BASIC CELL CHEMISTRY :

A Review of Chemical Compounds and their Interactions

Water

There is no life without water. In this chapter, water will be used to review some very basic ideas in chemistry, particularly as applies to cell and molecular biology. What is water? H_2O. Two hydrogen atoms and one oxygen atom (Fig. 1). Together they form a molecule of water. They are defined as a molecule by the presence of strong chemi-

If any of the first five pages is not a review, then your high school chemistry class failed to prepare you properly for college. Go take an intro chemistry class before proceeding with this course. Then go complain to your high school administrators and tell them to stop dumbing down courses and teaching to the lowest common denominator.
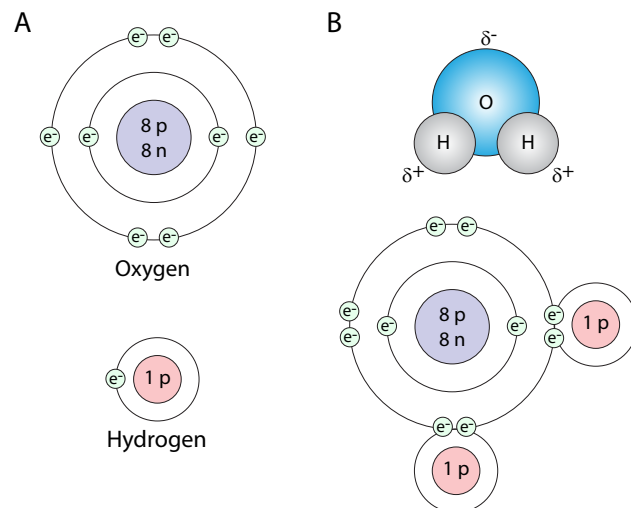

Figure 1. (A) Oxygen and hydrogen (B) Water.

cal bonds connecting each atom. In this case, each atom is connected to another by a *covalent* bond. These are the strongest type of chemical bonds, and form when two atoms are *sharing electrons* in order to fill their outermost (valence) electron shell and increase stability. In the case shown here, hydrogen (H) has only one electron, and for maximal stability of that electron shell, it should have two. Oxygen, on the other hand, has six electrons in its outer shell, and a filled shell would have eight. Thus, it would "like" to pull in two more electrons for maximal stability. As shown in Fig. 1B,
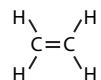
The volume of an atom is defined by electrons in a very fast and energetic obit around a nucleus. The electrons are very small negatively charged particles, and the nucleus is composed of neutrons (electrically neutral) and protons (positively charged), both relatively massive in comparison to electrons. The electrons' orbits around the nucleus can be approximated by "shells" or levels. These shells characteristically have limitations on the number of electrons that can fit within them: the first shell (closest to nucleus) holds only 2 electrons, while the second shell holds 8, and the third shell holds 18. The atom is most stable when its outer shell (and by extension, all inner ones also) is filled. The energy of the electrons also varies by level - innermost electrons have the least energy while the outermost electrons have the most.

both of those requirements are fulfilled when each of the hydrogen atoms shares an electron with the oxygen, which also shares an electron each with the hydrogen. The water molecule can also be written as H—O—H, in which the single solid line indicates a *pair of shared electrons*, i.e. a single covalent bond. The energy of an average single covalent bond is about 80 kcal/mol. However, as shown at left, double and even triple covalent bonds are possible. The strength of those types of bonds is slightly less than double (~150 kcal/mol) or triple (~200 kcal/mol) the energy of the single bonds.

Ethylene   Acetylene

Sharing electrons is not the only way to create bonds between atoms. *Ionic bonds* are created when an atom donates or receives an electron, rather than sharing one. When an atom gives up an electron, the electrical balance between the numbers of positively charged protons in its nucleus and negatively charged electrons is upset, and the overall atom now has a positive electrical charge. Similarly when an atom receives an extra electron, the balance in a neutral atom is upset, and the atom becomes negatively charged. An ionic bond is formed when one atom donates an electron to an adjacent atom, creating an ionic pair, one positively and one negatively charged. The electrical attraction between the oppositely charged atoms holds them together. Ionic bonds are weaker than covalent bonds, with an average bond energy of ~5.5 kcal/mol. Both covalent and ionic bonds are thermodynamically stable in dry, room temperature conditions (25°C, 298 K, 77°F). The average energy imparted when molecules collide at this temperature is only ~0.6 kcal/mol, far less than the energy needed to break a covalent or ionic bond.

Bond energy is a measure of the strength of the bond between two covalently joined atoms, and is proportional to the bond distance, which is determined by the atomic radii. It is not the same thing as bond dissociation energy, which is the energy released in a homolytic reaction (bond is split with electrons equally distributed) taking place at absolute zero, but they are similar in being measures of bond strength.

Although salts (such as NaCl) are ionic compounds, not all ionic compounds are salts. The chemical definition of a salt requires that the compound be formed by the substitution of a hydrogen ion (H^+) in the original compound. This usually occurs in neutralization reactions, such as the neutralization of hydrochloric acid, HCl (or H^+Cl^-) with sodium hydroxide ($Na^+(OH)^-$), which yields the salt NaCl, and water (HOH = H_2O).

Figure 2. (A) Individually, the Na atom and the Cl atom are electrically neutral. However, they are both very reactive chemically because both need only get rid of (Na) or take in (Cl) one electron to have a full outer shell. (B) Because an electron is completely transferred, the Na becomes Na^+ and Cl becomes Cl^-, reflecting the new charge imbalance. Although electrically no longer neutral, the thermodynamic enhancement from filling the outer shells makes both of these ions very stable.
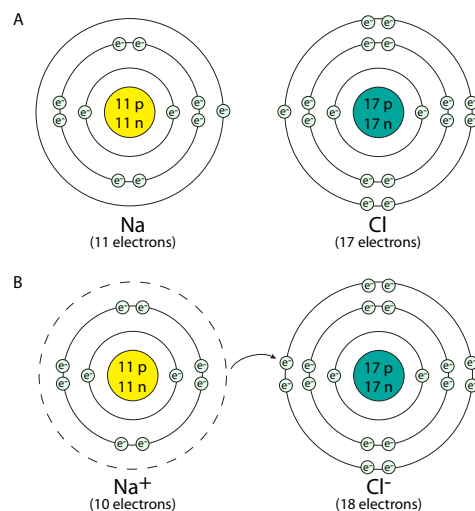
Covalent and ionic bonds between atoms are the only way to make molecules, which are stable collections of chemically bonded atoms. However, other attractive interactions between atoms and molecules exist, but they are significantly weaker, and can be disrupted with relatively small changes in temperature or environmental conditions. These are van der Waal's forces. They are very short-range interactions, requiring close apposition of the two atoms. As mentioned, an individual hydrogen bond (a specific type of van der Waal's force described below) or other van der Waal's interaction can be easily disrupted, but these types of interactions generally occur *en masse*. In a sense, they are like molecular Velcro® - each individual little plastic hook and individual loop of nylon could barely hold two hairs together, but a suit of velcro can hold a person on a vertical wall (a la Late Night with David Letterman, 1984).

