

# MACROMOLECULES, Part 1\*

## Carbohydrates, Lipids, and Nucleic Acids

**Introduction:** Living organisms are unique in being composed of long, massive molecules called macromolecules. Biochemistry, the study of the chemistry of life, is the study of the structure, function and interactions of the four types of macromolecules: carbohydrates, lipids, proteins and nucleic acids. All macromolecules are **polymers** (chains) of repeating links called **monomers**.

- Most different macromolecule types (carb., protein, nucleic acid) can be distinguished by their monomeric units. The exceptions are the lipids. These are not as large and complex as the other macromolecules, can be distinguished from the others based on their high ratio of C and H to O.
- Differences within general types of macromolecules (*i.e.*, between proteins, carbohydrates, fats and nucleic acids) result from variation in the identity, sequence, and bonding of their monomeric units.

**1. Lipids:** Lipids are compounds consisting almost entirely of carbon and hydrogen. As a result they are highly non-polar; likewise they are **highly reduced**<sup>1</sup>. Even in their most complex forms, they are not as large as some of the compounds that we will consider later and so perhaps we should not really consider them as "macromolecules". Lipids perform a variety of functions in organisms. These include:

- they are major constituents of cellular membranes
- they are the most important form of long-term energy storage
- some types of lipids are used as hormones
- in larger organisms they can be used to help waterproof the integument, provide thermal insulation, or act as cushions (amusing as this is to us!). The first two of these directly trace to their chemical properties.

We will divide lipids into two sub-groups:

- fats -- compounds primarily based on the presence of fatty acids. More about these in a moment
- sterols -- four ring compounds built up from groups called isoprenes.

Fats: the fundamental constituents of fats are fatty acids. These typically are a chain of 16 to 18 carbons with a carboxylic acid group at one terminal (where the "acid" term arises). Fatty acids come in two general types; saturated and

---

\* © 2006 by K.N. Prestwich and W.R. Healy, Dept. Biology, College of the Holy Cross, Worcester, MA 01610

<sup>1</sup> You probably remember a bit about oxidation and reduction from high school chemistry. For our purposes, a biological compound is reduced when it contains mostly C-C and C-H bonds -- *i.e.*, it contains little oxygen. Reduced compounds contain relatively large amounts of potential energy. We'll say a bit more about what it means when we say a compound is "reduced" or oxidized when we consider cellular respiration (so-called "catabolism" and "anabolism").





When one fatty acid is attached to glycerol, the resulting compound is a **monoglyceride**, if two a **diglyceride**, and three makes (surprise) a **triglyceride**. Fatty tissue (energy stores) is typically made up of di- and tri-glycerides.

On the other hand, if the fat consists of a mono- or di-glyceride with a phosphate group attached to the glycerol in the same place where a fatty acid could have been attached, the resulting compound is termed a **phospholipid**. Such compounds are the central constituents of cell membranes. Please look at your text to see an example of a phospholipid. Additionally, in mammals, they are very important in making it easy for us to inflate the alveoli in our lungs. But that's another story for another course.

**Sterols:** Do take a look at the general formula for a sterol (cholesterol in this case) in the textbook (Fig. 5.15). The main thing you should remember about them for the moment is that:

- they are important constituents of the membranes of animals and
- many hormones are sterols. Examples are: estrogen, progesterone, testosterone (all sex hormones); aldosterone (important in regulating blood volume and blood pressure), and cortisol (generally important for all of metabolism).

Most of us consume more cholesterol than is good for us. Although we need these compounds, we are more than capable of constructing all we need from compounds found in food such as carrots! Note that there is considerable variation in individuals in terms of the amount of cholesterol that they normally synthesize. In some rare cases, the individuals normally make so much cholesterol that it puts them at very high risk of cardiovascular disease, even early in life and even if they pursue a low cholesterol diet. Recently a number of very effective drugs have been developed to help control cholesterol levels when diet fails.

**2. Carbohydrates:** Compared to lipids, carbohydrates contain a considerable amount of oxygen and thus are said to be more oxidized. As a result they contain less potential energy per gram than fats.

