

Figure 12.10 Some simple sugars. (a) linear glucose, (b) cyclic glucose, (c) fructose, and (d) sucrose, a disaccharide formed by condensation of glucose and fructose.

hydroxyl groups for hydrogen atoms. Two of the simplest carbohydrates are the sugars glucose and fructose, both of which have the composition $C_6H_{12}O_6$. Both can exist as straight chains or cyclic structures (Figure 12.10), though the cyclic structures predominate. Glucose and fructose are examples of *monosaccharides*, the mono- prefix indicating they consist of single chains or rings. General names for these compounds are formed from the Greek prefix corresponding to the number of carbons and the suffix *-ose*. Thus fructose and glucose are pentose sugars and ribulose (a building block of nucleic acids) is a hexose sugar. Two monosaccharide units may be linked together by elimination of H_2O to form a disaccharide (another example of a *condensation* reaction). Sucrose, or common table sugar, is the condensation product of glucose and fructose.

Molecules consisting of ten or more monosaccharide units are called *polysaccharides*. Among the biologically most important polysaccharides are *cellulose* and *starch*. Cellulose, the basic structural material of plants, has the general formula of $(C_6H_{10}O_5)_n$ and consists of long (i.e., $n \geq 10,000$) chains of glucose units. The chains are cross-linked to each other by hydrogen bonds, building up fibrils. In cell walls of vascular plants these fibrils are interwoven or cemented by

other polysaccharides called hemicellulose. Cellulose is an example of a *homogenous polysaccharide*, that is, one that is formed by linkage of a single kind of monosaccharide. *Chitin*, the material forming hard structures in arthropods, molluscs, and some fungi and algae, is also a homogenous polysaccharide. It is related to cellulose by replacing one of the hydroxyl groups with an amido group. Starch, which serves to store energy in plants, is also a $(C_6H_{10}O_5)_n$ polysaccharide in its simplest form. Amylose, for example, which represents about 25% of starch, consists of long, unbranched chains of (cyclo-) glucose. Starches, however, also include heterogeneous polysaccharides, that is, polysaccharides containing more than one kind of monosaccharide unit. Water-soluble starches consist of relatively short chains ($n \approx 25$); insoluble starches are typically longer, up to 500,000u. Glycogen, a polysaccharide used for energy storage in animals, is closely related to starch. From a geochemical perspective, an important difference between cellulose and starch is that the former is much more stable and less readily metabolized. Although organisms generally store energy in the form of complex carbohydrates and lipids, these are always first converted to back to glucose before oxidation releases this energy.

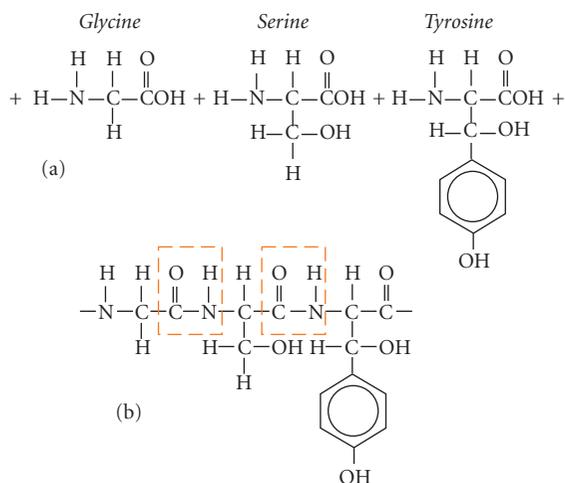


Figure 12.11 (a) Three of the 20 common amino acids that can combine to form proteins. (b) Peptide segment consisting of glycine, serine and tyrosine bound by peptide linkages (dashed boxes). Peptide linkage occurs between amine and carboxyl groups with the elimination of H₂O.

12.3.4.2 Nitrogen-bearing organic compounds: proteins, nucleotides, and nucleic acids

Amino acids are the basic building blocks of proteins. There are 20 common amino acids, the essential characteristic of which is the presence of both an amine and a carboxylic group (Figure 12.7). The simplest amino acids are *glycine*, which consists of the amino acid group with a hydrogen at the free position (Figure 12.11), and *alanine*, which has CH₃ at the free position. Amino acids may be characterized as neutral, acidic, or basic. *Acidic* amino acids have an additional carboxylic group, which acts as a proton donor. *Basic* amino acids, such as lysine (COOHCH((CH₂)₄NH₂) NH₂), have an additional amine group, which can act as a proton acceptor. Neutral amino acids, such as glycine and alanine, have equal numbers of carboxylic and amine groups.

Proteins are formed by condensation of many amino acid units into polymers called peptides. The simplest proteins consist of 40 amino acid units; the most complex ones consist of more than 8000 units. With 20 basic building blocks, the possible combinations are virtually limitless, making the diversity of life possible. The condensation reaction

forming peptides consists of linking the carboxylic group of one amino acid to the amine group of another with the elimination of water, as illustrated in Figure 12.11. This bond is referred to as a peptide linkage. The biosynthesis of proteins is performed in ribosomes by RNA molecules. The genetic information contained in DNA is essentially a set of blueprints for protein synthesis.

There are an immense variety of proteins, and they play a wide variety of roles in life. Proteins such as *collagen* (bone) and *keratin* (hair, claws) are the essential structural and connective materials of higher animals. It is contraction of proteins in muscles that provide movement. Enzymes, which are cell's catalysts, are often proteins, as are antibodies, which play an essential role in the immune system. Proteins also act to store and transport various elements and compounds; hemoglobin is a good example of such a protein. Hormone proteins serve as messengers and regulators.

This variety of function results from primary, secondary, and tertiary structures. The primary structure of proteins depends both on the kind of amino acid units composing them and on the order in which these units occur. These primary structures may then be folded. The folds are locked in by hydrogen bonds between adjacent parts of the chain (secondary structures). Other proteins are twisted into α -helix structures. Folding of the α -helix results in tertiary structures. All these structures contribute to the biological function of the protein.

Nucleotides are based on pyrimidine or purine groups (Figure 12.12). The nucleotides adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADP), illustrated in Figure 12.12, play key roles in both the creation and storage of chemical energy (photosynthesis) and its transfer and release (respiration) in organisms. Another nucleotide, nicotinamide adenine dinucleotide (NAD), plays an important role in respiration. ATP can be formed from ADP (adenosine diphosphate) by the addition of an inorganic phosphate ion. This process, called *phosphorylation*, involves a free energy change (ΔG) of about +40 kJ/mol. That energy is readily liberated on demand by the reverse reaction. Thus ATP serves as a general carrier of free energy within cells. NAD and NADP

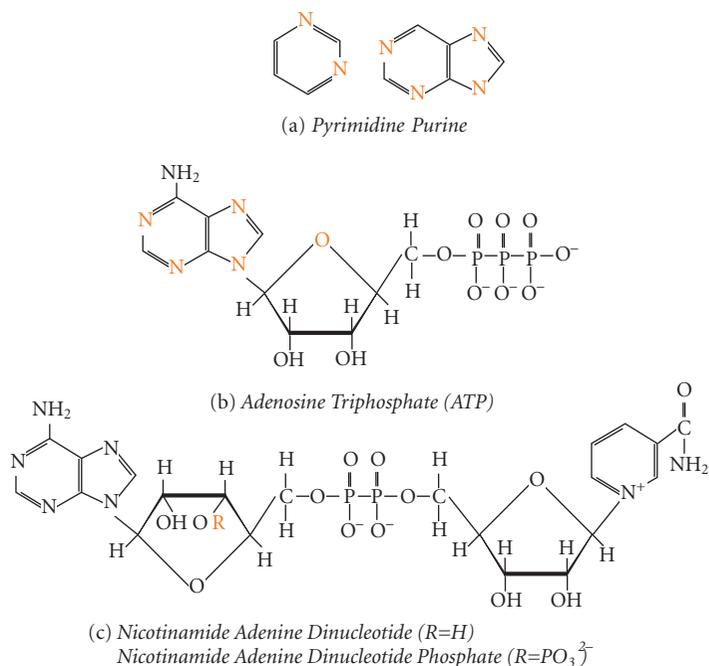


Figure 12.12 (a) The structure of pyrimidine and purine groups, essential components of nucleotides. (b) Structure of the nucleotide adenosine triphosphate (ATP). (c) Structure of NAD or nicotinamide adenine dinucleotide (when the radical labelled R is H) and NADP or nicotinamide adenine dinucleotide phosphate (when R is phosphate).

and their reduced equivalents (NADH and NADPH) serve as redox couples and as transport agents of reduced hydrogen.

Nucleic acids are related to proteins in the sense that they are nitrogen-containing polymers built from a variety of fundamental groups. The amine, phosphate (PO_4^3), and pyridinyl (Figure 12.8) groups are among the essential ingredients of nucleic acids. Unlike proteins, the carboxyl group is generally not present. The nucleic acids DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) contain the genetic code and control protein synthesis within the cell. DNA consists of two backbone strands of a polymer made up of phosphate and the pentose sugar β -D-ribofuranose connected to each other by pairs of the four nucleotides, cytosine, thymine, adenine, and guanine. The genetic information is encoded in the sequence of pairs.