In the case of hydrogen bonds, these occur when there is permanent asymmetric electron sharing within a covalently bonded molecule so that the shared electrons spend more time around one nucleus (thus imparting a negative character), than the other (which is therefore somewhat positive in character) to create a permanent electrical dipole. These dipole moments can interact with oppositely charged moments on other molecules or the same molecule. Van der Waals forces also include induced (non-permanent) dipole-dipole interactions in which a temporary shift in electron density as they orbit the nucleus forms a minute charge differential, that can induce an opposite and attractive charge differential in a very close neighboring atom. In fact, some texts define van der Waals forces exclusively as such, leaving hydrogen bonds as a separate category altogether. One of the arguments for that idea is that the bond length of the average H-bond is smaller than the sum of the van der Waal's radii of the two atoms.

As noted above, hydrogen bonds result from severely uneven sharing of electrons that generate permanent dipoles. In biological systems, this generally means that a hydrogen is covalently bound to either an oxygen or a nitrogen atom, which are both highly electronegative atoms, strongly attracting the shared electrons away from the hydrogen. Common hydrogen-bonding pairs are OH ···· :O, OH ···· :N, NH ···· :N, and NH ···· :O. Dotted lines are a common method for depicting hydrogen bonds in printed text and diagrams.

Water is a molecule that has a permanent dipole (i.e. it is a polar molecule), with the highly electronegative oxygen nucleus taking the lion's share of the shared electrons' time, leaving the hydrogen nuclei stripped bare down to their protons. The geometry of the water molecule (Fig. 1B) makes one side of the molecule somewhat negative with two pairs of free electrons, and the opposite side positive, because the shared electrons are only rarely near the hydrogen nuclei. This gives water the ability to hydrogen bond, and is the basis for several of water's most important qualities. The ability to form
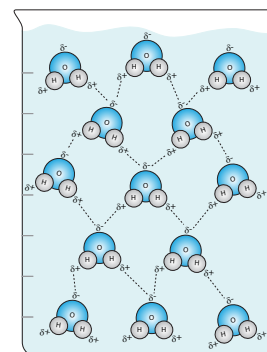

Figure 3. The hydrogen bonding of water molecules to one another is an important determinant of the physical properties of water.

many hydrogen bonds leads to a high specific heat of water, and enables it to act as a generous heat buffer. In order to get enough molecules of water moving faster and increase the temperature of the water, the energy put into the water must first be used to break apart the hydrogen bonds without generating heat. This is unlike most other liquids, which do not link internally with H-bonding. So the water is able to absorb more heat (energy) without a phase change than many other liquids.

Another important and unique characteristic of water is that the solid phase (ice) is less dense than the liquid phase. With most other liquids, as the temperature drops, the molecules have less energy, so they move less, and they stay closer together, increasing the density. Only part of that holds true with water. Again, the ability to form hydrogen bonds is directly related to this: as the temperature is lowered, the molecules move around less, affording them more opportunities to form hydrogen bonds. However, even though they are attractive, the H-bonds also act as spacers separating the water molecules more than if they were allowed to tumble about together in a liquid without forming H-bonds.

This aspect of water chemistry is actually more important to life in a geologic sense than at the cellular level. At the cellular level, the consequence is that freezing cells causes the water in them to expand and burst, killing them at low temperatures unless the cell has chemicals that act as antifreeze and lower the freezing temperature of the cytoplasm. On the other hand, at the geological level, when a pond or lake freezes in winter, the ice is less dense than water, thus staying on top of the pond, insulating deeper layers, and helping them stay liquid and able to support life (many organisms migrate deeper down in the winter). If water became more dense as it froze, as many other molecules, ice would sink, and eventually the entire pond would be completely solid, killing off most life in it once a year!

From a chemical standpoint, the polar nature of water makes it an excellent solvent for ionic and polar molecules. As you can see in the figure, the hydrogen side of water interacts with the negatively charged chloride ion, while the oxygen side of water interacts with the positively charged sodium ion, thus easily dissolving the salt. However, the polarity of water also makes it repel nonpolar molecules or by non-polar regions of molecules. This property, known as hydrophobicity, is crucial to life, since it is the basis for the formation of the biological membranes that define a cell. In general terms, the H-bonding between water molecules is very stable. Non-polar molecules cannot participate in H-bonding, and therefore create areas of instability wherever they are touching aqueous (water-based) solutions. The resolution to this problem is for hydrophobic molecules to aggregate, thus lowering the total surface area in contact with water. In living organisms, many protein and lipid molecules are amphipathic, with some portions hydrophobic, while other parts of the molecule are hydrophilic.

Acids and Bases

While it is easiest to think of water as H_2O, it is in fact in an equilibrium between the ionized molecules H^+ (which is simply a proton) and OH^- (the hydroxyl ion). The H^+ itself can be subsequently bound to a water molecule to form a hydronium ion, H_3O^+. The release of H^+ and OH^- are not limited to water molecules, and many compounds do so in aqueous solutions. These compounds can be classified as acids (raising the free H^+ concentration) or bases (increasing the free hydroxyl concentration]. The ex-

Water can dissociate from H_2O into the ions H^+ and OH^-, in which the departing hydrogen leaves its electron with the oxygen. However, H^+ is extremely reactive and almost immediately attaches to a nearby water molecule, forming the hydronium ion H_3O^+.

tent to which acids and bases donate or remove protons is measured on the pH scale, which is a logarithmic scale of relative H^+ concentration. Thus the Coca-Cola$^©$ that I am drinking, and which counts phosphoric, carbonic, and various other acids among its ingredients, has a pH around 3, which means that it liberates 10^4 times more H^+ than water, which has a pH of 7. Inside cells, the pH range is tightly restricted to slightly above neutral (neutral = pH 7), although in eukaryotes, various intracellular organelles (e.g. lysosomes) may have significantly different internal acidity/alkalinity. This is important biologically because changes in acidity or alkalinity can alter hydrogen and ionic bonds, thus potentially changing the shape and activity of enzymes and other biomolecules.

Sometimes, this can be used to an organism's advantage. For example, cells lining the stomach of an animal such as yourself secrete the enzyme pepsin into the stomach to help digest proteins. Pepsin has a pH optimum close to pH 2, which is great because stomach pH is also around 2. However, considering that cells themselves contain a lot of proteins, and we don't want pepsin-containing cells to digest themselves away, what is the solution? Because the pH inside the cell is close to 7.2, far above the pH optimum for pepsin, it is inactive inside the cell, and only works after it has been secreted into an acidic environment.