#### **Potential Energy and Oxidation -- Fats and Carbohydrates Compared**

In fact the difference in energy content is quite significant. When one gram of lipid is burned completely with  $O_2$  to yield  $CO_2$  and  $H_2O$  around 37.7 kJ of energy are released. By contrast, if one gram of dry carbohydrate is burned completely with  $O_2$ , the result is that only about 1/2 as much energy (about 18.8 kJ is released!

Moreover, since carbs contain a lot of oxygen, which has a very different electronegativity than C and H, they have many polar regions where they can form hydrogen bonds with other polar molecules. These are **hydroxyl** (most common) and **carbonyl** (as in **aldehydes** and **ketones**) groups. The result is that overall these molecules are good at forming H-bonds with other polar

substances, especially water. So, unlike lipids, carbohydrates are quite soluble and are nearly always found to have considerable amounts of water attached. We say these compounds are highly **hygroscopic**.

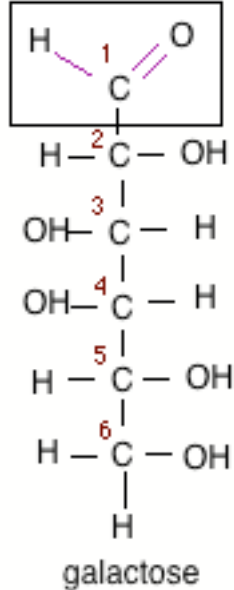
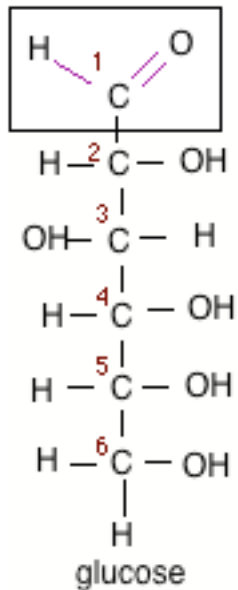
**Carbohydrate Chemical Formula:** As you almost certainly already know, the general formula for all carbohydrates is  $(\text{CH}_2\text{O})_n$  where **n** refers to the some number of "hydrated" carbons. For example, if  $n = 6$  you have an example of a hexose (6 C sugar) with a chemical formula,  $\text{C}_6\text{H}_{12}\text{O}_6$  (.g., glucose, fructose); if  $n = 5$  you have a pentose (5 C sugars, e.g., ribose) with the general formula  $\text{C}_5\text{H}_{10}\text{O}_5$ .

Unlike the case with fatty acids, there are many possible ways to arrange the atoms in a carbohydrate with a given chemical formula. The result is a bewildering number of compounds with the same chemical formula. However, recall that because they have different shapes and arrangements of reactive groups and polar molecules, their properties are different. Some of the many different possible arrangements (**isomers**) are covered in the text; be sure to review them (see text study questions).

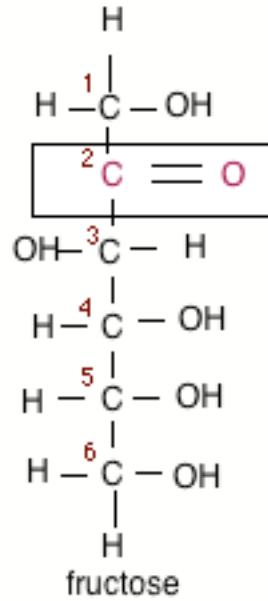
**Types of Carbohydrates** Carbohydrates exist either as individual units (**monosaccharides**) or they may be linked together to form larger molecules. Let's start with monosaccharides, also called simple sugars. Most simple sugars found in organisms are hexoses. There are also lesser amounts of pentoses.

Hexoses all have the formula  $\text{C}_6\text{H}_{12}\text{O}_6$  and therefore all contain essentially the same amounts of potential energy. On the other hand, there are many ways to arrange the 24 atoms found in hexoses. Two basic patterns exist with respect to the position of the carbonyl group. If it is on one of the ends of the molecule, the carbonyl is an aldehyde and the sugar is called an **aldose**; glucose and galactose are examples. Although they are both hexose aldoses, they also differ from each other in terms of the arrangement of their hydroxyl groups. These alternative arrangements, called **optical isomers**, have some noticeable differences in properties, especially in terms of how they are handled chemically within an organism. Glucose and galactose are optical isomers; they taste different and are metabolized somewhat differently. This traces to the specificity of the proteins that act as taste receptors or metabolic enzymes. In other sugars, the carbonyl is internal and thus is a ketone. These sugars are called **ketoses**; fructose is an example.