12.3.4.3 Lipids

Unlike carbohydrates and proteins, *lipids* are defined not by their composition and structure, but by their behavior: lipids are those

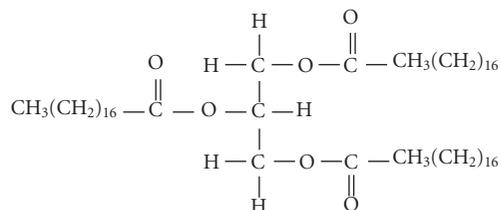


Figure 12.13 A triglyceride fat formed from the alcohol glycerol and three molecules of stearic acid.

organic substances that are water-insoluble but are soluble in organic solvents such as chloroform, toluene, acetone, and ether. Lipids include fats, oils, waxes, steroids, and other compounds. Fats and oils are generally *triglycerides*, which are esters of three *fatty acids* and *glycerol*, an alcohol (Figure 12.13). Fatty acids are straight-chained (aliphatic) carboxylic acids (i.e., an alkane with a carboxyl group at one end). They typically range in length from C_{12} to C_{36} . Because they are generally formed by successive additions of

acetyl (C_2) units, fatty acids have predominantly even numbers of carbon atoms. Unsaturated fats, such as *oleic acid* ($CH_3(CH_2)_7CH=CH(CH_2)_7CH(O)OH$), predominate in plants, whereas saturated fats, such as *stearic acid* ($CH_3(CH_2)_{16}CH(O)OH$), predominate in animals. In *phospholipids*, one of the fatty acids in the triglyceride is replaced by a phosphate unit, which is in turn often linked to a nitrogen base. Many *glycolipids*, which are combinations of lipids and carbohydrates, are triglycerides in which one of the fatty acids is replaced by a sugar. Ether lipids are glycerides formed from straight-chained alcohols, called *n-alkanols* or fatty alcohols, rather than fatty acids.

Waxes are a mixture of many constituents. Among the most important are wax esters, which are esters of straight-chained fatty acids and fatty alcohols, both of which have chain lengths generally in the range of C_{24} to C_{28} . The fatty alcohols also have predominantly even number of carbon atoms because they are synthesized from fatty acids. Plant waxes also contain long, straight-chained hydrocarbons (C_{23} to C_{33}). These hydrocarbons typically have odd numbers of carbons because they are formed by *decarboxylation** of fatty acids. *Cutin*, which forms protective coatings on plants, is a polymerized hydroxy fatty acid (commonly C_{16} or C_{18}).

Another important class of lipids is the *terpenoids*. Terpenoids display a great diversity of structure, but the basic unit of all terpenoids is the isoprene unit, a branched, 5-carbon chain with a methyl group attached to the second carbon atom (Figure 12.14a). Terpenoids are named on the basis of the number of isoprene units present: monoterpenoids have 2, sesquiterpenoids have 3, diterpenoids have 4, triterpenoids have 6, tetraterpenoids have 8. Molecules consisting of more than 8 isoprene units are termed polyterpenoids. Terpenoids may be cyclic (e.g., menthol, Figure 12.14b) or acyclic, saturated or unsaturated. Among other things, terpenoids serve as pheromones (scents, attractants), hormones, antibiotics, resins, and vegetable and animal oils. The noncyclic terpenoids, often referred to as acyclic *isoprenoids*, are important from a geochemical

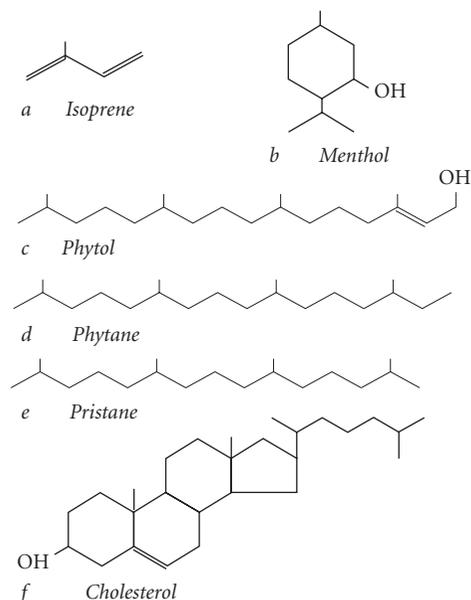


Figure 12.14 Terpenoids. (a) Isoprene, the building block of all terpenoids. (b) Menthol, a simple cyclic monoterpene. (c) Phytol, an acyclic diterpene. (d) Phytane, an acyclic diterpene derived from phytol. (e) Pristane, one of the possible products of diagenesis involving phytane. (f) Cholesterol, illustrating the ring system shared by all steroids. The carbon atoms at the apices of the phenols and the hydrogens bound to them are not shown.

perspective because they are common components of sedimentary organic matter. Furthermore, they are also found in petroleum, apparently having survived diagenesis. Many isoprenoids in sedimentary organic matter are derivatives of *phytol*, a diterpene that forms part of chlorophyll-a (Figure 12.15), a key compound in photosynthesis. Under strongly reducing conditions during diagenesis, phytol (Figure 12.14c) is converted to *phytane* (Figure 12.14d), whose backbone is a 16-carbon chain, through the loss of the OH functional group and hydrogenization (loss of the C-double bond). Under less reducing conditions, phytol is converted to *pristane* (Figure 12.14e), which has one less carbon in its backbone. Another important isoprenoid is squalene, a triterpene ($C_{30}H_{50}$). Squalene is abundant in both plants and animals. Among

* *Decarboxylation* is a process whereby a CO_2 molecule is lost from an organic compound. *Carboxylation* is the addition of a CO_2 molecule to an organic molecule.

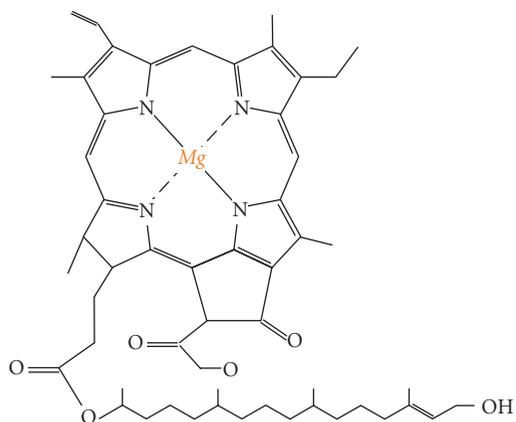


Figure 12.15 Structure of chlorophyll-a. The linear chain is a phytol unit.

other things, it is the precursor of a biologically important class of compounds called *steroids*. *Cholesterol* (Figure 12.14f) is a triterpenoid that occurs in both plants and animals. Most cholesterol is found in cell membranes and in lipoproteins. Cholesterol also serves as the precursor of other animal steroids.

Like proteins, lipids play a variety of roles in life. Phospholipids are the primary constituents of cell membranes. Per unit weight, lipids release twice as much energy as carbohydrates upon oxidation. Thus fats and oils serve as efficient stores of energy for both plants and animals. Lipids include pigments that are essential in photosynthesis in plants (e.g., chlorophyll) and vitamin A production in mammals (e.g., carotenoids). Waxes such as cutin form protective barriers. Other lipids act as sex pheromones (i.e., providing scent) or hormones, or assist in digestion. From a geochemical perspective, lipids are important because they are thought to play a dominant role in petroleum formation. Furthermore, many oils and pigments are unique to groups of organisms. Some of these, isoprenoids in particular, can survive diagenesis, and can be used in reconstructing the origin of hydrocarbon deposits. Such compounds are called *biomarkers*.

12.3.4.4 Lignin and tannins

Lignin is another important structural material in higher plants. It forms a network

around cellulose to provide structural rigidity in wood. It is second only to cellulose as the most abundant organic molecule in the biosphere. It is a rigid, high molecular weight polyphenol. As such, it is quite stable and resistant to bacterial decomposition. As a result, it is a very important contributor to soil and terrestrial sedimentary organic carbon. Because marine plants are almost exclusively algae, most of which do not produce lignin (red algae, which are often multicellular seaweeds, are the exception), lignin is a much less important contributor to marine organic carbon.

Tannins, whose name derives from their use in tanning leather, are another class of compounds found only in higher plants. They occur predominantly in bark and leaves and function to make the plant less palatable to herbivores. Like lignin, they are high molecular weight (500–3000) polyphenol compounds. Functional groups include carboxyl as well as OH.

12.4 THE CHEMISTRY OF LIFE: IMPORTANT BIOCHEMICAL PROCESSES

Our main concern in this chapter is the role played by organic compounds outside living tissue. Nevertheless, given the remarkable impact of life upon the surface of the Earth and its chemistry, and the interplay and constant exchange between living and non-living organic matter, a very brief survey of the more important biochemical processes is appropriate at this point.

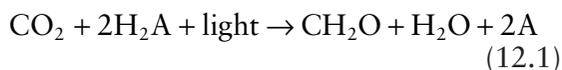
As was mentioned earlier, autotrophs produce the energy they need by synthesis of organic compounds from inorganic ones. Most autotrophs are phototrophic: they use light energy to synthesize organic compounds, a process known as photosynthesis. Some bacteria use chemical energy, derived from the oxidation of H_2S or other reduced species, to synthesize organic compounds in a process called chemosynthesis.

12.4.1 Photosynthesis

Like most biochemical processes, photosynthesis is a complex one, involving many steps catalyzed by a variety of compounds. The details of the photosynthetic processes also

vary somewhat between photosynthetic bacteria and true plants. In true plants, photosynthesis takes place within specialized intracellular organelles called chloroplasts. In prokaryotic bacteria such specialized intracellular units are absent. In these organisms, the site of photosynthesis may be in internal membranes within the cell protoplasm, as in the case of the cyanobacteria, or on the cell's plasma membrane, as in the case of the halophiles.

We can describe photosynthesis with the following reaction:



In photosynthetic eukaryotes and cyanobacteria, A is oxygen, and hence the hydrogen donor is water. However, in anaerobic photosynthesis, carried out by some photosynthetic bacteria (for example the purple sulfur bacteria), A may be sulfur, so that H_2S is the hydrogen donor; alternatively, hydrogen may be taken up directly. In oxygenic photosynthesis, carried out by plants and some photosynthetic bacteria, molecular oxygen is a byproduct of photosynthesis and it is this process that is to a large degree responsible for free oxygen in the atmosphere. Some bacteria can fix CO_2 without liberating oxygen, a process called non-oxygenic photosynthesis. Our brief description will focus on oxygenic photosynthesis.