Carbon

The major constituent molecules in all living organisms are based on carbon. Carbon has versatility stemming from its four outer shell electrons allowing the possibility of four covalent bonds with a variety of partners, incuding very stable carbon-carbon covalent bonds. Because of this, long carbon chains can form the backbone of more complex molecules, and makes possible the great diversity of macromolecules found in the cell. The carbon chains themselves are not very reactive, but they often have reactive chemical groups attached to them.

Common groups are the hydroxyl ($-OH$), carbonyl ($-CO$), carboxyl ($-COOH$), and phosphate ($-PO_4$). Carbon chains may even have other carbon chains attached to them: the smaller ones behave and are named as groups also: methyl ($-CH_3$), ethyl, ($-C_2H_5$), propyl ($-C_3H_7$), and so forth. Figure 4B (right) depicts several functional groups that can be found in the simple molecule, acetic acid (very dilute acetic acid is the primary component of vinegar).

Carbon is also the basis for the four major classes of biological molecules: sugars, nucleotides, amino acids, and fatty acids. The first three are classes of molecules that can
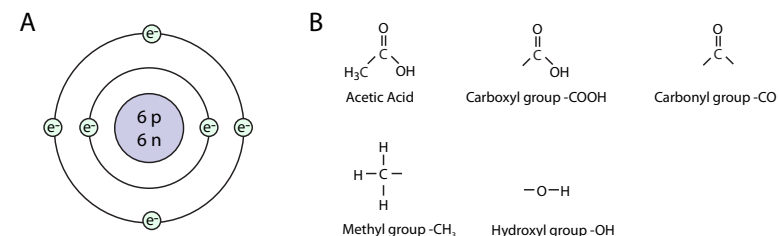

Figure 4. (A) The carbon atom has four electrons in its outer shell. (B) Functional groups that can be identified from a molecule of acetic acid.

be strung together by covalent bonds to make important large biomolecules: simple sugars can form large polysaccharides such as starch, cellulose, or glycogen, nucleotides can form RNA (ribonucleic acids) or DNA (deoxyribonucleic acids), and amino acids can form proteins. Fatty acids, on the other hand, are acid derivatives of long chains of carbons linked to one another, with hydrogens taking up most of the other bonding positions.

Sugars

Sugars, and glucose in particular, are important molecules for cells because they are the primary energy source. Sugars have the general chemical formula CH_2O, and can be joined together almost infinitely for storage. However, because they are hydrophilic, they allow water molecules to intercalate between them, and cannot pack as efficiently as fats, which are hydrophobic and thus exclude water. On the other hand, the sugars can be mobilized for use more quickly. Therefore, polysaccharides are usually short-term reservoirs of energy for an organism, while fats are used for longer-term storage.

The general chemical formula cannot fully define a particular sugar, because the same set of atoms, e.g. $C_6H_{12}O_6$ can refer to glucose, fructose, mannose, or galactose, and that doesn't even include the stereoisomers. Isomers are rearrangements of the same atoms, such as with glucose and fructose (fig. 5), while stereoisomers are much more similar: they are mirror-images of one another. Thus glucose can exist as L-glucose or D-glucose, depending on whether it is a "left-handed" or "right-handed" isomer. This may seem like an esoteric distinction, but it becomes important in intermolecular interactions, because many are based on recognition of specific shapes, so an L-conformation molecule may not be recognized by an enzyme that recognizes its D- isomer.

Another important aspect of sugar chemistry is whether it is an aldose or a ketose, based on the type of carbonyl group it carries. This is easiest to understand looking at the position of the carbonyl group in the linear structure: put simply, an aldehyde is a terminal carbonyl group, while a ketone is an internal carbonyl group. Sugars in aqueous solution exist in an equilibrium between the linear form and the ring form, which is formed by intramolecular attack by a hydroxyl group on the carbonyl. Technically, the cyclic sugar is a pyranose (6-membered ring) or a furanose (5-membered ring), so that D-glucose cyclizes into D-glucopyranose. However, in most cell biology courses, the cyclic sugar will still be referred to as its non-cyclic alter ego. Note that due to the difference between the $C_6H_{12}O_6$ aldose glucose, and the $C_6H_{12}O_6$ ketose fructose, cyclization generates a pyranose in one case, and a furanose in the latter (fig. 5), although the
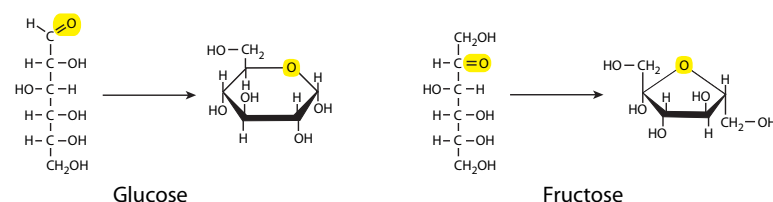

Glucose   Fructose

Figure 5. Glucose is an aldose (terminal carbonyl) that cyclizes into a pyranose, fructose is a ketose (internal carbonyl) that cyclizes into a furanose.

number of carbons (and other atoms) are the same. These two molecules are therefore recognized differently by the enzymes of the cell, leading to different metabolic pathways.

Simple sugars can be joined together by *condensation reactions* to form glycosidic bonds. These reactions are called condensation reactions because they form water as a by-product. The glycosidic bond is an —O— linkage between carbons of two sugars. The bond is usually named with the specific linkages: for example in cellulose, glucoses are linked by β(1,4) linkages, which means in a standard ring diagram, the upward-facing β-hydroxyl on the 1-carbon interacts with the —OH on the 4-carbon of a neighboring glucose (fig. 6B). [Technically, since only two glucoses are shown here, this is a molecule of cellobiose, not cellulose.] In contrast, the maltose shown in the same figure (fig. 6A), while also showing two glucoses linked together, is an α(1,4) linkage, with a downward-facing α-hydroxyl on the 1-carbon.

Large polysaccharides generally have one of two functions: as a very strong structural component of a cell, and as a storage molecule for readily accessible energy. The two major structural polysaccharides made by cells are *cellulose* and *chitin*. Cellulose is primarily synthesized by plants, while chitin is mostly synthesized by invertebrates (think crab shells), though it is also made by many fungi and algae. As we just saw, cellulose is an array of parallel lengths of glucose monomers joined together by β(1,4) glycosidic bonds (fig. 7). These long glucans are stacked closely on one another so that many H-bonds can form along their lengths, which are virtually limitless, determined by the needs of the organism. Interestingly, chitin is also a homopolymer linked by β(1,4) glycosidic bonds, but instead of glucose, the monosaccharide used is N-acetylglucosamine (often abbreviated GlcNAc, see chapter 11). However, the macromolecular structure is very similar to cellulose, and like cellulose, it is very strong.

