
2.1 Small alkanes. Skeletal isomerism.

The simplest alkane, CH$_4$, is methane. Replace one of the H atoms of methane with a C atom, and attach enough H atoms to the C to satisfy its valence (three). We now have a two-carbon compound, ethane. Replace one of the H atoms of ethane (it doesn't matter which, they're all equivalent) with another CH$_3$ group. We now have a three-carbon compound, propane.

Propane has a property that is not present in ethane or methane. There are two different kinds of C atoms in propane, and there are two different kinds of H atoms. The terminal CH$_3$ groups in propane (which are indistinguishable, or equivalent) are called methyl groups, and the middle CH$_2$ group is called a methylene group. The H atoms on the CH$_3$ groups are called primary (1°), and the H atoms on the methylene group are called secondary (2°).
Suppose we want to add one more C atom to propane. We have two different kinds of H we can replace, a methylene H or a methyl H. If we replace the methyl H, we have *butane*. If we replace the methylene H, we have a *skeletal isomer* of butane. This isomer of butane has a common or trivial name, *isobutane*. More systematically, we can think of it as propane with a methyl group on the second carbon, or 2-methylpropane (all one word).

Skeletal isomers, like isobutane and butane, have the same *chemical formula*, but the *atom-to-atom connections* are different. Skeletal isomers have different properties — taste, smell, melting point, boiling point, everything. They are as different as if their elemental compositions were different. This concept is incredibly important. The different properties of skeletal isomers is one reason why there is such a diversity of organic compounds.

We will soon learn about other kinds of isomers. For now, note that “isomer” is a relationship word, like “cousin”. It makes no sense to say that a compound is an isomer. It does make sense to say that one compound is an isomer of another.

Butane has two different kinds of H atom, and so does isobutane. If we replace one H atom in butane with CH₃, we get two different kinds of pentane. If we do the same in isobutene, we also get two different kinds of pentane. So it seems there are four skeletal isomers with the formula C₅H₁₂. However, one of the forms of pentane that we generated from butane is the same as one of the forms we generated from isobutene, so in fact there are only three skeletal isomers: pentane, 2-methylbutane, and 2,2-dimethylpropane.

There are many different ways to draw a compound such as pentane. The mutual orientations of the different lines can change, and the compound is still pentane. The same is true of any other compound. *What matters is the connections between the atoms, not the relative positions of the atoms!*
One more point. Methane has the formula CH$_4$, ethane C$_2$H$_6$, propane C$_3$H$_8$, butane and isobutane C$_4$H$_{10}$, and the three pentanes C$_5$H$_{10}$. The formula for any acyclic alkane is C$_n$H$_{2n+2}$.

### 2.2 Higher alkanes.

We now arrive at one of my least favorite subjects: nomenclature. We have learned about methane, ethane, propane, butane, and pentane. Above pentane, the roots of the names derive from the Greek word for that number: hexane, heptane, octane, nonane, decane, undecane, dodecane, etc. The class of compounds is called alkanes. The root of the alkane's name tells you how many C atoms there are; the number of H atoms in an alkane that has no rings is given by the formula C$_n$H$_{2n+2}$.

A branched alkane is named as follows. Find the longest straight chain in the molecule. That provides the root of the name. Then describe the substituents on that long chain. If the substituent has one carbon, it is methyl; if two, ethyl, etc. The position of the substituent is described by a number; if there is a choice, the number of the first substituent should be as low as possible. If there is more than one substituent of the same kind, the prefix di-, tri-, etc. is used, and the position of every group must be listed. If there is more than one substituent of different kinds, they are listed alphabetically, and the first substituent in the name should have the lowest number. You can use i-Pr, s-Bu, i-Bu, and t-Bu as substituent names.