Oxygenic photosynthesis can be divided into a light and a dark stage. The former involves two separate photoreactions. The first of these, governed by Photosystem II (or PS II) is the photodissociation of water. This reaction can be described as:



Among the enzymes mediating this process is a Mn-bearing protein, which makes up for the liberated electrons by oxidation of Mn. The hydrogen ions and electrons produced by PS II travel along distinct chemical pathways, which in the case of the latter include Fe- and Cu-bearing proteins, to the site of the dark reactions. The energy from the electrons liberated in PS II is used for phosphorylation of ADP to ATP. In Photosystem I (PS I), NADP^+

is reduced to NADPH. In the subsequent dark reactions, this NADPH acts as an electron donor in the reduction of CO_2 , and the energy for this reaction is supplied by ATP.

Energy to drive both PS I and PS II is captured by chlorophyll (Figure 12.15), a member of a class of lipids called *porphyrins*, or some other light-absorbing pigment. There are several varieties of chlorophyll, several of which are generally present within a given plant. Chlorophyll-a is the principal photosynthetic pigment in plants and cyanobacteria. Some other bacteria, such as the green sulfur bacteria, utilize a closely related substance called bacteriochlorophyll. The halophiles, members of the archeobacteria, use retinol (the same light-sensitive pigment in the human retina) rather than chlorophyll to capture light energy. All chlorophylls strongly absorb light in the red and blue parts of the visible spectrum (the green color of plants results from a lack of absorption of green light).

It is in the dark stage of photosynthesis that carbohydrate is actually synthesized. At this point, there is a divergence in the chemical pathways. In C_3 plants the initial carbohydrate produced (3-phosphoglycerate) is a 3-carbon chain. This process, known as the Calvin Cycle (illustrated in Figure 12.16) is used by all marine plants and about 90% of terrestrial plants. The first step is *phosphorylation*, or the addition of a phosphate group. In this reaction, ribulose 5-phosphate, a C_5 sugar containing one phosphate, is converted to ribulose 1,5-bisphosphate, with the additional phosphate coming from the ATP generated during the light stage. In the next step, an enzyme called *ribulose bisphosphate carboxylase oxygenase* (RUBISCO) catalyzes a reaction in which ribulose 1,5-bisphosphate reacts with one molecule of CO_2 to produce three molecules of 3-phosphoglyceric acid. Then an additional phosphate is added to each of these molecules to form 1,3-bisphosphoglycerate. This phosphate is then replaced by hydrogen supplied by NADPH to form glyceraldehyde 3-phosphate. It is in this step that the new CO_2 in the phosphoglycerate is reduced. The ΔG for reduction of CO_2 to CH_2O is about +480 kJ/mol. Most of this energy is supplied by the oxidation of two molecules of NADPH to NADP^+ . Most of the resulting triose phosphate is converted back

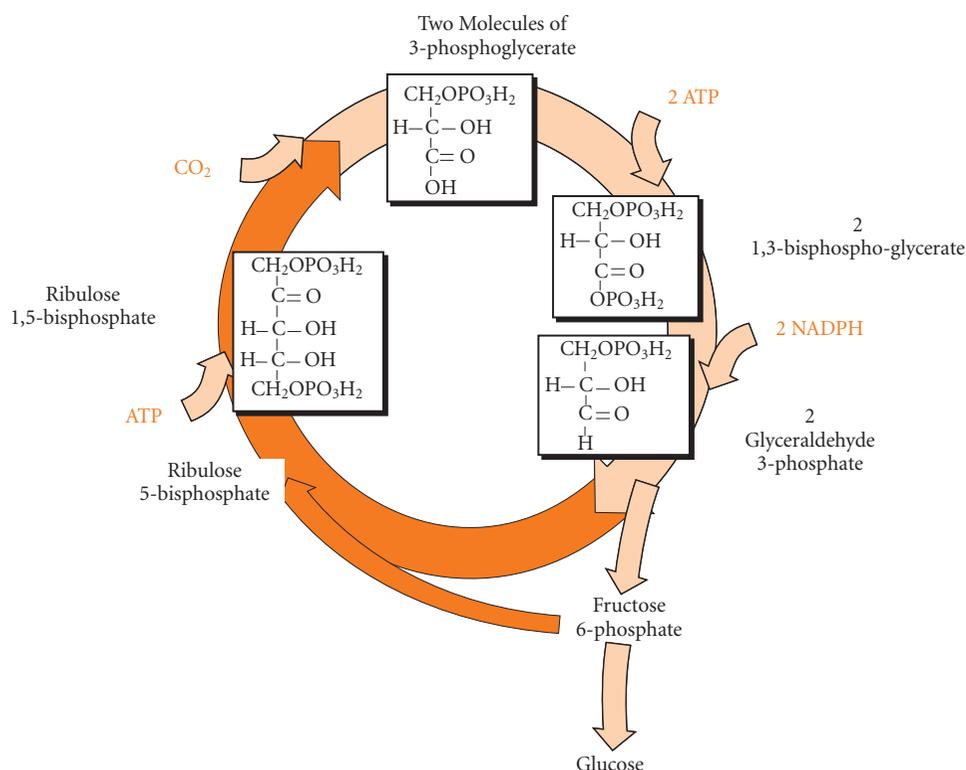


Figure 12.16 The Calvin Cycle of dark reactions for the production of glucose in C_3 plants.

to ribulose 5-phosphate for further synthesis, but some is used for synthesis of fructose or glucose. Monosaccharides produced in this way, if not immediately metabolized, are eventually converted to more complex carbohydrates or other essential compounds.

The other photosynthetic pathway is the Hatch-Slack Cycle, used by the C_4 plants, which include hot-region grasses and related crops such as maize and sugarcane. These plants use *phosphoenol pyruvate carboxylase* (PEP) to initially fix the carbon and form oxaloacetate, a compound that contains 4 carbons. CO_2 is fixed in outer mesophyll cells as oxaloacetate and is then transported, either as malate or aspartate, to inner “bundle sheath” cells, where it is decarboxylated and refixed in the Calvin Cycle (Figure 9.14). C_4 photosynthesis appears to have evolved only recently; the oldest known C_4 plants are of late Miocene age (though there has been speculation they may have first evolved in the Cretaceous). Because the C_4 pathway is some-

what more efficient at low ambient concentrations of CO_2 , there has been speculation that their appearance may reflect an evolutionary response to low atmospheric CO_2 concentrations of the late Cenozoic.

From a geochemical perspective, the most important aspect of the difference between C_3 and C_4 plants is the difference in carbon isotope fractionation during photosynthesis. The C_3 pathway produces a $\delta^{13}C$ fractionation of from -20 to -30% , whereas the C_4 pathway produces a much smaller fractionation of about -13% . Further aspects of this are discussed in Chapter 9.

Some bacteria, the chemoautotrophs, obtain the energy necessary to reduce CO_2 to carbohydrate from oxidation of sulfide to sulfate or of ferrous iron to ferric iron. As in the light stage of photosynthesis, the oxidation reactions, catalyzed by an enzyme called dehydrogenase, provide both protons (for reduction of CO_2) and energy in the form of electrons. Although they require reduced

species for energy, chemosynthetic bacteria are aerobic. Thus they live at boundaries between oxic and anoxic conditions.

12.4.2 Respiration

Respiration is the process by which organic carbon is oxidized to H_2O and CO_2 , liberating energy in the process. Polysaccharides serve as the primary energy storage in both plants and animals. In plants, these generally take the form of starches, while in animals, glycogen serves as the primary energy store. In both cases, single glucose molecules are first liberated from these chains before being oxidized. This is accomplished through the catalytic action of an enzyme.* There are a very large number of enzymes, many designed to catalyze the breaking (or formation) of a single specific bond. Glucose molecules liberated from complex carbohydrates in this fashion are then converted to two molecules of pyruvic acid in a multi-step process called *glycolysis*. The first steps involve phosphorylation and conversion of glucose to fructose diphosphate. The fructose diphosphate is then cleaved into two triose phosphates and converted to pyruvic acid (CH_3COCOOH) upon removal of the phosphates. This liberates 2 hydrogen atoms, which are consumed by the reduction of NAD^+ to NADH , per glucose molecule. The reduction of each NAD^+ consumes about 219 kJ/mol. This energy can be later used to produce ATP. Although ATP is consumed in the initial phosphorylation steps, it is produced in subsequent steps and the entire process results in a net production of 2 ATP molecules per glucose molecule (and an additional energy gain of 80 kJ/mol).

Glycolysis does not release CO_2 . Energy released in the process comes primarily from oxidation of hydrogen in the sugar to hydroxyl and resulting conversion of a sugar (glucose) to a carboxylic acid (pyruvic acid). Thus glycolysis releases only a small fraction of the energy stored in the glucose molecule. Under aerobic conditions, further energy may be obtained by oxidation of the pyruvic acid

molecule. This oxidation of pyruvic acid occurs within the mitochondria of eukaryote cells in a process called the *Krebs Cycle* (illustrated in Figure 12.17), or alternatively the *citric acid cycle* (in contrast, glycolysis occurs within the general cell cytoplasm). In the Krebs Cycle, pyruvic acid is first oxidized to acetyl coenzyme A (a reactive thioester) plus CO_2 . This acetate is then combined with an oxaloacetate produced in an earlier Krebs cycle to form citrate. Two additional CO_2 molecules as well as additional hydrogens are then successively stripped, ultimately producing an oxaloacetate. Thus one pass through the Krebs cycle completely oxidizes one pyruvic acid molecule, producing 4 molecules of NADH , one of FADH_2 (FAD , flavin adenosine dinucleotide, is an oxidizing agent similar to NAD), and one of ATP. The hydrogen stored in FADH_2 and NADH is passed along through a series of protein carriers until it is finally combined with free oxygen to form H_2O . The last step generates 3 molecules of ATP per water molecule produced. In all, 36 molecules of ATP can be produced per molecule of glucose, 34 of these being produced in the Krebs Cycle.