### Hexose Aldoses

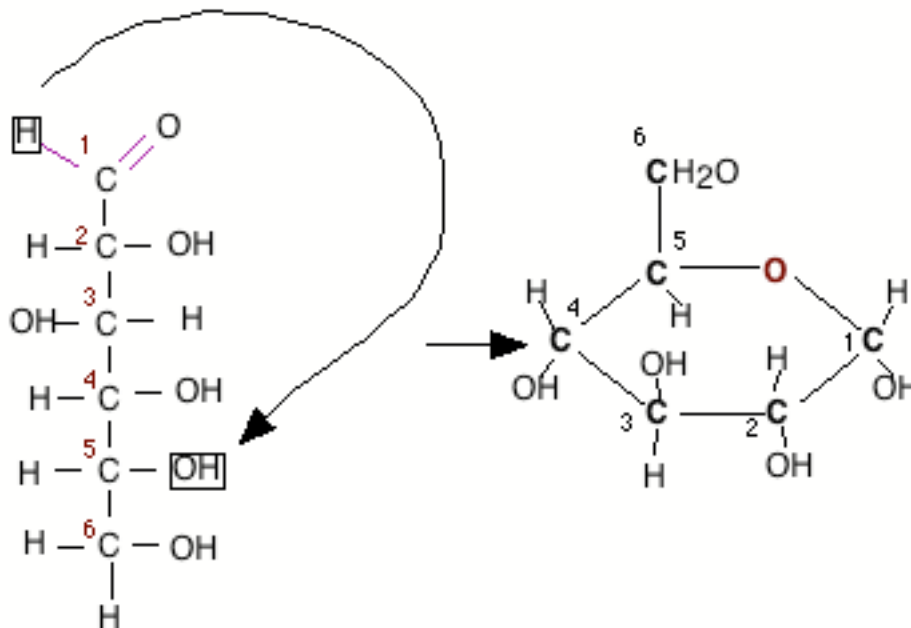


### Hexose Ketose



In most cases it is easy for the body to isomerize one hexose into another (but it takes an enzyme to direct each such conversion). One other thing to note is that the carbon atoms in the main "skeleton" of organic molecules are typically numbered (see previous diagram). Among other things, this makes it easy to keep track of where bonds form, as we will see in a moment.

One last point about simple sugars. The previous diagrams imply they are all linear structures. In fact, **sugars generally exist as rings**. The carbonyl group typically reacts with a hydrogen on the number 5 carbon. If the original compound was an aldose, the result is a six membered ring consisting of 5 carbons (#s 1 to 5) and one oxygen. The #6 carbon projects off the ring:



On the other hand, if the compound is a ketose, the result is typically a five membered ring.

Understand where these ring structures come from and be aware of the fact that hexoses can come in 5 and 6 membered rings. Don't memorize the details of the reaction. That's for biochemistry.

**Pentoses** on the other hand are used primarily in molecules like DNA, RNA and as constituents of high-energy compounds such as ATP and GTP or as signal molecules such as 3',5' cyclic AMP. are the units from which more complex carbohydrates. In all cases they can be seen as being a part of an overall structure. For instance, they and phosphate together form the continuous strand that makes up most of a molecule of DNA or RNA.

**Di- and Polysaccharides:** It is common for **hexoses** to polymerize by the same dehydration reaction that we saw earlier. The result is larger and larger carbohydrates -- **disaccharides** and especially the **polysaccharides** such as **starches** and **chitin**. Polysaccharides tend to be used as:

- energy storage molecules (primarily in the starch form)
- structural molecules (for example cellulose and chitin)
- identification molecules (as part of mucco-polysaccharides in cell membranes)

Two or more monosaccharides can be linked together via an **ether bond (C-O-C)**. In carbohydrates, this bond is given the name **glycosidic linkage**. The properties of the resulting molecule will be determined by:

- the type of sugars on the bond

- the linear order of the sugars
- the types of bonds that are formed between the monosaccharide units (the carbons involved, the actual shape of the bond).