![Chemical structures](image)

2-methylhexane 3-ethylpentane 3,3-dimethylhexane

5-butyl-4-ethyl-2-methyldecane 2,3,3-trimethylpentane

I do not expect you to name compounds. I do expect you to be able to draw a compound that I name.
2.3 **Cycloalkanes.**

Let's take pentane and remove one H from each of the terminal C atoms. Then let's join the terminal two C atoms. The structure we have now has the formula C\(_5\)H\(_{10}\). We call it cyclopentane. In fact, we can make cyclic alkanes of any size. A three-membered ring is cyclopropane, a four-membered ring is cyclobutane, a six-membered ring is cyclohexane, etc. The number of H atoms in a cycloalkane is given by the formula C\(_n\)H\(_{2n}\).

When we have a cycloalkane with a single substituent, for example a fluorine atom on a six-membered ring, we call it fluorocyclohexane. No number is needed, because the ring has no beginning and no end. We can have alkyl substituents on an alkane: for example, isopropylcyclopropane. If we have more than one substituent, the alphabetically first substituent is at position 1, and the other substituent is numbered relative to the first one and as low as possible: for example, 1-fluoro-3-iodocyclooctane (not 1-fluoro-5-iodocyclooctane).

Many organic compounds have more than one ring. A compound can rings that are isolated, spiro, fused, or bridged, depending on whether they share 0, 1, 2, or >2 atoms.

2.4 **Alkenes and alkynes.**

So far, all the compounds that we’ve encountered have only C–C single bonds. However, C has the ability to make second and third bonds to its neighbors, to give double and triple bonds. Compounds that contain C=\(\equiv\)C double bonds are called alkenes, and compounds that contain C=\(\equiv\)C triple bonds are called alkynes. (Those of you whose native languages are not English need to be careful about pronouncing these words correctly: al-CAIN as in “Cain and Abel”, al-KEEN as in “keen eyesight”, and al-KYNE as in “you’re very kind”. In Spanish and other European languages, these words are pronounced al-CON as in “con artist”, al-KANE as in “Cain and Abel”, and al-KEEN as in “keen eyesight”.) Alkenes and alkynes are named like alkanes are named, except that the suffix is -ene or -yne instead of -ane. However, there is an additional caveat: in compounds with a double or triple bond in a chain of four or more C atoms, you must specify the position of the double or triple bond with a prefix. So there are two kinds of butene,
1-butene and 2-butene. (There is no 3-butene, because it is 1-butene.) The same is true of butyne.

\[
\begin{align*}
\text{H}_2\text{C}=\text{CH}_2 & \quad \text{CH}_2=\text{CHCH}_3 & \quad \text{CH}_2=\text{CHCH}_2\text{CH}_3 & \quad \text{CH}_3\text{CH}=\text{CHCH}_3 \\
ethene & \quad \text{propene} & \quad 1\text{-butene} & \quad 2\text{-butene} \\
(a\text{.k}\text{.a. ethylene}) & \quad (a\text{.k}\text{.a. propylene}) & & \\
\text{HC}=\text{CH} & \quad \text{HC}!\text{CCH}_3 & \quad \text{HC}!\text{CCH}_2\text{CH}_3 & \quad \text{CH}_3\text{C}!\text{CCH}_3 \\
ethyne & \quad \text{propyne} & \quad 1\text{-butyne} & \quad 2\text{-butyne} \\
(a\text{.k}\text{.a. acetylene}) & & & 
\end{align*}
\]

Compounds can have more than one double or triple bond. For example, \(\text{CH}_3\text{CH}_2\text{CH}=\text{CHCH}_2\text{CH}=\text{CHCH}_3\) has eight \(C\) atoms and \(\text{two}\) double bonds. As an eight-carbon compound, its name has the root \(\text{oct-}\). As a compound with two double bonds, its suffix is \(-\text{diene}\). Because \(\text{octadiene}\) is awkward, we call it \(\text{octadiene}\). But it’s not just any octadiene, it’s an octadiene with the double bonds in specific locations. We call it \(2,5\text{-octadiene}\) or \(\text{octa}-2,5\text{-diene}\). It’s not \(3,6\text{-octadiene}\) because we count from the end that gives the first double bond as low a number as possible. \(\text{CH}_3\text{CH}_2\text{CH}=\text{CHCH}_2\text{C}=\text{CCH}_3\), on the other hand, would be called \(3\text{-octen}-6\text{-yne}\), because for some reason we prefer to give the double bond a lower number than the triple bond. (I will \text{never} test you on naming a compound. There are free Web sites that will name any compound you draw.)