Fats serve as secondary energy storage in both plants and animals. Because they contain less oxygen, they store approximately twice as much energy per unit weight as carbohydrates. Respiration of fats involves breaking them into their component fatty acid and glycerol units. The glycerol is converted to triose phosphate, much as glucose, and then enters the Krebs Cycle as pyruvate. The fatty acids undergo successive loss of C_2 units that subsequently enter the Krebs Cycle as acetyl coenzyme A.

When oxygen is absent in cells, hydrogen carried by NADH cannot be oxidized to H_2O and the Krebs Cycle cannot operate. Some organisms that normally live under aerobic conditions can switch to an alternative metabolism, called *fermentation*, which does not require oxygen. Such organisms are called *facultative anaerobes*. When such cells are deprived of oxygen, NAD^+ is regenerated from NADH by reduction of pyruvic acid.

* Enzymes are denoted by the ending “-ase”. Classes of enzymes that catalyze hydrolysis reactions are named by combining the name of the class of compounds whose hydrolysis they catalyze with the suffix “-ase”. Thus lipases catalyze the hydrolysis of lipids, amylases catalyze the hydrolysis of starches, etc.

Lactic acid* ($\text{CH}_3\text{CH}(\text{OH})\text{COOH}$) is one of several possible products of this process. In alcoholic fermentation, carried out by yeast, CO_2 is removed from pyruvate leaving acetaldehyde (CH_3CHO), which is then reduced

to ethanol ($\text{CH}_3\text{CH}_2\text{OH}$). Many bacteria can live only by fermentation and are poisoned by free oxygen. Such organisms are called *obligate anaerobes*. These bacteria, of which there is a great variety, carry out cellular respiration

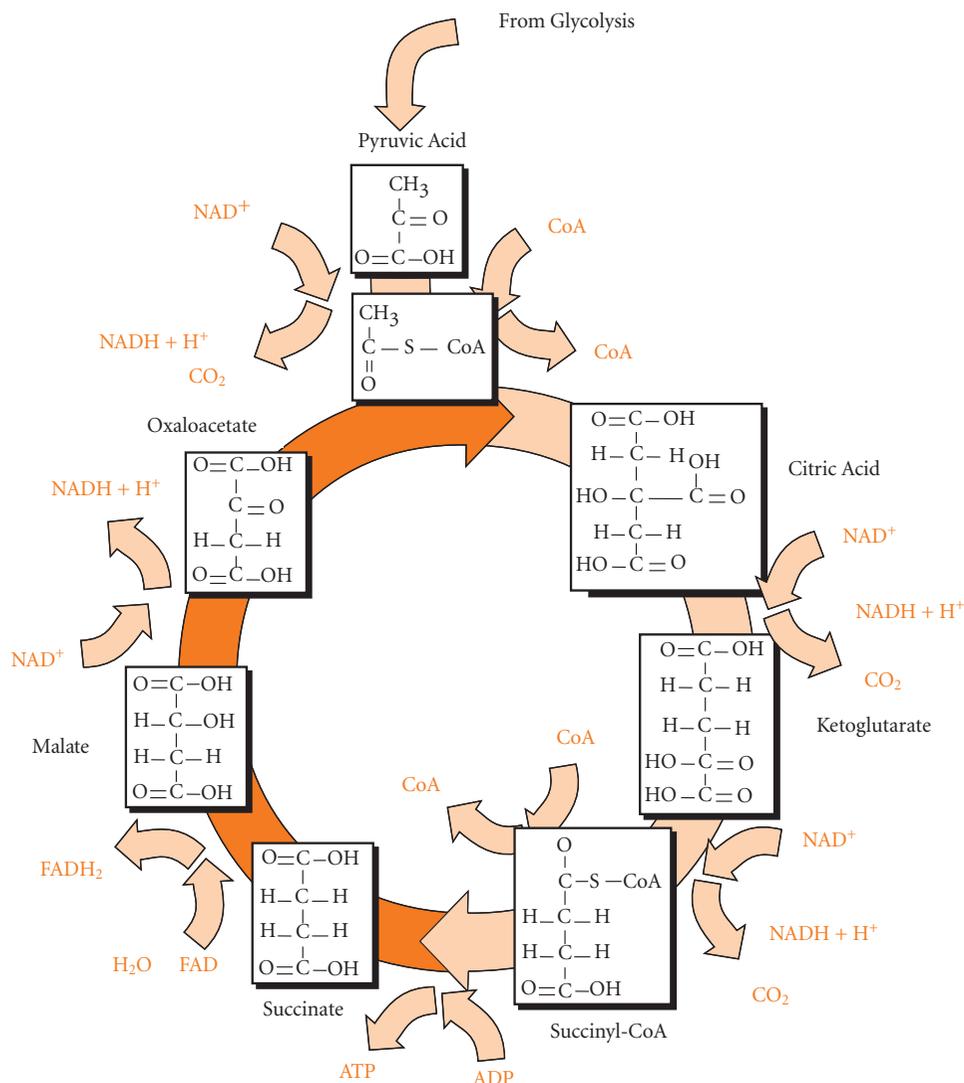


Figure 12.17 Illustration of the Krebs Cycle, or citric acid cycle, in which organic matter is oxidized by organisms to produce energy in the form of ATP. New pyruvate enters the cycle at the top, where one carbon is removed from the structure. The remaining two carbons join with a four-carbon oxaloacetate chain from a previous cycle to form citric acid, from which two additional carbons are successively stripped.

* Animal muscle cells have the ability to switch to fermentation when oxygen cannot be supplied fast enough for NAD^+ to be regenerated. Under these circumstances, lactic acid, the formal name for which is 2-hydroxy-propanoic acid, builds up in muscle tissue. The presence of lactic acid contributes to the sore muscle feeling that often accompanies vigorous exercise. "Conditioning" involves a number of processes, including increasing efficiency of oxygen delivery and a build-up of glycogen stores in muscle tissue, but also increasing tolerance to lactic acid. Nerve and brain cells, however, cannot respire anaerobically and quickly die when deprived of oxygen.

by using a compound other than oxygen as an electron acceptor. The most common such receptors are nitrate, nitrite, and sulfate, but some bacteria can reduce Fe^{3+} and Mn^{4+} .

12.5 ORGANIC MATTER IN NATURAL WATERS AND SOILS

12.5.1 Dissolved organic substances

Virtually all water at the surface of the Earth contains organic substances. These can be divided into *particulate organic matter* (POM) and *dissolved organic matter* (DOM). In a sense, though, there is a continuum, with colloids being intermediate between dissolved and particulate forms. By definition, dissolved substances are those that pass through a 0.45 μm filter; particulate organic matter is anything coarser than this. Together, dissolved and particulate organic carbon are called *total organic matter* (TOM). The carbon fraction of this organic matter is called *total organic carbon* (TOC) and can likewise be divided into *particulate organic carbon* (POC), which is also called *suspended organic carbon* (SOC), and *dissolved organic carbon* (DOC). In some instances, several other distinctions are made. Those organic substances that are volatile at ambient temperature constitute volatile organic carbon (VOC). Colloids have sizes in the range of 1 nanometer to 0.45 μm , corresponding to molecular weights greater than about 10,000. They may be specifically referred to as colloidal organic matter (COM).

Figure 12.18 illustrates the range in concentrations of dissolved and particulate organic carbon. Groundwater and seawater typically have the lowest organic carbon concentrations, while wetland waters (bogs, swamps, marshes) have the highest. The interstitial waters of the upper layers of soils often have DOC concentrations in the range of 20–30 mg/l. Organic carbon concentrations in natural waters generally show strong seasonal variations, as factors such as rainfall, biological productivity and microbial activity all may vary seasonally. One might also note that even rain and snow typically contain some DOC. The concentration is generally 1 mg/l or less, but much higher concentrations have been measured.

Organic substances dissolved in natural waters can be divided into *simple compounds* and *humic substances*. Simple compounds include *biomolecules*, which we will define as

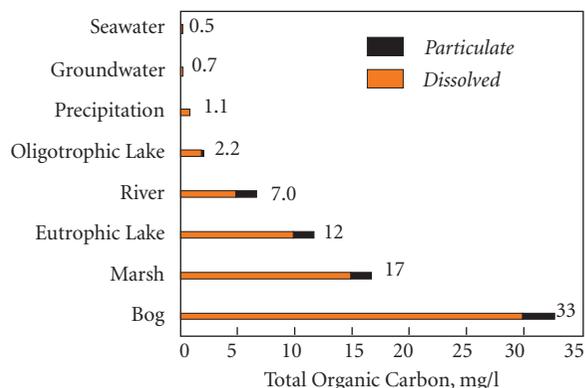


Figure 12.18 Average concentration of dissolved, particulate, and total organic carbon in various natural waters. After Thurman (1985), Figure 4.1, p. 105. With kind permission from Springer Science+Business Media B.V.

molecules synthesized by organisms to directly support life, and breakdown products of biomolecules, as well as excreted waste products and metabolites and substances released, or exuded, by cells of organisms to assist in breaking down and assimilating nutrients. Simple compounds include carboxylic acids (including fatty acids), phenols, carbohydrates, amino acids, nucleic acids, and hydrocarbons. As we shall see, humic substances are complex and poorly understood substances found in water and soil. Their exact origin is not well understood and is still somewhat controversial, but they probably include both residues of complex biomolecules and polymers that form from condensation of simpler biomolecules. Humic substances are further divided into fulvic and humic acids based on their solubility. *Hydrophilic acids* are a third set of compounds closely related to humic substances. Hydrophilic acids are, however, more soluble (as their name implies) and generally simpler than humic and fulvic acids. Humic, fulvic, and hydrophilic acids are the most abundant organic substances in river water, as Figure 12.19 illustrates. Simple organic compounds compose a smaller fraction of DOM in deep ocean water than in river water.