Thus, two hexoses, glucose and fructose, linked together between their number 1 and 2 positions gives the disaccharide sucrose. Or, maltose, another disaccharide is formed when two molecules of glucose are linked via their 1 and 4 positions. Note that in both cases the resulting molecules have the formulas  $C_{12}H_{24}O_{12}$  and they contain similar amounts of potential energy. However, a number of their properties are quite different. For instance, sucrose tastes quite sweet while maltose is not especially so.

When more sugars are linked together the result is a polysaccharide. Here are a number of important polysaccharides:

**1. Built entirely from glucose:**

- If all links are 1 to 4; the result is the common plant starch **amylose**
- If most links are 1 to 4 but some are 1 to 6, the result is **amylopectin**, also a plant starch. The 1 to 6 linkages are called **branch points**.
- **Glycogen** is like amylopectin except that the branch points occur more often. Glycogen is the type of starch that animals make and store; if you are in a good nutritional state, you have large amounts of it in your liver and muscles and also some in your brain. It is the source of most of the glucose you use in cellular metabolism.

It is worthwhile to note that to digest these three different macromolecules requires some different enzymes.

- **Cellulose** is a straight polymer of glucose but it differs in the exact structure of the glucose is slightly different than in the starches. The result is a molecule that most organisms are not able to digest. It is strong flexible and is certainly the most abundant life-derived compound on the planet.

Look at the figures in the text on pages 64-68 but do not learn any of the structures. Instead simply gain an appreciation for the ways that different starches can be made and for their sources and functions. It is not important that you learn actual structures.

When does something become big enough to be a starch or polysaccharide? (The same problem arises when discussing polypeptides and nucleic acid chains).

2. **Substituted carbohydrates:** chitin is a very important polysaccharide that differs from normal carbs. in that each sugar composing it contains N. There are a number of kinds of these "substituted" molecules.

Chitin, of course, is a principal constituent of the exoskeleton of arthropods and some other animals and of the cell walls of fungi.

**3. Nucleic Acids:** As you know, nucleic acids are the molecules of heredity. In most organisms, one type of nucleic acid, DNA, is used to store the hereditary information that codes for every protein in the body. This information is present in regions referred to as structural genes. Moreover, information that helps to determine when proteins will be synthesized is present in other areas (so called regulatory sequences and structural genes). More about this later. In some organisms (certain types of viruses, if indeed they are organisms) DNA is replaced by RNA. However, more commonly RNA functions in transferring information from DNA out of the nucleus to guide the formation of proteins, Moreover RNA also helps in more direct ways to synthesize proteins. More about this later when we cover anabolism.

Both types of nucleic acid are polymers of nucleotides. Nucleotides are therefore used to synthesize DNA and RNA. However, nucleotides also have other functions, primarily in energy transfer (ATP is a nucleotide) and as signal molecules (cyclic AMP is a nucleotide).

Let's therefore start with the structure of nucleotides. All varieties of nucleotide have a tripartite structure consisting of :

1. **a 5-carbon (pentose) sugar.** The sugar will either be ribose ( $C_5H_{10}O_5$ ) or deoxyribose ( $C_5H_{10}O_4$ )
2. **a nitrogenous base.** Nitrogenous bases are ring compounds formed principally of C and N. They either have one or two rings.
  - If a single ring, they are termed pyrimidines of which there are three types: **cytosine (C)**, **thymine (T)** or **uracil (U)**. Adenine and thymine will attach to deoxyribose and adenine and uracil will attach to ribose.
  - If a double ring structure, they are called **purines** and they go by the name of either **adenine (A)** or **guanine (G)**. Both purines will attach to either type of pentose sugar.
  - The nitrogenous bases differ from each other due to differences in the atoms attached to them.
3. also attached to the sugar are **one to three phosphate groups**. These are attached in series with each other -- one directly attaches to the pentose and the other one or two (when present) attach to the first phosphate group. When there are two or three phosphate groups present, **the outer two are very reactive** and are termed **high-energy phosphates**.

