edn, revised printing, Prentice-Hall, Upper Saddle River, NJ.

Wagner, G. (1997) An account of NMR in structural biology. *Nature Struct. Biol.,* NMR Suppl. (October), 841–844.

A short, clear account of the development of NMR as a biochemical tool, current applications, and future prospects.

Prebiotic Evolution

Cavalier-Smith, T. (1987) The origin of cells: a symbiosis between genes, catalysts, and membranes. *Cold Spring Harb. Symp. Quant. Biol.* **52,** 805–824.

Darnell, J.E. & Doolittle, W.F. (1986) Speculations on the early course of evolution. *Proc. Natl. Acad. Sci. USA* **83,** 1271–1275.

A clear statement of the RNA world scenario.

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.

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.

Ferris, J.P. (1984) The chemistry of life's origin. *Chem. Eng. News* **62,** 21–35.

A short, clear description of the experimental evidence for the synthesis of biomolecules under prebiotic conditions.

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.

Hager, A.J., Pollard, J.D., & Szostak, J.W. (1996) Ribozymes: aiming at RNA replication and protein synthesis. *Chem. Biol.* **3,** 717–725.

A short review of studies on RNA catalysis of RNA synthesis and amide bond synthesis, and the relevance of these findings to the RNA world scenario.

Hirao, I. & Ellington, A.D. (1995) Re-creating the RNA world. *Curr. Biol.* **5,** 1017–1022.

This and the article by Robertson and Ellington (1997), listed below, describe research aimed at reproducing in the laboratory the evolutionary rise of catalytic RNA.

Horgan, J. (1991) In the beginning . . . *Sci. Am.* **264** (February), 116–125.

A brief, clear statement of current theories on prebiotic evolution.

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 recent 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.

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.

Miller, S.L. & Orgel, L.E. (1974) *The Origins of Life on the Earth,* Prentice-Hall, Inc., Englewood Cliffs, NJ.

The New Age of RNA. (1993) *FASEB J.* **7** (1).

A collection of about 15 short articles related to the RNA world scenario.

Robertson, M.P. & Ellington, A.D. (1997) Ribozymes: red in tooth and claw. *Curr. Biol.* **7,** R376–R379.

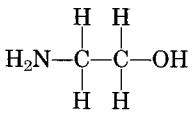
Schopf, J.W. (ed.) (1983) *Earth's Earliest Biosphere,* Princeton University Press, Princeton, NJ.

A comprehensive discussion of geologic history and its relation to the development of life.

problems

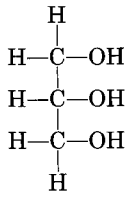
1. Vitamin C: Is the Synthetic Vitamin as Good as the Natural One? A claim sometimes put forth by purveyors of health foods is that vitamins obtained from natural sources are more healthful than those obtained by chemical synthesis. For example, pure L-ascorbic acid (vitamin C) extracted from rose hips 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?

2. Identification of Functional Groups Figures 3–5 and 3–6 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 constituent functional group.



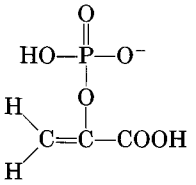
Ethanolamine

(a)



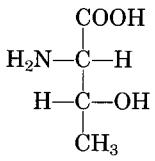
Glycerol

(b)



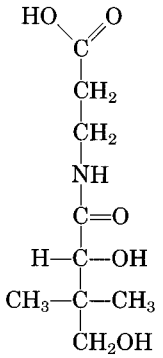
Phosphoenolpyruvic acid, an intermediate in glucose metabolism

(c)



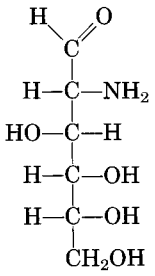
Threonine, an amino acid

(d)



Pantothenic acid, a vitamin

(e)