Cyclic compounds can also contain double and triple bonds. If the compound contains a single double or triple bond, you don’t need to give it a prefix indicating the position of the multiple bond. If the ring has a substituent, the multiple-bond \(C\) atoms are given the numbers 1 and 2 so that the substituent has the lowest number possible. If a compound contains more than one double or triple bond, then you do need to indicate the positions of the multiple bonds with prefixes. However, \(1,3,5\text{-cyclohexatriene}\) has its own name, benzene, both for historical reasons and because of its special chemical properties.
For reasons that we will discuss later, the second bond of a double bond, and the second and third bonds of a triple bond, are called $\pi$ bonds. (The first bond is a $\sigma$ bond.)

### 2.5 Degrees of unsaturation of hydrocarbons.

We’ve seen that an acyclic hydrocarbon with only single bonds has the formula $C_nH_{2n+2}$. If we remove two H atoms from adjacent C atoms and make a double bond between the C atoms, the resulting alkene will have the formula $C_nH_{2n}$. If we remove two H atoms from nonadjacent C atoms and make a new single bond between the C atoms, the resulting cycloalkane will have the formula $C_nH_{2n}$ as well. As a result, any compound with the formula $C_nH_{2n}$ must have one ring or one $\pi$ bond. As a shorthand, we call it one degree of unsaturation.

If we start with our alkene and convert the double bond to a triple bond, we remove two more H atoms, so we have the formula $C_nH_{2n-2}$. We get a compound with the same formula if we start with our alkene and make a ring, or if we put in a second double bond, or if we start with our cycloalkane and make another ring. In other words, compounds with two degrees of unsaturation (rings or $\pi$ bonds) have the formula $C_nH_{2n-2}$.

You can begin to see the pattern. A hydrocarbon with $n$ C atoms and $m$ degrees of unsaturation has the formula $C_nH_{2n+2−2m}$, or $C_nH_{2(n−m+1)}$. Furthermore, given the formula of a hydrocarbon, you can determine the number of degrees of unsaturation (rings and $\pi$ bonds) that the compound has as follows:

1. From the number of C atoms in the compound, compute the number of H atoms the compound would have if it had zero degrees of unsaturation ($2n + 2$).
2. Subtract the actual number of H atoms and divide by 2.

Once you know the number of degrees of unsaturation in a compound, it is easy to draw possible structures for that compound. You no longer need to worry about the number of H atoms in your structure, because it’s determined by the number of degrees of unsaturation. So if I ask you to
draw a compound with formula \( \text{C}_{12}\text{H}_{20} \), you would draw a compound with twelve C atoms and three degrees of unsaturation: three rings, a ring and a triple bond, two rings and a double bond, etc.

Note: A compound that contains at least six C atoms and at least four degrees of unsaturation often contains a benzene ring. For example, for \( \text{C}_7\text{H}_8 \), you could draw methylbenzene, a.k.a. toluene. \( \text{C}_9\text{H}_{10} \), with five degrees of unsaturation, can contain a benzene ring plus one more double bond or ring.

### 2.6 Common hydrocarbon substituents and abbreviations

The word "methyl" refers to a \( \text{CH}_3 \)– group that is attached to the rest of a molecule. In general, the root "meth-" means "a one-carbon group". The abbreviation Me is often used for \( \text{CH}_3 \). So MeI is the same as \( \text{CH}_3\text{I} \).

The word "ethyl" refers to a \( \text{CH}_3\text{CH}_2 \)– group that is attached to the rest of a molecule. In general, the root "eth-" means "a two-carbon group". The abbreviation Et is often used for \( \text{CH}_3\text{CH}_2 \). So EtBr is the same as \( \text{CH}_3\text{CH}_2\text{Br} \).