**About names<sup>2</sup>.** Nucleotides are named according to the sugar, nitrogenous base, and the number of phosphates present. Here are the conventions; be able to at least recognize what is being talked about:

<sup>2</sup> note that the names of the nucleotides have slightly different names than the nitrogenous bases -- this need not concern us the moment (wait till genetics). The different endings are used so that you can quickly tell you are talking about a nucleotide as compared to a nitrogenous base.

1. **Phosphates:**

- **one** phosphate = a **monophosphate** (abbreviated **MP**)
- **two** phosphates = a **diphosphate** (abbreviated **DP**)
- **three** phosphates = **triphosphate** (abbreviated **TP**)

2. **Sugars:**

- If the sugar is ribose, it is understood (no name is added)
- If deoxyribose, the nucleotide is prefixed as "deoxyribose blah-blah TP" or whatever. When abbreviated, a small d is used -- e.g., dCTP is deoxy cytosine triphosphate (*i.e.*, deoxyribose + three phosphates + cytosine as the nitrogenous base).

3. **Nitrogenous base** -- the main part of the name -- either the full name (e.g., uracil) or the abbreviation (U) plus a special ending that we will not learn (see note at the bottom of this box):

? OK, give the name of nucleotide that consists of:

- (a) ribose + three phosphates + adenine
- (b) ribose + two phosphates + thymine
- (c) deoxy ribose + guanine + one phosphate

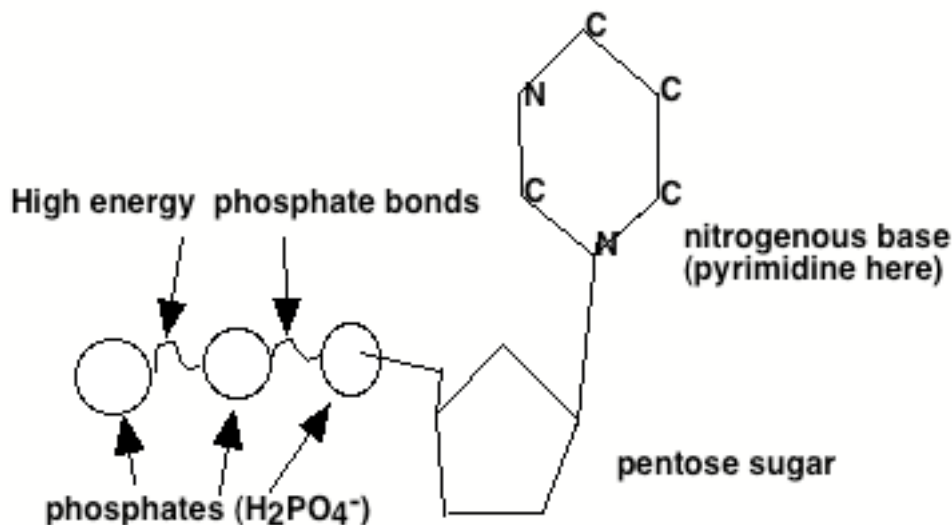
also give the number of high-energy phosphates

ANS: (a) ATP (adenosine tri phosphate) -- (two high energy phosphate)

(b) generally, no such compound is found -- why?

(c) dGMP (deoxyribose Guanine Monophosphate) (no high energy phosphates)

So, by way of review, here are a couple of nucleotides with the high-energy portions shown:



**Polymers of Nucleotides:** A nucleic acid polymer is formed when the sugar of one nucleotide binds with the phosphate group of its neighbor. The nitrogenous

bases constitute the part of part of the nucleic acid molecule that makes it unique. The situation is absolutely analogous to that in proteins; we will cover this in more detail in the next class. The **sugar phosphate backbone** holds the molecules together and therefore holds a given **sequence of nucleotides**. This sequence of nucleotides is what provides the molecule with its unique information and function.