12.5.1.1 Simple compounds

Simple compounds typically constitute 10 to 60% of dissolved organic carbon in natural

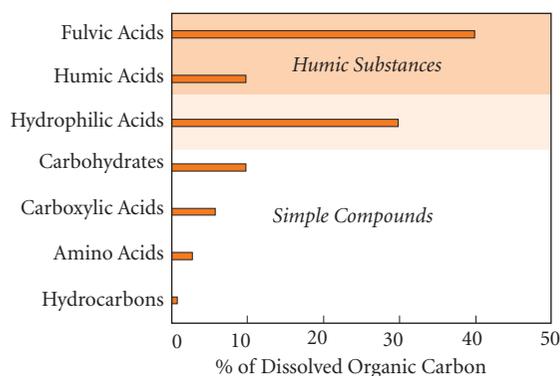


Figure 12.19 Components of dissolved organic carbon in typical river water. After Thurman (1985), Figure 4.1, p. 105. With kind permission from Springer Science+Business Media B.V.

waters, with an average of about 25%. Simple compounds are derived directly from cells or the breakdown of dead cells. Most of them are readily metabolized by bacteria, hence their residence time in water is usually quite short, probably a matter of minutes to days. Concentrations are maintained by continuous production by the biota. We might expect under these circumstances that the dissolved concentrations of these substances would be quite variable, and this is indeed the case. A number of instances of regular diurnal or seasonal variations in concentrations of carbohydrates and other simple compounds have been documented (reviewed in Thurman, 1985).

Fatty acids are among the most abundant of the dissolved simple organic compounds. They are also among the most important for several reasons. First, the carboxyl group acts as an acid, and hence these substances can significantly affect pH. Second, carboxylate can complex metal ions and will thus affect the concentration and activity of metal cations in solution. Third, they apparently form the basic building blocks of marine and aquatic humic substances.

The short-chained carboxylic acids (C_1 to C_5), which include acetic acid (the most common), formic acid, propionic acid (CH_3CH_2COOH), butyric acid ($CH_3(CH_2)_2COOH$), and valeric acid ($CH_3(CH_2)_3COOH$), have relatively high vapor pressures and hence are known as volatile fatty acids. Carboxylic acids are produced by microbial fermentation of more complex compounds. They are typi-

cally present in concentrations of 10 to $100\mu\text{g/l}$ in natural waters, but can reach higher concentrations in petroleum-related ground waters.

The longer-chained, or non-volatile, fatty acids are produced by hydrolysis of triglycerides. Though these may range from C_{12} to C_{36} , C_{16} and C_{18} fatty acids (e.g., palmitoleic acid, its saturated equivalent palmitic acid, stearic acid, and its saturated equivalent linoleic acid) are most common. The concentrations of non-volatile fatty acids are typically 5 to $500\mu\text{g/l}$ and generally somewhat higher than those of volatile fatty acids (Thurman, 1985).

Hydroxy and keto acids, particularly lactic, glycolic (also called hydroxy-acetic acid, $OHCH_2COOH$), and pyruvic acid, also occur in solution in natural waters in concentrations in the range of 10 to $250\mu\text{g/l}$. They can originate from breakdown of plant matter, but most are probably directly excreted by algae and other organisms; hydroxy acids are produced in both the Krebs and Calvin cycles, as we have seen. Other related compounds that have been identified in solution in natural waters include dicarboxylic acids such as oxalic acid, and aromatic carboxylic acids such as salicylic and benzoic acid. These generally do not exceed concentrations of a few $\mu\text{g/l}$.

Phenols and other monomeric or dimeric aromatic compounds are generally present in concentrations of a few $\mu\text{g/l}$ or less in uncontaminated waters. Natural phenols result primarily from breakdown of lignins and tannins. Anthropogenic activity (i.e., pollution), however, can result in higher concentrations of phenols. An interesting example of “natural pollution” was the occurrence of high concentrations of phenol and cresol (a phenol unit with a methyl group attached to the ring) in South Fork Castle Lake in Washington State, US. This lake was created when debris flows from the 1980 catastrophic eruption of Mt. St. Helens blocked the South Fork of Castle Creek. Concentrations of phenol approached $1000\mu\text{g/l}$ and that of cresol approached $100\mu\text{g/l}$ (McKnight *et al.*, 1982). These compounds were presumably produced by decomposition of the many dead trees floating in the lake.

Carbohydrates constitute about 10% of dissolved organic matter. Carbohydrates are found in a variety of other forms and are also present as sugars bound to humic substances.

Table 12.1 Total dissolved carbohydrates in natural waters.

| | Concentration μg/l | % DOC |
|-------------|-----------------------|-------|
| Groundwater | 65–125 | 1–4 |
| Seawater | 100–1000 | 5–10 |
| River water | 100–2000 | 5–10 |
| Lake water | 100–3000 | 8–12 |

From Thurman (1985). With kind permission from Springer Science+Business Media B.V.

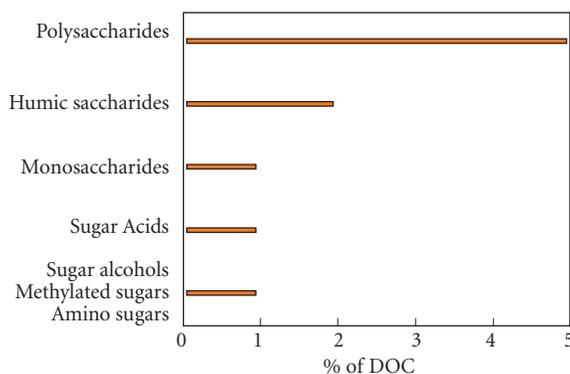


Figure 12.20 Abundances of dissolved carbohydrates in seawater as a percentage of total dissolved organic carbon. After Thurman (1985), Figure 7.4, p. 187. With kind permission from Springer Science+Business Media B.V.

Table 12.1 summarizes typical carbohydrate concentrations in various natural waters. In fresh waters, most carbohydrates are probably derived from leaching of terrestrial plant matter. In seawater, phytoplankton are thought to be the principal source of dissolved carbohydrates. Figure 12.20 illustrates the concentrations of various carbohydrates in seawater. As this figure shows, polysaccharides are the most abundant carbohydrates, followed by saccharides associated with humic substances. Monosaccharides, sugar acids, amino sugars, methylated sugars and sugar alcohols are less abundant. Of the monosaccharides, glucose is generally the most abundant. Other important monosaccharides include fructose, arabinose, xylose, and galactose.

Carbohydrates are readily assimilated and metabolized by bacteria; hence their concentration in water is probably regulated by bacterial activity. A number of studies have

demonstrated diurnal variations in carbohydrate concentration, with minimums occurring in the early morning and the maximum concentration occurring in the late afternoon. These results suggest active production of dissolved carbohydrate by phytoplankton and rapid consumption by bacteria. Thus the residence time of a carbohydrate molecule in solution may be as short as a few hours in many instances (Thurman, 1985).

Amino acids are present in natural waters both as free molecules and combined in peptides and proteinaceous substances, and are also important constituents of humic substances. The concentrations of free amino acids are typically a factor of four or so lower than the concentrations of combined amino acids. The simpler amino acids, glycine, serine, alanine, lysine, and valine, are the ones most abundant as free amino acids. Typical concentrations of free and combined amino acids are 15–200 μg/l in seawater, 15–250 in groundwater, 40–800 in rivers and lakes, 250–5000 in eutrophic lakes and marshes, and 400–9000 μg/l in interstitial waters of sediments (Thurman, 1985). Since amino acids are essential to all life, the sources of amino acids are quite varied.

A variety of hydrocarbons are also present in natural waters. These may be divided into volatile and non-volatile hydrocarbons, depending on the vapor pressure and boiling point. Short-chained hydrocarbons tend to be volatile, and this volatility limits their abundance in natural waters. Longer-chained hydrocarbons are not volatile. Their abundance in natural waters is often limited by their solubility, which is generally low. Hydrocarbons may originate both from natural and anthropogenic sources. Here we focus on the naturally occurring hydrocarbons.

Of the volatile hydrocarbons, methane is by far the most abundant. Some methane can be produced in the digestive tracts of higher animals, but the principal source of natural methane is methanogenic bacteria in reducing sediments. The amount of methane produced in such environments (swamps, reducing marine and lake sediments, etc.) is a significant, though minor, part of the global carbon cycle. The amount of methane released to the atmosphere from natural sources annually, primarily reducing sediments in wetlands, is 1.5×10^{14} g; anthropogenic activities release

Table 12.2 Dissolved methane in natural waters.

| | Concentration μg/l |
|--------------------|-----------------------|
| Groundwater | 10–10,000 |
| Seawater | 10–100 |
| Lake water | 10–10,000 |
| Interstitial water | 100–10,000 |

From Thurman (1985). With kind permission from Springer Science+Business Media B.V.

about 3.6×10^{14} g (Graedel and Crutzen, 1993). The abundance of methane in natural waters is summarized in Table 12.2. The solubility of methane in water is 800 μg/l (at 20°C and 0.1 MPa). At concentrations above this level, methane bubbles will form and this process is undoubtedly important in the transport of methane from sediment interstitial waters of swamps and shallow lakes to the atmosphere. Some of this methane redissolves in the lake water and is oxidized by methanotrophic bacteria.

Among the most common non-volatile and semi-volatile hydrocarbons in seawater are C₁₅ and C₁₇ n-alkanes (pentadecane and heptadecane) and isoprenoids. C₁₅ and C₁₇ originate by decarboxylation of C₁₆ and C₁₈ fatty acids (palmitic and stearic acids), which are derived from zoo- and phytoplankton. The isoprenoids, primarily pristane and phytane, are derived from the phytol chain in the chlorophyll molecule. *Halogenated hydrocarbons* (i.e., hydrocarbons where one or more hydrogens are replaced by a halogen) are of particular interest because of their toxicity. Most halogenated hydrocarbons in natural waters are anthropogenic, having been directly manufactured and discarded or leaked into natural waters. Others, such as chloroform and trichloromethane, can arise indirectly through chlorination of drinking water. However, some halogenated hydrocarbons do occur naturally in seawater at very low abundances.