As with structural polysaccharides, there are also two primary energy-storage polysaccharides: *starch*, which is synthesized by plants, and *glycogen*, which is synthesized by animals. Starch is actually a mixture of two slightly different polysaccharides. One is α-amylose, which is a glucose homopolymer like cellulose, but connected by α(1,4) glycosidic linkages, which makes it completely different structurally. Unlike the linear and highly stackable cellulose polysaccharides, α-amylose takes on a twisting α-helical shape. The other starch polysaccharide is amylopectin, which is like α-amylose with the addition of branches formed from α(1,6) glycosidic bonds every 24-30 residues (fig. 6C). The storage polysaccharide for animals, glycogen, is essentially amylopectin with a higher frequency of branching, approximately every 8-14 residues. Whereas the tight packing of the structural polysaccharides renders them waterproof, this is certainly not the case for starch or glycogen, both of which can interact with many water molecules
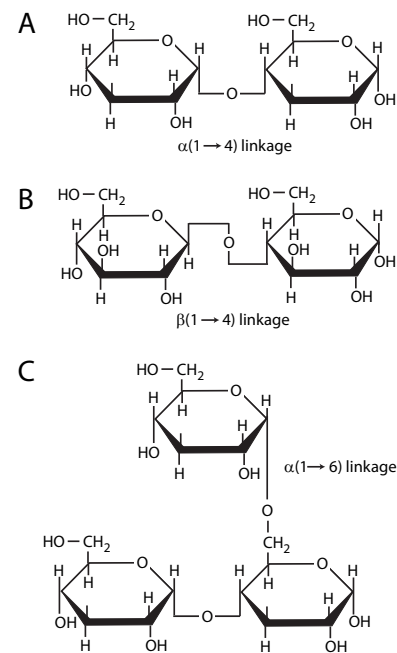

Figure 6. (A) the α(1,4) glycosidic bond of maltose, (B) the β(1,4) bond of cellobiose, and (C) the α(1,6) bond in branching glycogen.
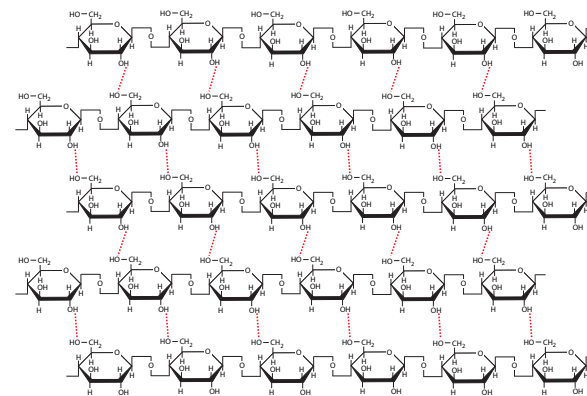

Figure 7. Cellulose is a very strong material due to the many hydrogen bonds (in red) possible when strands of β(1,4)-linked glucoses are aligned.

simultaneously, and swell up with the hydration, as any cook who has ever made a pudding (the thickening ingredient is starch from corn) can attest.

Nucleotides

Nucleotides, the building blocks of RNA and DNA, are themselves composed of a pentose sugar attached to a nitrogenous base on one side and a phosphate group on another. The sugar is either the 5-carbon sugar ribose or its close cousin, deoxyribose (the "deoxy" refers to a "missing" hydroxyl group on the 2-carbon, which has an H instead). The attached nitrogenous base can be a *purine*, which is a 6-member ring fused to a 5-member ring, or a *pyrimidine*, which is a single 6-membered ring. These bases are usually adenine (purine), guanine (purine), thymine (pyrimidine), and cytosine (pyrimidine) for DNA, with a substitution of uracil for thymine in RNA bases. However, there are also some unconventional and modified bases that show up in special situations, such as in tRNAs. In addition to being the monomer components of DNA and RNA, nucleotides have other important functions as well. The best known, adenosine triphosphate, or ATP, is the primary "instant" energy source for the cell by the energy released through hydrolysis of its terminal phosphate group.

DNA or RNA are built from nucleotides through linkages of the sugars, and the polymerization occurs by condensation reactions, but these bonds are not glycosidic bonds like with polysaccharides. Instead, bonds form between the 5' phosphate group of one nucleotide and the 3' hydroxyl group of another. These are phosphodiester bonds, and a quick glance at the structure (fig. 8) explains the naming: an ester bond is a carbon-oxygen linkage, and the phosphodiester bond is a C-O-P-O-C, so there are two esters with a phosphorus linking them. With the purine or pyrimidine base on the 1-carbon, this arrangement places the bases on the opposite side of the sugar from the polymerizing phosphodiester bonds. This forms a sugar-phosphate backbone to the DNA/RNA, which then has the bases projecting out from it.

The bases will then likely interact with the bases of other nucleotides, whether part of another nucleic acid strand or free-floating. Not only do they interact, but they interact with great specificity and consistency: adenines base-pair with thymines (or uracils) through two hydrogen bonds, while guanines interact with cytosine through three H-bonds. Note that while one extra hydrogen bond does not appear to be particularly significant, the attraction between G-C is 50% stronger than between A-T, and over long stretches of DNA, areas high in G-C content are significantly more difficult to unzip (separate strands) than areas high in A-T pairs. This specific base-pairing, known as *Chargaff's rules*, is the basis for life: base-pairing is needed to make DNA double-
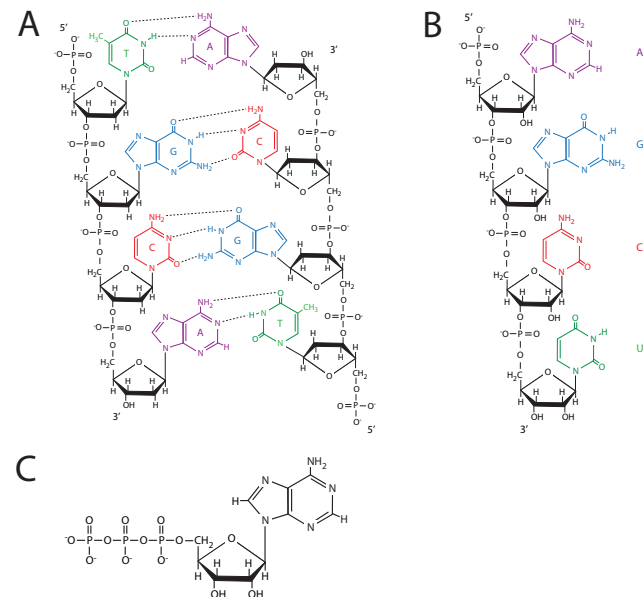

Figure 8. (A) DNA and (B) RNA differ by the presence of —OH on the 2-carbon of ribose but not deoxyribose and the use of uracil in RNA instead of thymine. Both are constructed from nucleotides like adenosine triphosphate (C).