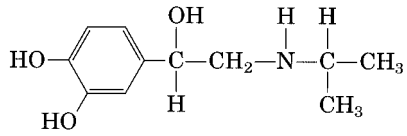


D-Glucosamine

(f)

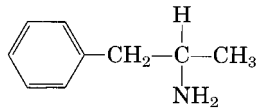
Problem 2

3. 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

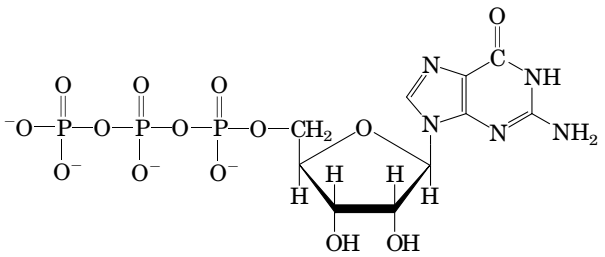
4. Drug Action and Shape of Molecules Some years ago two drug companies marketed a drug under the trade names Dexedrine and Benzedrine. The structure of the drug is shown below.



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.

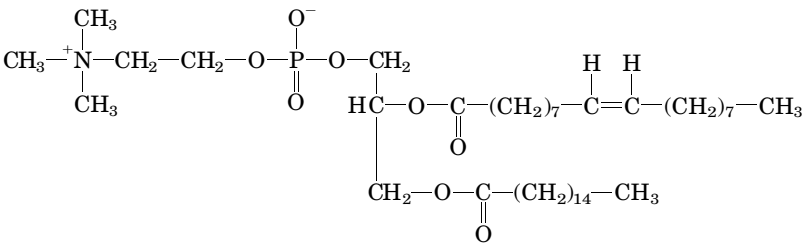
5. Components of Complex Biomolecules Figure 3–24 shows the major components of complex biomolecules. For each of the three important biomolecules below and at right (shown in their ionized forms at physiological pH), identify the constituents.

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

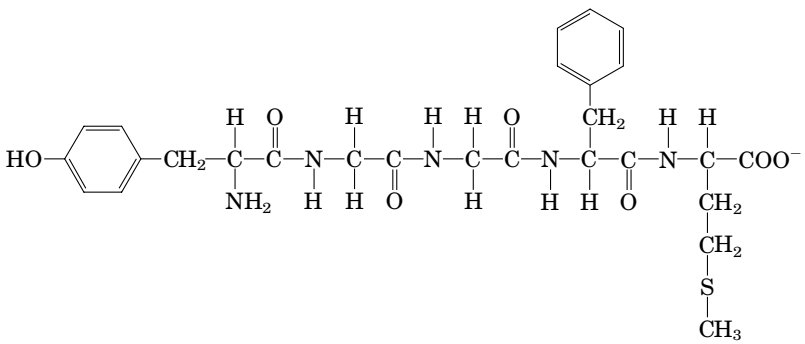


Problem 5

(b) Phosphatidylcholine, a component of many membranes:



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



6. 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 a mass spectrometer, was 90.00 amu. An infrared spectrum 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) does this observation eliminate? Which structures 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 now eliminated? Which structures are consistent with the observation?

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

7. Separating Biomolecules In laboratory biochemistry, it is first necessary to separate the molecule of interest from the other biomolecules in the sample—that is, to *purify* the protein, nucleic acid, carbohydrate, or lipid. Specific purification techniques will be addressed later in the text. However, just by looking at the monomeric subunits from which the larger biomolecules are made, you should have some ideas as to what characteristics of those biomolecules would allow you to separate them one from another.

(a) What characteristics of amino acids and fatty acids would allow them to be easily separated from each other?

(b) How might nucleotides be separated from glucose molecules?

8. 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 written based on the premise of silicon-based life. Is this realistic? What characteristics of silicon make it *less* well adapted to performing as the central organizing element for life? To answer this question, use the information in this chapter about carbon's bonding versatility, and refer to a beginning inorganic chemistry textbook for silicon's bonding properties.