12.5.1.2 Humic substances

Humic substances are high molecular weight (>500 u) compounds that are produced by partial degradation of complex biomolecules and recombination of these with simple bio-

molecules and their breakdown products. They are complex molecules: their exact structures are not known, and in any case are variable. As we noted earlier, humic substances dominate dissolved organic matter in natural waters. They are also important contributors to particulate organic matter.

The exact definition of humic substances is analytical and varies somewhat between dissolved humic substances and soil humic substances. Aquatic and marine humic substances are defined as colored (yellow, brown, or black), polyelectrolytic acids that can be removed from solution through absorption on weak-base ion exchange resins (e.g., the acrylic-ester resin XAD-8) or through some similar procedure. Humic substances tend to be highly refractory in the biological sense, that is, they resist decomposition by organisms. As a result, their residence time in water is on the scale of weeks to thousands of years. Their exact molecular structures are variable and therefore difficult to characterize, but in general they consist of a flexible chain with limited branching and cross-linkage to which smaller molecules and molecular fragments such as sugars and amino acids are attached. While the overall structure is poorly characterized, the chemical composition and functional groups are well characterized. The most important of these functional groups are carboxylic, phenolic, alcoholic, carbonyl, amino, and sulfhydryl (SH) groups. Because of the predominance of the carboxyl groups, humic substances are acidic (i.e., they act as proton donors). *Hydrophilic acids* are closely related to humic substances: they are molecules too complex to fully describe, but are not absorbed as readily on to ion exchange resins. They are slightly colored, highly branched, and highly substituted organic acids. They appear to have lower molecular weight and a greater number of acid functional groups than humic substances.

Soluble humic substances in waters are divided into *fulvic acid* and *humic acid*. The definition of these two is again analytical. Humic acids are defined as those humic substances that precipitate when the solution is acidified with HCl to a pH of 1. Fulvic acids are those substances remaining in solution at this pH.

The concentration ranges of humic substances are given in Table 12.3. As a proportion

of dissolved organic carbon, hydrophilic acids predominate in seawater and groundwater, constituting about 50% of DOC. They are only slightly less important in lake waters. In wetlands, streams and rivers, fulvic acids are the most important fraction of dissolved organic carbon, constituting 60% of wetland DOC and 40% of stream and river DOC. Humic acids are always less abundant than fulvic and hydrophilic acids. Humic acids comprise 15% of wetland DOC and less than 10% of DOC in other waters (Thurman, 1985).

While the definition of humic and fulvic acids is based on their acid solubility, there is nevertheless a compositional difference between the two. Table 12.4 compares the compositions of humic and fulvic acids from several environments. The table shows that on the whole, humic substances are H-poor compared with most biomolecules. For example, carbohydrates have a C:H weight ratio of 6; the C:H ratios in Table 12.4 range from 7 to 17. Humic acid is on average poorer

in hydrogen and richer in nitrogen than fulvic acid and usually poorer in oxygen as well. These compositional differences are better developed in soil humic substances than in dissolved ones.

Fulvic acid tends to have a lower average molecular weight (800–2000 u) than humic acid (>2000 u). In addition, fulvic acid has a higher content of carboxyl acid groups than does humic acid, while humic acid is richer in phenolic groups (Table 12.5). The higher content of carboxyl groups in fulvic acid and the greater abundance of aromatic structures in humic acid help to explain the higher solubility of fulvic acid. In general, dissolved humic substances have higher concentrations of carboxyl groups than soil humic substances. The average concentration of carboxyl groups in dissolved fulvic acid, 5.5 mM/g, corresponds roughly to one carboxyl group per 6 carbon atoms. In contrast, dissolved humic acid has about one carboxyl group per 12 carbon atoms.

In addition to functional groups, humic substances incorporate a variety of biomolecules in their structures. Carbohydrates account for 1% of the carbon in dissolved humic substances, and their abundances appear to be higher in humic than in fulvic acids. Many of these carbohydrates appear to be bound to the humic structure only through hydrogen bonds. Amino acids are also present in humic substances, although in smaller amounts. The most abundant are glycine, aspartic acid, glutamic acid, and alanine. Aquatic humic acids contain about 120 nM/mg amino acid;

Table 12.3 Humic substances in natural waters.

| | Concentration mg C/l |
|-------------|-------------------------|
| Groundwater | 0.03–0.10 |
| Seawater | 0.06–0.60 |
| Lake water | 0.5–4.0 |
| Rivers | 0.5–4.0 |
| Wetlands | 10–30 |

From Thurman (1985). With kind permission from Springer Science+Business Media B.V.

Table 12.4 Composition of fulvic and humic acids dissolved in natural waters.

| | C | H | O | N | P | S | Ash |
|-------------------------|------|------|------|------|------|------|------|
| Groundwater | | | | | | | |
| Biscayne Aquifer fulvic | 55.4 | 4.2 | 35.4 | 1.8 | | | 0.04 |
| Biscayne Aquifer humic | 58.3 | 3.4 | 30.1 | 5.8 | | | 10.4 |
| Seawater | | | | | | | |
| Saragasso Sea fulvic | 50.0 | 6.8 | 36.4 | 6.4 | | 0.46 | |
| Lake water | | | | | | | |
| Lake Celyn humic | 50.2 | 3.1 | 44.8 | 1.9 | | | |
| Lake Celyn fulvic | 43.5 | 2.7 | 51.6 | 2.2 | | | |
| Stream water | | | | | | | |
| Ogeechee Stream fulvic | 54.6 | 4.97 | 38.2 | 0.87 | 0.62 | 0.74 | 0.86 |
| Ogeechee Stream humic | 55.9 | 4.19 | 36.5 | 1.27 | 0.25 | 0.93 | 1.13 |

Data from Aiken *et al.* (1985). Concentrations in weight percent.

aquatic fulvic acids generally contain less: 15–120 nM/mg. Soil humic substances contain higher amino acid concentrations: roughly 500 nM/mg for soil humic acids and 150 nM/mg for soil fulvic acids. Hydrogen bonding appears to play a role in molecular aggregation.

Soil humic substances have core structures that are highly aromatic (20 to 70% of total C). It was originally thought that this was also true of aquatic humic substances. However, a variety of studies over the last 30 years have led to the view that the core structures of aquatic and marine humic substances are primarily aliphatic rather than aromatic (Malcolm, 1985; Thurman, 1985; Steinberg and Muenster, 1985; Harvey and Boran, 1985). Nuclear magnetic resonance studies

Table 12.5 Functional groups in dissolved fulvic and humic acids. From Thurman (1985). With kind permission from Springer Science+Business Media B.V.

| | Carboxyl meq/l | Phenolic meq/l |
|-------------------------|----------------|----------------|
| Groundwater fulvic | 5.1–5.5 | 1.6–2.0 |
| Groundwater humic | | 2.5 |
| Seawater fulvic | 5.5 | – |
| Lake fulvic | 5.5–6.2 | 0.5–2.1 |
| Lake humic | 5.9 | 3 |
| Stream and river fulvic | 5.5–6.4 | 1.5–2.1 |
| Stream and river humic | 4.0–4.7 | 1.9–2.0 |

suggest only 15–20% of the carbon in aquatic fulvic acids and about 30% of the carbon in aquatic humic acids is aromatic. Figure 12.21 illustrates one proposed hypothetical structure of aquatic fulvic acid. This molecule has a molecular weight of about 1000 u.

There are differences between marine and aquatic humic substances. Derivatives of lignin appear to be important in the backbone of aquatic humic substances, but not marine. Marine humic acids appear to have an even smaller proportion of aromatic carbon than aquatic ones, and marine fulvic acids have essentially none. Aromatic structures are far more common in terrestrial plants than marine ones, so this suggests that terrestrial humic matter does not contribute substantially to marine humic matter.

The shape of humic substances is strongly influenced by pH. At low pH, humic and fulvic molecules are fibrous. At neutral pH, these fibers tend to mesh together to form a sponge-like structure. This mesh structure is capable of trapping smaller molecules. For example, fatty acids and acyclic alkalenes, which are found in most humic and fulvic acids, may be simply trapped rather than bonded to the molecular backbone. Under alkaline conditions the structure becomes plate-like.

The origin of dissolved humic substances is not well understood. At one time it was widely held that aquatic humic substances consist of

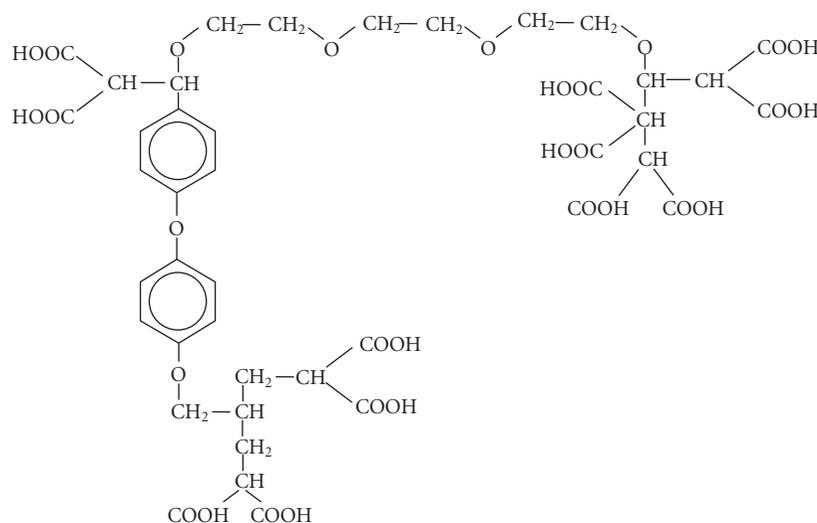


Figure 12.21 Hypothetical structure of aquatic fulvic acid. From Bergmann (1978). Reproduced with permission.

fulvic acid that had been leached or eroded from soils. However, most recent works on the subject express the view that at least a substantial fraction of aquatic humic substances is autochthonous, that is, produced within the body of water itself. The evidence favoring this view includes the compositional and structural differences between aquatic and soil humic substances. Aquatic humic substances are primarily aliphatic, whereas soil humic substances are primarily aromatic. This observation in itself does not exclude derivation of aquatic humic substances from soil (since leaching may preferentially remove aliphatic-rich molecules), but the difference is consistent with soil humic substances being derived primarily from higher plants (which are rich in aromatic molecules such as lignin and tannins), and aquatic humic substances being derived primarily from algae and zooplankton (which are poor in aromatic molecules). An additional question is the degree to which humic substances are composed of partially degraded macromolecules or are condensed from simpler fragments. The latter view is supported by several observations. One is that clay particles, metal oxides, and metal cations can all promote polymerization of organic molecules. Bacteria, or enzymes released by them, have also been shown to promote condensation and polymerization of organic molecules. Hydrophilic acids may be precursors of the more complex fulvic and humic acids. As we noted, soil humic substances are highly aromatic, and this suggests that lignins and tannins derived from higher plants are important contributors to soil humic substances. Lignins and tannins are relatively resistant to decomposition, although they are probably modified to some degree by decomposers before incorporation into humic molecules. Molecular byproducts of decomposers and their remains as well as waxes from higher plants are probably the primary contributors to the aliphatic components of soil humic substances (Killops and Killops, 2005).