stranded which gives an organism a built-in backup of genetic information and it is also the basis for transforming that information into proteins that form the bulk of a cell.
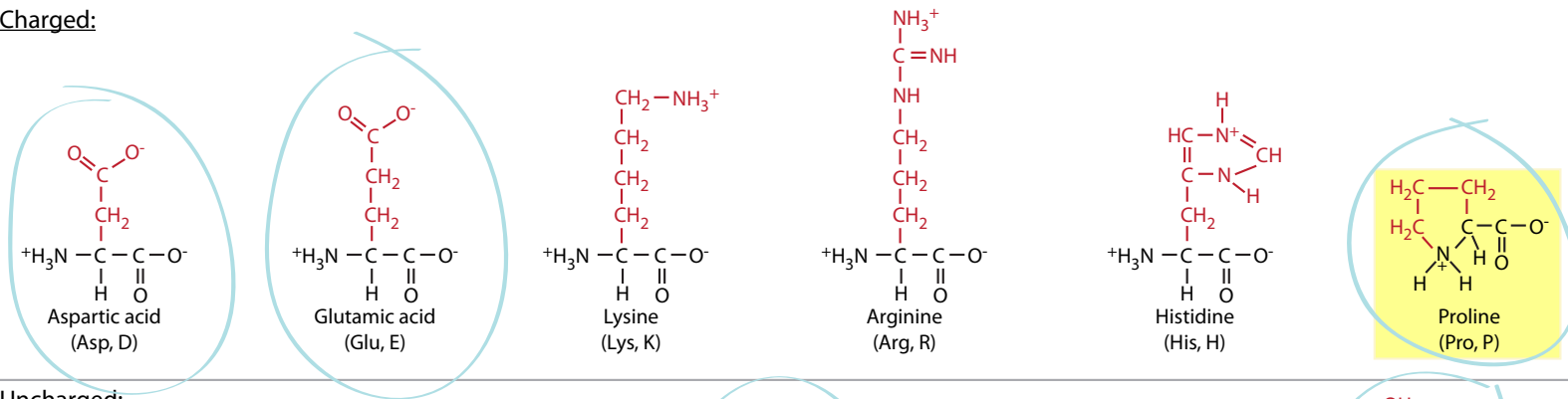
Nucleic acids, the long polymers of nucleotides, exist in either single or double stranded forms in vitro. However, in the cell, most RNA is single-stranded, and most DNA is double-stranded. This difference is important to their function: RNA is a temporary information transfer molecule for a particular gene, DNA is the permanent repository of all genetic information needed to make an organism. Therefore, RNA needs to be easily read, meaning that the bases need to be accessible, and not locked to a complementary strand. Its long-term stability is not particularly important because when it is made, usually many copies are made at the time, and it is only needed while the cell needs to make the protein it encodes. Conversely, the same strand of DNA is read over and over to make the RNA, and since there are only two copies of each chromosome (a chromosome is a single double-stranded DNA molecule) in a cell, the ability to maintain the integrity of the DNA is crucial. Because of base pairing, each strand of DNA contains all the information necessary to make a complete exact copy of its complementary strand.

Of course, the point of the genetic information in DNA is to encode the production of proteins that can then carry out the functions that define cellular life. Some of those functions, such as DNA replication, gene regulation, transcription, and translation, require the proteins to interact with a nucleic acid. Usually, part of the recognition process involves apposition of a positively charged region of the protein to the DNA (or RNA), which is a very negatively charged molecule, as expected from all the phosphates in the sugar-phosphate backbone. RNA, but not DNA (with some exceptions), can also interact with itself by complementary base-pairing. If a stretch of RNA sequence comes into contact with a stretch of RNA with a complementary sequence on the same molecule, then base-pairing can occur. Depending on the number of nucleotides between the complementary areas, secondary structures such as stem-and-loops and hairpins can form.
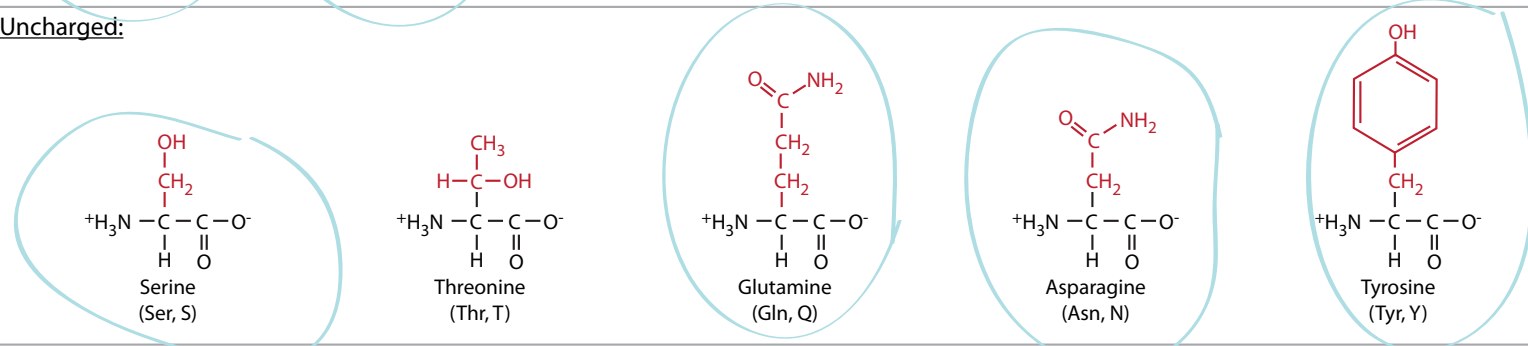
Amino Acids