A three-carbon group attached to the rest of a molecule could be attached through its terminal C atom or its internal C atom. In the former case, the \( \text{CH}_3\text{CH}_2\text{CH}_2 \)– group is called “propyl” or “\( n \)-propyl” (the \( n \) stands for “normal”) and is often abbreviated as \( n \)-Pr. In the latter case, the \( (\text{CH}_3)_2\text{CH} \)– group is called “isopropyl” and is often abbreviated as \( i \)-Pr.

A four-carbon group attached to the rest of a molecule could be straight or branched, and either of these groups could be attached through its terminal C atom or its internal C atom. As a result, there are kinds of four-carbon groups, which are called butyl, isobutyl, \( sec \)- or \( s \)-butyl (\( sec \) stands for “secondary”), and \( tert \)- or \( t \)-butyl (\( tert \) stands for “tertiary”).
Another common substructure is the phenyl group, \( {\text{C}}_6{\text{H}}_5^- \), abbreviated as Ph. It consists of a benzene ring in which five of the six C atoms are attached to H, and the sixth C atom is attached to the rest of the compound. Another one is the acetyl group, \( \text{CH}_3\text{(=O)}^- \), abbreviated as Ac.

\[
\begin{array}{c}
\text{phenyl, Ph} \\
\text{acetyl, Ac}
\end{array}
\]

2.7 Functional groups.

Hydrocarbons that contain only single bonds are fairly unreactive except under drastic circumstances (like heating them very hot in the presence of \( \text{O}_2 \), as in your car’s engine). If a hydrocarbon contains \( \text{C}==\text{C} \) double or \( \text{C}≡\text{C} \) triple bonds, reactions tend to occur at or near those bonds. Reactions also tend to occur at or near heteroatoms (any atoms other than C or H) in organic compounds. The structures in organic compounds that contain heteroatoms or \( \pi \) bonds (or both) are called functional groups. The functional group approximation says that the chemistry of an organic compound can be understood by looking at the chemistry of its functional groups. Even the behavior of such a large compound as cholesterol can be understood as similar to the behavior of a small alcohol and a small alkene.

\[
\begin{array}{c}
\text{cholesterol}
\end{array}
\]

Important: Just because two compounds behave similarly doesn't mean that they behave the same. The more similar two compounds are, the more similarly they behave, but how similar two compounds must be for you to predict that they have similar behavior is a matter of judgment. For example, you might predict that cholesterol would have behavior similar to ethylene, \( \text{CH}_2=\text{CH}_2 \), but because the two alkene C atoms in cholesterol are attached to a total of three other C atoms, a better analogy would be \((\text{CH}_3)_2\text{C}=\text{CHCH}_3\), and an even better analogy would be 1-methylcyclohexene. Similarly, even though methanol, \( \text{CH}_3\text{OH} \), has an OH group like cholesterol, the OH-bearing C atom in cholesterol is attached to two other C atoms, so a better analog would be \((\text{CH}_3)_2\text{CHOH}\), or isopropanol.
As I survey the different functional groups, I don’t want you memorizing names of functional groups or worrying about whether a particular compound contains this or that functional group. The purpose of this section is to familiarize you with the kinds of structures that you will see as the semester proceeds.

### 2.8 Functional groups containing C–X single bonds.

The heteroatoms that are found most often in organic compounds are N, O, the halogens, and, to a lesser extent, P and S.

#### 2.8.1 Haloalkanes.

Organic compounds that contain halogen are called haloalkanes or alkyl halides. Halogen-containing compounds are not frequently found naturally in terrestrial organisms, but they are often found in artificial organic compounds, and they play an important role in medicinal chemistry. They are also widely used in organic synthesis.