As mentioned previously, there are two types of nucleic acids that are distinguished in name by the sugar in their nucleotides. These are deoxyribonucleic acid (DNA), the genetic material confined largely to the nucleus that contains the sugar deoxyribose, and ribonucleic acid (RNA), found both in the nucleus and the cytoplasm, which contains ribose. Other differences between these two macromolecules will be discussed below.

**DEOXYRIBONUCLEIC ACID -- DNA:** DNA is the substance of genes and along with protein forms the chromosomes located in the nucleus. It is a double-stranded nucleic acid. This means that it consists of two chains wrapped around each other to form a double helix. The two strands or chains are held together by chemical bonds between the nitrogenous bases in the respective strands.

DNA contains four different nitrogenous bases. These fall under two headings or classes:

- purines, which have a double ring structure, and
- pyrimidines, which consist of a single ring.

Recall that there are two purines (adenine and guanine) that comprise respectively the nucleotides adenosine and guanine, and two pyrimidines (thymine and cytosine) that comprise respectively the nucleotides thymidine and cytidine. The nucleotides that constitute a single chain are free to vary in sequence, but the base sequence between opposing strands is restricted by the phenomenon of complementary base pairing.

According to the rule of complementary base pairing, a purine on one strand can only bind with a pyrimidine on the other strand and vice versa. Furthermore, the rule is even more specific - adenine always binds with thymine and guanine always binds with cytosine. Thus, if the base sequence of one strand is known, the sequence of the opposing strand can be determined by using the rule of complementary base pairing. For example, if the base sequence in chain #1 is CCAGCTTTG, the base sequence in chain #2 must be GGTCGAAAC. The key to understanding the structure and function of DNA is the rule of complementary base pairing!

Diversity among genes is related not to between-strand bases, but to variation in the base sequence which makes up a single strand. Thus, molecules of DNA vary because the base sequences within their strands vary.

**RIBONUCLEIC ACID – RNA:** RNA exists in a number of different forms which we will discuss later and it differs from DNA in three ways:

- it is a single stranded molecule, and as mentioned previously:
- it contains the pentose sugar ribose instead of deoxyribose, and
- it lacks the base thymine but contains the base uracil.

More about RNA and DNA in a few weeks -- I'm sure you've had enough for now!

---

## **STUDY QUESTIONS FROM THIS HANDOUT AND FROM CAMPBELL, Ch. 5: MACROMOLECULES**

1. Carbohydrates: look at their structures and using what you have already learned (or reviewed) in this course, explain how they should interact with polar substances such as water. DO NOT MEMORIZE THE STRUCTURES OF CARBOHYDRATES. However, you should be able to:
  - (a) identify a compound as being a carbohydrate
  - (b) identify isomers of monosaccharides -- not by name, but simply tell that they are isomers of each other.
  - (c) know how di- and polysaccharides are formed and be able to list some of the factors that make them different from each other.
  - (d) know in general terms what glycogen, amylose, amylopectin and cellulose are (but you do not need to be able to draw them). What are their biological functions and where are they found in cells and what types of organisms produce them and are able to use them?
  - (e) give the key molecular difference between Chitin and other polysaccharides. What is chitin used for and what are the two large groups of organisms that use it?
2. What makes fats more fluid?
3. What are the names of purines and pyrimidines? Which are found in DNA and which are in RNA? What is the most obvious structural difference between these two types of nitrogenous bases? What pentose would be found in a compound consisting of the pyrimidine thymine and two phosphates? What would be the initials of this compound? How many high-energy bonds would it possess?
4. Why are lipids so hydrophobic? What does adding phosphate and other groups to fats do the properties of a molecule -- for instance, if phosphate-choline is substituted for a fatty acid in a triglyceride (see fig. 5.12) what happens to the properties and shape of the molecule?

5. Distinguish between each of the following:

sterol and fat  
saturated and unsaturated fatty acid  
hydrolysis and dehydration  
amylose vs. glycogen vs. cellulose  
monomer and polymer  
adenine and adenosine  
purine and pyrimidine  
uracil and thymine  
ribose and deoxyribose

4. Identify each of the following:

diester  
chitin  
hygroscopic  
nitrogenous base

cytosine  
guanine  
complementary base pairing  
transcription