Most of the major molecules of the cell - whether structural, like cellular equivalents of a building's girders and beams, or mechanical, like enzymes that take apart or put together other molecules, are proteins. Proteins interact with a wide variety of other molecules, though any given interaction is usually quite specific. The specificity is determined in part by electrical attraction between the molecules. So, what determines the charge of different regions of a protein?
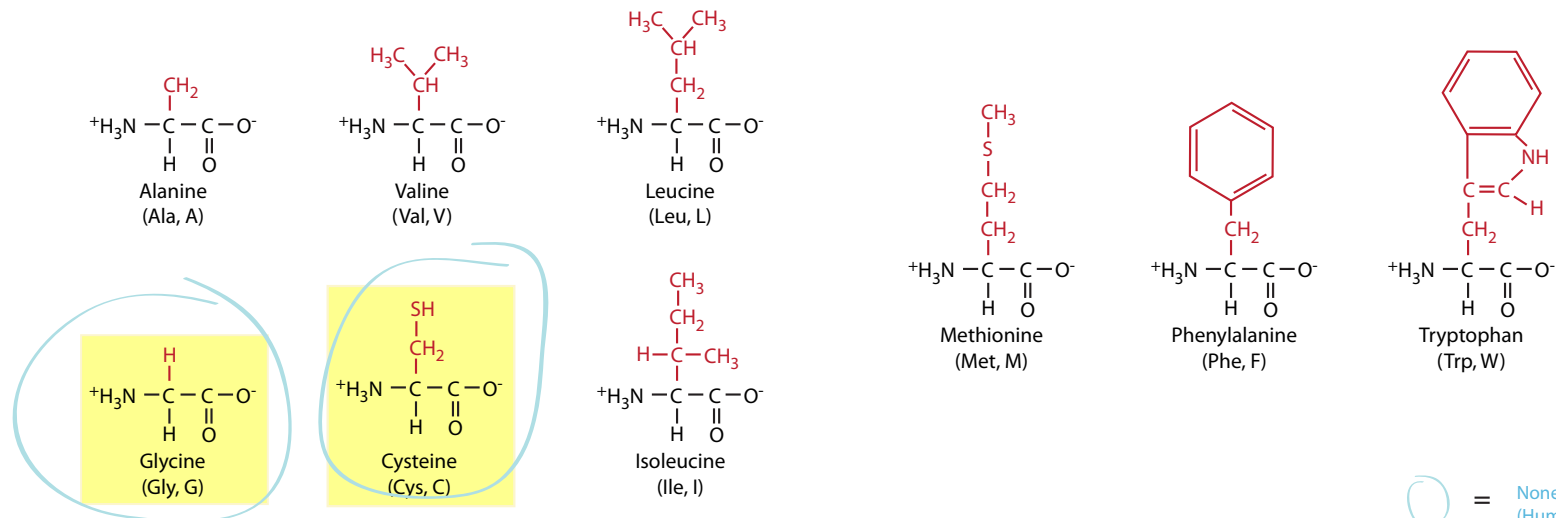
Polar Charged:

Aspartic acid (Asp, D) · Glutamic acid (Glu, E) · Lysine (Lys, K) · Arginine (Arg, R) · Histidine (His, H) · Proline (Pro, P)

Polar Uncharged:

Serine (Ser, S) · Threonine (Thr, T) · Glutamine (Gln, Q) · Asparagine (Asn, N) · Tyrosine (Tyr, Y)

Nonpolar:

Alanine (Ala, A) · Valine (Val, V) · Leucine (Leu, L) · Methionine (Met, M) · Phenylalanine (Phe, F) · Tryptophan (Trp, W) · Glycine (Gly, G) · Cysteine (Cys, C) · Isoleucine (Ile, I)

= Nonessential a.a. (Human)

Figure 9. The Amino Acids. The backbone is shown in black, while the side chains are colored red. Amino acids circled in blue can be synthesized by humans, while the uncircled amino acids must be ingested. Amino acids with a yellow background have unique structural considerations: the extremely small side chain of glycine allows it to fit into tight spaces, the sulfhydryl group of cystein allows the formation of disulfide bonds, and the cyclic structure of proline introduces a forced bend in the polypeptide chain.

Amino acids (figure 9), which are joined together to make proteins, may be positively charged (basic), negatively charged (acidic), polar, or nonpolar, based on the characteristics of their side chains. The charge on the amino or carboxyl end of each amino acid does not play a role in the overall character of any particular region of the protein, because they are effectively neutral, having been linked, the amino group of one amino acid to the carboxyl group of another, by a peptide bond. Note the figure of the amino acid: it is one carbon, called the α carbon, linked to amino and carboxyl groups on opposite sides, and to hydrogen, and a side chain, denoted by R. These side chains, of which there are twenty common ones, can be as simple as a hydrogen atom (glycine), or could be quite complex, involving extended ring structures (histidine, phenylalanine). The variety in their size, shape, and charge all add up to an extremely versatile set of building blocks for some of the most important working molecules of the cell. In the cell, a peptide bond is formed between two amino acids with enzymatic help from the ribosome. Like the previous two polymerizing reactions, formation of peptide bonds is a condensation reaction in which the carbon of the carboxyl group and the nitrogen from the amino group of their respective amino acids are bonded together (fig. 10). This is a very stable bond due to resonance of the amide group. In the cell, peptide bonds are mostly nonreactive, except when attacked by proteolytic enzymes.

Figure 10. Peptide Bond Formation. A condensation reaction between the carboxyl group of alanine and and the amino group of valine generates a peptide bond linking the two amino acids, with a molecule of water as a byproduct.

Alanine

Valine
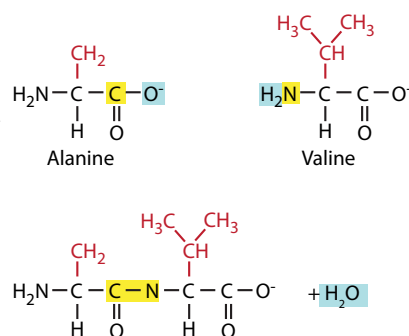
A peptide is an inexact term used for relatively few (usually <30) amino acids joined together. Each amino acid in a polypeptide or protein may also be referred to as a "residue" which can sometimes be confusing because the same term is also applied to monomers of nucleic acids and of polysaccharides. Larger polymers are known as polypeptides or as proteins, although polypeptide has more of a structural connotation and may be used to indicate an unfinished or not-yet-functional state, whereas protein generally implies some physiological function. On of the key characteristics of proteins is the ability to form secondary, tertiary, and for proteins, quaternary structure by means of specific folding patterns. If you think of a long piece of thread, yarn, or rope, you

Chirality

Almost all amino acids (glycine is the exception) are optically active, which means that they are asymmetric in such a way that it is impossible to superimpose the original molecule upon its mirror image. There is a "handed-ness" about them, much as your right hand cannot be superimposed on your left hand if both palms must face the same direction. In fact, in the figure here, you can also understand why glycine is an exception, since its R-group is a simple hydrogen atom.
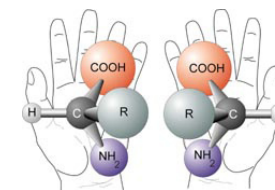

figure released to public domain by NASA

Chiral pairs, or enantiomers, not only have the same atomic components like all isomers, they also have the same bonds and bond order. The term "optically active" comes from the discovery that polarize light is rotated in different directions by enantiomers. Amino acids are often labeled as either D- (dextrorotatory) or L- (levorotatory) depending on their atomic configuration in relation to the enantiomers of glyceraldehyde. This is a common naming system, but not always logical, in that almost half of the L-amino acids are in fact dextrorotatory (clockwise rotation of light), but their molecular configurations resemble the levorotatory isomer of glyceraldehyde.

Ribosome-created proteins and peptides are all constructed with L-amino acids. However, D-amino acids do exist in nature, and can be incorporated into peptides through non-ribosomal means. An excellent example is found in the cell walls of some bacteria. Because most proteolytic enzymes only act on proteins with L-amino acids, the incorporation of D-amino acids into the cell wall can protect the bacteria from harm. These D-amino acids are incorporated by transpeptidase. Transpeptidase is also the target of the antibiotic, penicillin, which is an irreversible inhibitor of that enzyme.

can probably imagine an infinite number of different ways to arrange it, from spirals to loops to random tangles. This is essentially what can happen with a protein with the constraints put upon it by the size and charge of the amino acids that compose it.