![isoflurane (anesthetic)](image1)
![thyroxine (hormone)](image2)

![fluoxetine (antidepressant)](image3)
![DDT (insecticide)](image4)

![Teflon (material)](image5)

Start with propane, and replace an H atom with a Br atom. The Br atom could end up on a terminal C atom, or it could end up on the internal C atom. These two compounds, called 1-bromopropane and 2-bromopropane, have the same formula but different atom-to-atom connections, so they are skeletal isomers. Their physical, chemical, and biological properties are different. They are different compounds.
Now start with methylcyclohexane. How many different skeletal isomers can you draw by replacing a single H atom with a Cl atom?

2.8.2 Alcohols and thiols.
Organic compounds that contain a C–OH group are called alcohols. (Exception: C(=O)–OH groups are not called alcohols, but are instead called carboxylic acids.) Alcohols are one of the most widely found functional groups in nature. We name alcohols by adding the suffix -ol to the end of the name of the parent alkane, using a numerical prefix if the position of the OH group on the chain is ambiguous. So 1-pentanol, 2-pentanol, and 3-pentanol are all structural isomers, and CH$_2$=CHCH$_2$CH$_2$OH is 3-buten-1-ol. (Note that the numbering gives the OH group priority over the alkene group.)

![Chemical structures of alcohols](image)

Many compounds in our bodies contain the alcohol functional group. Sugars, also called carbohydrates, are polyalcohols. Ribonucleic acids (RNA) and deoxyribonucleic acids (DNA) contain alcohol groups that ultimately derive from sugars (ribose or deoxyribose) as well. The amino acids serine, threonine, and tyrosine contain alcohol groups as well.
Alcohols can be classified as primary, secondary, or tertiary ($1^\circ$, $2^\circ$, or $3^\circ$), depending on whether the C atom bearing the OH group is bound to one, two, or three nonhydrogen atoms. Menthol and isopropanol are $2^\circ$ alcohols, whereas ethanol and 2-ethyl-1-hexanol are $1^\circ$ alcohols.

Organic compounds that contain a C–SH group are called thioles or, more archaically and poetically, mercaptans. The most noticeable property of thioles is their stench. Methanethiol, CH$_3$SH, is one of the compounds that gives flatulence its delightful odor. The odiferous compounds in skunk spray are primarily thioles as well. The amino acid cysteine, the common biological laboratory reagent dithithreitol (DTT, Cleland’s reagent), and the important enzyme cofactor coenzyme A contain thiol groups as well.
2.8.3 Ethers and sulfides.

Organic compounds that contain an O atom connected to two C atoms, neither of which is doubtfully bound to another O, are called *ethers*. The very first ether to be prepared was diethyl ether, EtOEt, which at the time (1600s) was called simply ether. (It was called ether because it was so prone to evaporate that the chemists thought that they may have isolated Plato’s fifth element.) Other common ethers include *t*-butyl methyl ether, which was used in gasoline to reduce smog formation until it started contaminating groundwater, and tetrahydrofuran, a common solvent. Several of the haloalkanes that I drew previously also contain ether groups in them.

We often name ethers as if they were alkanes modified by addition of an O atom bound to another alkyl group. An alkyl group plus an O atom is called an *alkoxy* group. If the alkyl group contains one C atom, the *methoxy*; if it contains two, *ethoxy*; and so on.

\[ \text{diethyl ether (solvent)} \]
\[ \text{*t*-butyl methyl ether (antismog additive)} \]
\[ \text{tetrahydrofuran (solvent)} \]

2-(3-methylbutoxy)-4-methoxyoctane
Although they are not as common as alcohols among naturally occurring compounds, ethers are still quite common. Both THC (the psychoactive ingredient of marijuana) and codeine contain both ether and alcohol groups.

The sulfur analogs of ethers are called sulfides or thioethers. The most important sulfide in living organisms is methionine.

2.8.4 Amines.

Organic compounds that contain an N atom with only single bonds to C atoms, none of which is doubly bound to an O, are called amines. Amines are further classified as primary, secondary, or tertiary (1°, 2°, or 3°), depending on whether the N atom is bound to one, two, or three nonhydrogen atoms.