Marine fulvic acids may arise by autoxidative cross-linking of polyunsaturated lipids, perhaps catalyzed by light and transition metals (Harvey and Boran, 1985). Olefinic carbons (i.e., those doubly bonded to other carbons) may be particularly susceptible to autoxidation. Figure 12.22 illustrates this

process. A number of laboratory experiments have demonstrated the plausibility of production of marine humic substances in this manner. In this model of humic substance generation, aromatic units in marine humic acid arise from cross-linking of fatty acid chains. Another possibility is that aromatic units in marine humic acids are derived from terrestrial sources: degradation products of lignin and tannin washed into the sea. Most likely, however, terrestrial material is only a locally important contributor to marine humic substances.

12.5.2 Organic matter in soils

Organic matter can constitute a third or more of the mass of poorly drained soils, though fractions from 6% to 10% are more common in well-drained soils. Concentrations of organic matter are highest in the surface layers (O and A) and low in the deeper layers (C). In the soil solution, a range of 2–30 mg/l DOC is common (Thurman, 1985). Soil organic matter, collectively called *humus*, includes biomolecules as well as humic substances. Biomolecules can be released from dead cells, excreted from living ones, or leached from foliage and rinsed into the soil by rain (i.e., throughfall). In addition, soil organic matter includes a variety of substances, called *exudates*, released by macro- and microorganisms expressly to break down both organic and inorganic components in soils so that they may be assimilated. Among other reasons, the exudates are necessary because many biomolecules are too large to pass through microbe cell walls. Among the most important of the exudates are simple carboxylic acids such as acetic and oxalic acid (Figure 12.7), formic acid (HCOOH), tartaric acid (2,3-dihydroxybutanedioic acid: $\text{COOH}(\text{CH}(\text{OH}))_2\text{COOH}$) and citric acid, and a variety of phenolic acids. The OH groups of phenols may also dissociate, and hence these compounds also contribute to soil acidity and rock weathering, though less so than the carboxylic acids. These simple organic acids are commonly present in relatively high concentrations around plant roots in soils, though on average their concentration is less than 1 mM in the soil solution (Drever and Vance, 1994). Because of the presence of both these simple acids and the more complex fulvic and

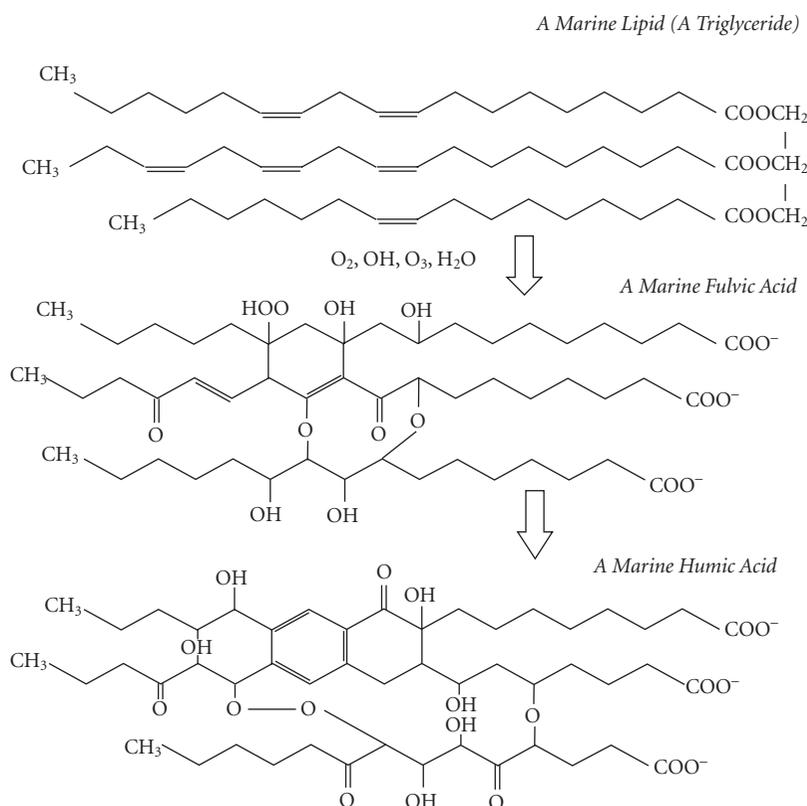


Figure 12.22 Possible mechanism for generation of marine fulvic and humic acids from triglycerides. Harvey and Boran (1985). With permission from John Wiley & Sons.

humic acids, most soils are slightly acidic. These acids contribute both directly (through surface complexation reactions) and indirectly (as proton donors, by increasing the solubility of cations through complex formation) to the weathering of rocks, although the overall extent to which weathering is accelerated by plants is unclear. As most biomolecules are readily metabolized by bacteria, their residence time in the soil is likely to be quite short, a matter of days or less. Concentrations of these molecules are maintained by continuous production by the biota. In contrast, the residence time of refractory humic substances in soils can be as long as thousands of years.

The definition of soil humic substances differs slightly from that of aquatic and marine humic substances. Soil humic substances are divided into *fulvic acid*, *humic acid*, and *humins* based on their solubility. By definition, the organic material that remains insoluble when 500 mol/m^3 NaOH is added to soil is

called humin. The material dissolved by this procedure may be separated into fulvic and humic acids by adjusting the pH to 1 by addition of HCl: the material precipitated at that point is *humic acid*, that remaining in solution is *fulvic acid* (the procedure for separating fulvic and humic acid is the same for aquatic and marine humic substances).

As was the case for aquatic humic substances, there is a systematic compositional difference between soil humic and fulvic acids; indeed the compositional differences between fulvic and humic acid appears greater in soils than in water. As is summarized in Table 12.6, humic acid is richer in carbon and poorer in oxygen and sulfur than fulvic acid, and poorer in hydrogen and richer in nitrogen as well. The higher content of polar groups such as carboxyl accounts for the higher solubility of fulvic acid. Soil fulvic acids tend to have molecular weights $< 2 \times 10^3$, while the molecular weight of soil humic acids can exceed 10^6 .

Based on these compositions, the approximate chemical formula for average humic acid is $C_{187}H_{189}O_{89}N_9S$ and that of fulvic acid is $C_{68}H_{91}O_{48}N_3S$. Compared with the composition of living organisms, humic and fulvic acids have substantially lower ratios of hydrogen, oxygen, and nitrogen to carbon. Since humic substances are ultimately derived from

biomolecules, this indicates that H, O, and N are lost in the humification process. Humin, the insoluble organic matter in soil, tends to be even richer in carbon and nitrogen and poorer in sulfur than either fulvic or humic acid. From these formulas we can deduce that there is a substantial degree of unsaturation (i.e., double carbon bonding) in humic substances, which partially accounts for their stability and biologically refractory nature. Stevenson and Vance (1989) estimated the average content of functional groups in soil humic and fulvic acids as 7.2 and 10.3 meq/g respectively. One possible structure for soil humic acid is illustrated in Figure 12.23.

Table 12.6 Composition of soil humic and fulvic acids.

| | Humic acid | | Fulvic acid | |
|----------------------------|------------|---------|-------------|-----------|
| | Mean | Range | Mean | Range |
| Elemental comp. | wt% | | | |
| C | 56. | 53–59 | 45.7 | 40.7–50.6 |
| H | 4.6 | 3.0–6.5 | 5.4 | 3.8–7.0 |
| N | 3.2 | 0.8–5.5 | 2.1 | 0.9–3.3 |
| O | 35.5 | 32–38.5 | 44.8 | 39–50 |
| S | 0.8 | 0.1–1.5 | 1.9 | 0.1–3.6 |
| Functional groups | meq/g | | | |
| Total acidic groups | | 5.6–8.9 | | 6.4–12.2 |
| Carboxyl | 3.6 | 1.5–6.0 | 8.2 | 5.2–11.2 |
| Phenolic OH | 3.1 | 2.1–5.7 | 3.0 | 0.3–5.7 |
| Alcoholic OH | | 0.2–4.9 | | 2.6–9.5 |
| Quinoid/keto C=O | | 0.1–5.6 | | 0.3–3.1 |
| Methyloxy OCH ₃ | | 2.1–5.7 | 3.0 | 0.3–5.7 |

After Schnitzer (1978). With permission from Elsevier.

The origin of soil humus is not precisely known. The relatively high proportion of aromatic units suggests the most important contributors to humic substances are lignins and tannins. As we noted earlier, these are polyaromatic substances that are quite refractory in a biological sense. These are partially degraded by soil microbes. Monomer or smaller polymer units may then condense, perhaps catalyzed by clays, metal ions, or bacteria. To this basic structure other components, which include amino acids, carbohydrates, and alkanes derived from fatty acids, may be attached. Soil microbes are probably the primary contributors of these units.