The *primary structure* of a protein is simply the sequence of amino acids that compose the protein. These amino acids are joined by peptide bonds from the carboxyl terminal of one amino acid to the amino terminal of the next. Secondary structure refers to the localized, simple, shapes that can be formed, such as alpha-helices, or beta-sheets. These come about primarily through hydrogen bonding to nearby (relative to the primary structure) residues.

Tertiary structure is 3-dimensional structure that is built upon arrangements of second-
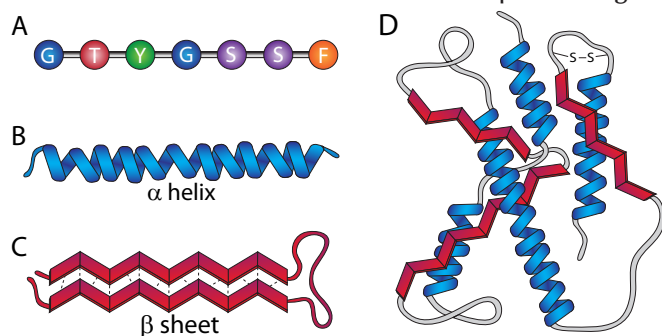

Figure 11. (A) Primary structure, (B) Secondary structure: an α-helical region, (C) Secondary structure: a β-pleated sheet region, (D) Tertiary structure.

ary structures, often through disulfide bonds and hydrophobic interactions in addition to hydrogen bonding. In the context of structural stability, cysteine plays a special role. Beyond the primary structure, most protein folding is held in place by hydrogen bonds. Although strong enough in most situations, they can be disrupted without extraordinary energy. Disulfide bonds (−S−S−) are covalent bonds that form between the sulfhydryl groups of two cysteines that effectively locks the local protein structure in place, making the protein extremely stable.

Finally, quaternary structure is the arrangement of different individual polypeptides (subunits) into a functional protein. Obviously, only multi-subunit proteins have a quaternary structure.

Not surprisingly, the proteins of thermophilic archaebacteria such as Thermophilus aquaticus (Taq) or Pyrococcus furiosus (Pfu) have a high proportion of cysteines and disulfide bonds, since they live in deep sea volcanic ocean vents under high pressure and temperatures.
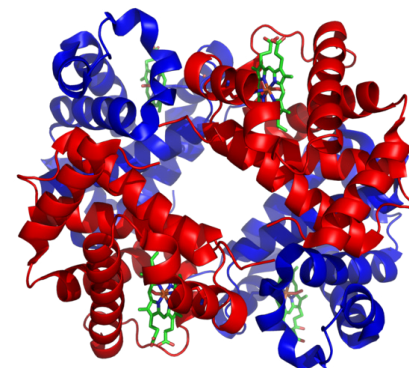

Figure 12. Quaternary structure is illustrated here by hemoglobin, which is composed of four independent polypeptide subunits that come together in a specific conformation to make a functional hemoglobin protein.

Fatty Acids

Unlike monosaccharides, nucleotides, and amino acids, fatty acids are not monomers that are linked together to form much larger molecules. Although fatty acids can be linked together, for example, into triacylglycerols or phospholipids, they are not linked directly to one another, and generally no more than three in a given molecule. The fatty acids themselves are long chains of carbon atoms topped off with a carboxyl group. The length of the chain can vary, although most are between 14 and 20 carbons, and in higher order plants and animals, fatty acids with 16 and 18 carbons are the major species. Due to the mechanism of synthesis, most fatty acids have an even number of carbons, although odd-numbered carbon chains can also be generated. More variety can be generated by double-bonds between the carbons. Fatty acid chains *with no double bonds are saturated*, because each carbon is saturated with as many bonded hydrogen atoms as possible. Fatty acid chains with double bonds are unsaturated (fig. 13). Those with more than one double bond are called polyunsaturated. The fatty acids in eukaryotic cells are nearly evenly divided between saturated and unsaturated types, and many of the latter may be polyunsaturated. In prokaryotes, polyunsaturation is rare, but other modifications such as branching and cyclization are more common than in eukaryotes. A table of common fatty acids is shown below.
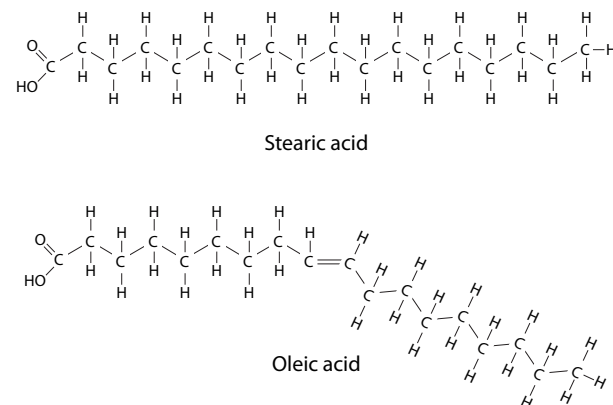

Figure 13. Fatty acids. (Top) Stearic acid is a fully saturated fatty acid with no carbon-carbon double bonds. (Bottom) Oleic acid is an unsaturated fatty acid.

Myristic Acid	14:0 (14 carbons, no double bonds)
Palmitic Acid	16:0
Stearic Acid	18:0
Arachidic Acid	20:0
Palmitoleic Acid	16:1
Oleic Acid	18:1
Linoleic Acid	18:2
Arachidonic Acid	2:4

There are significant physical differences between the saturated and unsaturated fatty acids due simply to the geometry of the double-bonded carbons. A saturated fatty acid is very flexible with free rotation around all of its C-C bonds. The usual linear diagrams and formulas depicting saturated fatty acids also serve to explain the ability of saturated fatty acids to pack tightly together, with very little intervening space. Unsaturated fatty acids, on the other hand are unable to pack as tightly because of the rotational constraint impoarted by the double bond. The carbons cannot rotate around the double bond, so there is now a "kink" in the chain. Generally, double-bonded carbons in fatty acids are in the cis- configuration, introducing a 30-degree bend in the structure.
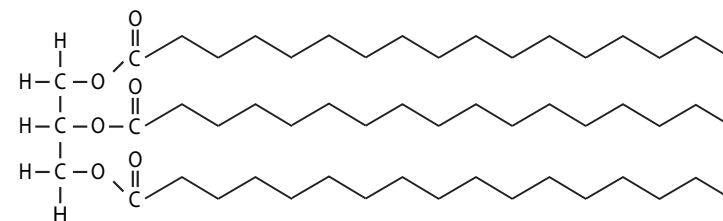

Figure 14. Triglycerides. These lipids are formed by conjugation of a glycerol to three fatty acyl chains through ester bonds from each glycerol oxygen.