As we found in Chapter 13, organic compounds in soils, particularly carboxylic acids such as oxalic acid, play an important role in podzolization – the depletion of Fe and Al in the upper soil horizons and their enrichment in lower horizons. This occurs as a result of the ability of carboxylic acids to form soluble

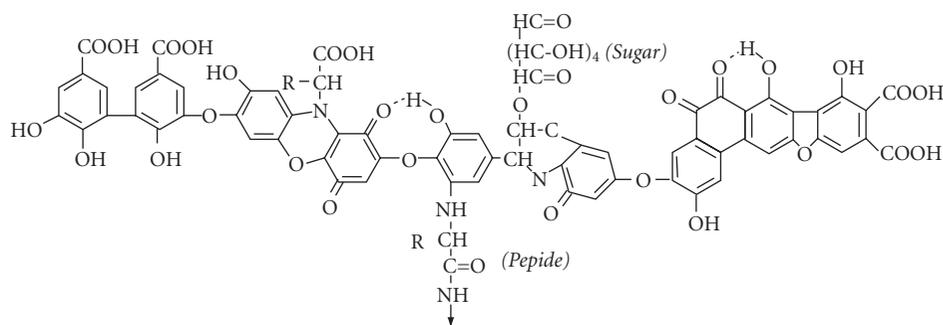


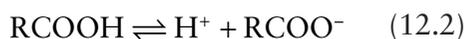
Figure 12.23 Possible structure of soil humic acid. Dashed lines represent hydrogen bonds. Stevenson (1982). With permission from John Wiley & Sons.

complexes with Fe and Al. Fe and Al carboxylate complexes form in the upper soil layers, where organic acid concentrations are high, then are carried to deeper levels by water flow. At deeper levels, bacteria oxidize the carboxylate, and the Fe and Al precipitate as hydroxides. We examine metal–organic complexation further in the following section.

12.6 CHEMICAL PROPERTIES OF ORGANIC MOLECULES

12.6.1 Acid–base properties

As we noted above, the carboxyl group can dissociate to give up a hydrogen atom:



(we use R here as a general representation for the remainder of the molecule). Like other acids, organic acids will increasingly dissociate as pH increases. As for other reactions, we can write an equilibrium constant expression:

$$K_a = \frac{a_{\text{H}^+} a_{\text{RCOO}^-}}{a_{\text{RCOOH}}} \quad (12.3)$$

The equilibrium constant is commonly reported as $\text{p}K_a$, which, analogous to pH, is the negative of the logarithm of the equilibrium constant. The Henderson-Hasselback equation relates $\text{p}K_a$, activity quotient, and pH:

$$\text{p}K_a = -\log K_a = -\frac{a_{\text{RCOO}^-}}{a_{\text{RCOOH}}} - \text{p}H \quad (12.4)$$

Expressed in this way, the $\text{p}K_a$ is the pH where half of the acid molecules are dissociated and half undissociated. Most carboxylic acids have $\text{p}K_a$ values between 1 to 5, which is below the pH of most streams and lakes, and similar to or lower than the pH range of soils (generally 3 to 8). Thus carboxylic acids will be dissociated in most environments. As we have seen, carboxylic acids (both simple and as functional groups on humic substances) are important components of dissolved and soil organic matter. The effect of increasing concentrations of organic molecules will thus be to increase the concentration of protons, lowering the pH of natural waters.

The OH groups of phenols can also dissociate and hence contribute to solution or soil acidity. Phenols are, however, much weaker acids than carboxylic acids. One measure of this is the $\text{p}K_a$. Whereas the $\text{p}K_a$ values of carboxylic acids are typically 3 or so, $\text{p}K_a$ for phenols are more typically 8. Thus at pH typical of most natural waters and soils, phenols will be only partially dissociated or not dissociated.

In contrast to carboxyl and phenol groups, nitrogen-containing groups, such as amino groups, are basic as they tend to bind free protons. By analogy to $\text{p}K_a$, defined above, we can define a $\text{p}K_b$, which is the pH when half the molecules or groups will be protonated and half unprotonated. Aliphatic amines are the most strongly basic, with typical $\text{p}K_b$ values of 10 to 12. At pH values below this, they will be protonated. Thus in most natural waters and soils they will bear a positive charge and behave as cations. Because particle surfaces are typically negatively charged, these organic cations are readily bound to particle surfaces and removed from solution. Aromatic amines are typically weaker bases, with $\text{p}K_b$ values around 4 to 6. They will be protonated only in acidic waters and soils.

As we have seen, humic and fulvic acids are generally the most abundant organic substances in natural waters and soils. They often contribute significantly to the acidity of waters and soils. Under some circumstances, such as lowland tropical rivers or swamps, they are the principle negative ions present. They typically contain 10^{-2} eq/g ionizable acid groups per weight of organic carbon. Carboxyl groups are most common, but other functional groups are also present. As a result, humic substances cannot be characterized by a single $\text{p}K_a$. Their titration curves (Figure 12.24) typically have a “smeared out” appearance, a result both of the variety of functional groups present, and electrostatic interactions between these groups.

12.6.2 Complexation

Another important geochemical property of organic molecules is their ability to form complexes with metals, especially transition metals and aluminum. Complexation between metal ions and organic anions is similar, for the most part, to complexation between metals

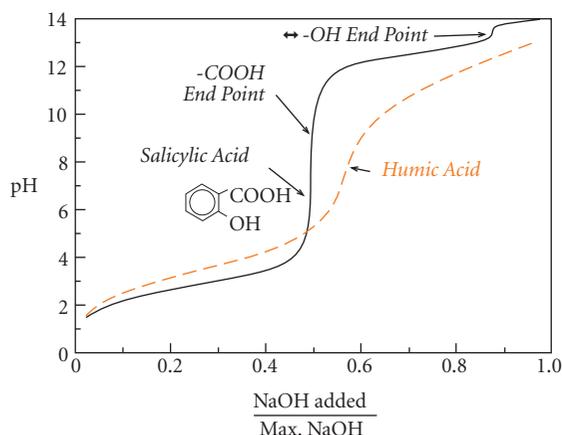


Figure 12.24 Comparison of titration curves of salicylic and humic acid. The salicylic acid shows two end-points, corresponding to its carboxyl and phenol functional groups. The humic acid shows a smeared-out titration curve corresponding to a continuously changed pK_a . Morel (1983). With permission from John Wiley & Sons.

and inorganic anions. One important difference is that many organic compounds have more than one site that can bind to the metal. Compounds having this property are referred to as *multidentate*. Complex formation with multidentate ligands is called *chelation*, and the complexes formed are called *chelates*. A simple example is the oxalate ion, $(\text{COO})_2^{2-}$, which consists of two carboxyl groups (Figure 12.6) and is *bidentate*. Citric acid and glutamic acid (an amino acid) are tridentate ligands (although the amino group will not dissociate and take on a positive charge, it nevertheless has an electron pair available to share). A second important difference is that organic functional groups are only weakly acidic, meaning they have a high affinity for protons and are often only partially dissociated at the pH of natural waters. As a result, metal–organic complexation is strongly pH-dependent.

A metal-oxalate complex results in the formation of a ring (Figure 12.25), with the two oxygens that are singly bonded to carbon each binding to the metal. In the oxalate complex, the ring has five members; a 6-member ring would be formed in a metal-malonate; 7-member rings are formed by phthalate or succinate chelates. A metal

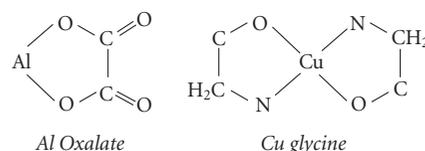


Figure 12.25 Examples of rings formed by chelates.

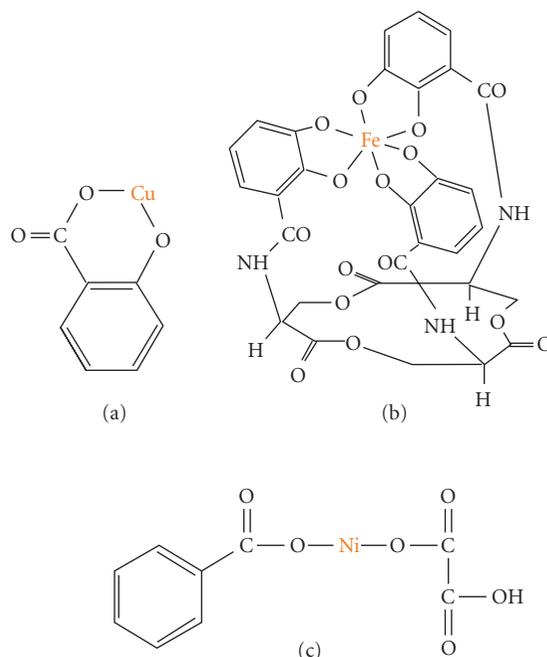


Figure 12.26 (a) Copper salicylate complex. The Cu ion is bound to both the carboxylic and phenol groups. (b) Enterobactin, a natural iron-chelating agent. (c) Ni complexed by separate benzoate and acetate groups.

glycine complex forms two rings on opposing sides of the metal (Figure 12.25). Salicylate is another example of a bidentate ion (Figure 12.26a). In this case, the binding sites are formed by two different functional groups: carboxyl and hydroxyl. Not all organic anions are multidentate, of course. Benzoate, acetate, and phenol are examples of unidentate organic anions. It is also possible for a single metal ion to bind to more than one organic ligand, as illustrated in Figure 12.26c.

Multidentate ions can very strongly bind trace metals. A few natural chelators that are specific for Fe have been characterized. One, enterobactin, is illustrated in Figure 12.26b. These Fe-specific chelators may have stability