Fatty acids inside cells are usually parts of larger molecules, rather than free acids. Some of the most common lipids derived from fatty acids are triacylglycerols, phosphoglycerides, and sphingolipids. Triacylglycerols, as the name implies, is three fatty acid (acyl) chains connected to a glycerol molecule by ester bonds (fig. 14). Triacylglycerols, also known as triglycerides, may have fatty acids of the same (simple triacylglycerols) or varying types (mixed triacylglycerols). Mixtures of these are the primary long-term energy storage molecules for most organisms. Although they may be referred to colloquially as fats or oils, the only real difference is the degree of saturation of their constituent fatty acids. Mixtures with higher percentages of saturated fatty acids have a higher melting point and if they are solid at room temperature, they are referred to as fats. Triacylglycerol mixtures remaining liquid at room temperature are oils.

In human medicine, a common test for heart disease risk factors is measurement of triglyceride levels in the blood. Although various cell types can make and use triglycerides, most of the triglycerides in people are concentrated in the adipose tissue, which is made up of adipocytes, or fat cells, though liver is also a significant fat store. These cells have specialized to carry fat globules that take up most of the volume of the cell. When triglyceride levels in the blood are high, it means that fat is being produced or ingested faster than it can be taken up by the adipocytes.

Phospholipids (also called phosphoglycerides or glycerophospholipids), are also based on attachment of fatty acids to glycerol. However, instead of three fatty acyl tails, there are only two, and in the third position is a phosphate group (fig. 15). The phosphate group also attaches to a "head group" . The identity of the head group names the molecule, along with the fatty acyl tails. In the example figure, 1-stearoyl refers to the stearic acid on the 1-carbon of the glycerol backbone; 2-palmitoyl refers to the palmitic acid on the 2-carbon of the glycerol, and phosphatidylethanolamine refers to the phosphate group and its attached ethanolamine, that are linked to the glycerol 3-carbon. Because of the negatively-charge phosphate group, and a head group that is often polar or charged, phospholipids are amphipathic - carrying a strong hydrophobic character in the two fatty acyl tails, and a strong hydrophilic character in the head group. This amphipathicity is crucial in the role of phospholipids as the primary component of cellular membranes.

Sphingolipids (fig. 16) are also important constituents of membranes, and are based not upon a glycerol backbone, but on the amino alcohol, sphingosine (or dihydrosphingosine). There are four major types of sphingolipids: ceramides, sphingomyelins, cerebrosides, and gangliosides. Ceramides are sphingosine molecules with a fatty acid tail attached to the amino group. Sphingomyelins are ceramides in which a phosphocholine or phosphoethanolamine are attached to the 1-carbon. Cerebrosides and ganglio-
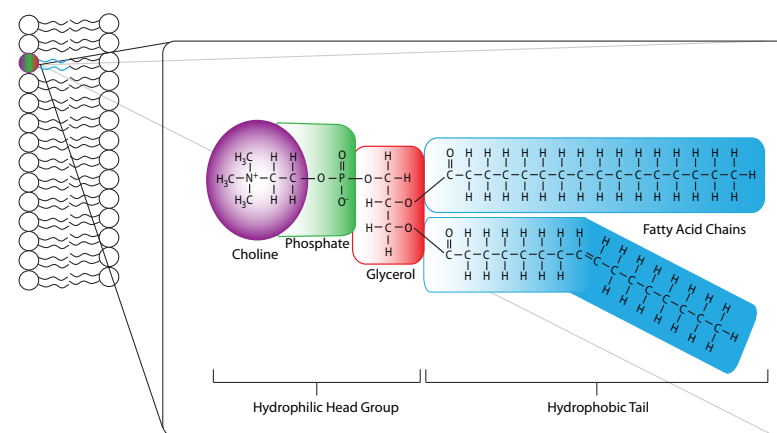

Figure 15. A phospholipid: the glycerol backbone (red) connects to two fatty acids and to a phosphate and polar head group.
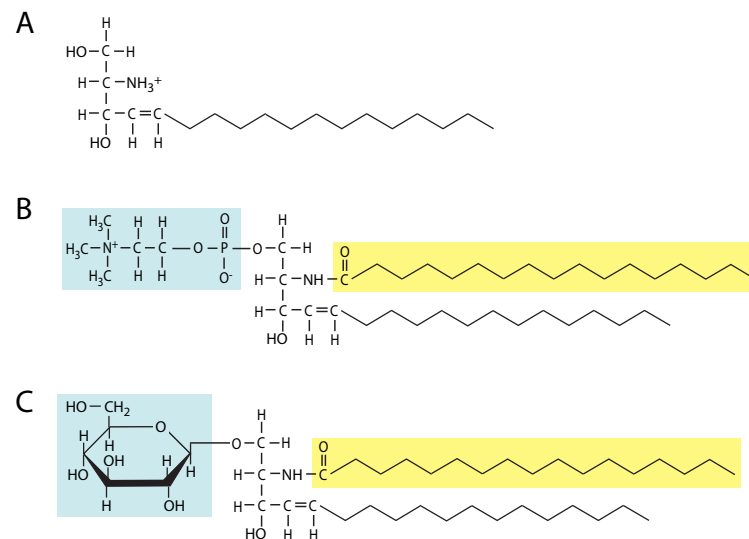

Figure 16. Sphingolipids are based on the amino alcohol, sphingosine (A). Ceramides have a fatty acid tail attached, and a ceramide with a phosphocholine head group is a sphingomyelin (B). If the head group is a sugar, then the molecule is a cerebroside. (C).

sides are glycolipids - they have a sugar or sugars, respectively, attached to the 1-carbon of a ceramide. The oligosaccharides attached to gangliosides all contain at least one sialic acid residue. In additional to being a structural component of the cell membrane, gangliosides are particular important in cell to cell recognition.

Lipids are vaguely defined as biological compounds that are insoluble in water but are soluble in organic solvents such as methanol or chloroform. This includes the fatty acid derivatives listed above, and it includes the final topic for this chapter, cholesterol. Cholesterol (fig. 17) is the major biological derivative of cyclopentanoperhydrophenan-threne, a saturated hydrocarbon consisting of four fused ring formations. It is an important component of plasma membranes in animal cells, and is also the metabolic precursor to steroid hormones, such as cortisol or β-estradiol. Plant cells have little if any cholesterol, but other sterols like stigmasterol are present. Similarly, fungi have their particular sterols. However, prokaryotes do not, for the most part, contain any sterol molecules.
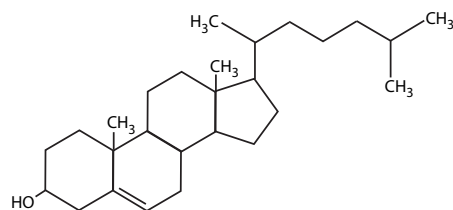

Figure 17. Cholesterol is an important lipid both as a membrane component and as a steroid precursor.