Amines are among the most common functional groups in living organisms. All of the amino acids, the building blocks of proteins, contain amines. (That’s why they’re called amino acids.) All of the amino acids that are coded by DNA are primary amines, with the exception of proline, which is a secondary amine. The amino acid lysine contains two amine groups. Other amines
commonly found in living (and decomposing) organisms are putrescine, cadaverine, and spermidine. (Amines, which were first isolated from rotting fish, are known for their unpleasant odors, as you can tell from the names of some of these compounds.)

\[
\begin{align*}
\text{glycine} & : \quad \text{H}_2\text{N}\text{–COOH} \\
\text{lysine} & : \quad \text{H}_2\text{N}\text{–C}_9\text{H}_{15}\text{–COOH} \\
\text{putrescine} & : \quad \text{H}_2\text{N}\text{–C}_9\text{H}_{17}\text{–NH}_2
\end{align*}
\]

\[
\begin{align*}
\text{cadaverine} & : \quad \text{H}_2\text{N}\text{–C}_9\text{H}_{15}\text{–NH}_2 \\
\text{spermidine} & : \quad \text{H}_2\text{N}\text{–C}_9\text{H}_{17}\text{–NH}\text{–C}_9\text{H}_{17}\text{–NH}_2
\end{align*}
\]

Three of the four DNA bases also contain the amine functional group.

\[
\begin{align*}
\text{adenine} & : \quad \text{N}_4\text{H}_2\text{–NH}_2 \\
\text{guanine} & : \quad \text{N}_4\text{H}_2\text{–NH}\text{–NH}_2 \\
\text{cytosine} & : \quad \text{N}_4\text{H}_2\text{–NH}_2 \\
\text{thymine} & : \quad \text{N}_4\text{H}_2\text{–NH}\text{–C}_3\text{H}_2\text{–NH}_2
\end{align*}
\]

Many amines cross the blood–brain barrier and have central nervous system (CNS) activity. Some of these amines are natural neurotransmitters, some are used for legitimate medical purposes, and some have no medicinal applications. Some are primary amines, some are secondary, and some are tertiary.
Because many amines were isolated before their structures could be determined, many of them have common names that do not reflect their structures.

### 2.9 Functional groups containing C=X and C≡X bonds.

#### 2.9.1 Carbonyl compounds.

The central functional group in organic chemistry is the carbonyl group, in which a C atom bears a double bond to O. Carbonyl compounds can be divided into those whose carbonyl C atoms are
singly bound to zero, one, or two heteroatoms. The first group includes the aldehydes and the ketones. The second group includes carboxylic acids, esters, carboxamides (often called amides for short), thioesters, acyl anhydrides, and acyl chlorides. The third group includes carbonate, carbamate, and ureas. Almost all of these different functional groups are widely distributed in nature (exceptions are in red).

The carbonyl groups all share the C=O group. We have already seen that one can draw the C=O group in the resonance form, in which the C atom is electron-deficient. We have also seen that the more good resonance structures one can draw, the lower in energy the compound. Therefore, the energy of different kinds of carbonyl compounds is determined by the ability of the groups attached to the carbonyl C to donate electrons to the partially electron-deficient carbonyl C atom and stabilize it, or, conversely, by their tendency to pull electrons away from the electron-deficient carbonyl C atom, thereby destabilizing it. The order of carbonyl compound energy is

\[
\text{ureas} < \text{carbamates} < \text{amides} < \text{esters} = \text{carboxylic acids} < \text{ketones} < \text{aldehydes} \sim \text{thioesters} < \text{acyl anhydrides} < \text{acyl chlorides}
\]
Carbonyl compounds that bear N and O are lower in energy than ketones and aldehydes because both N and O have lone pairs that they can share well with the neighboring C atom, but N is even more stabilizing than O because it is less electronegative. On the other hand, because the S–C and Cl–C bonds are long, and because the S and Cl atoms are considerably larger than the C atom, neither is a good resonance donor, and electronegativity takes over, making thioesters and acyl chlorides higher in energy than ketones and aldehydes. Although acyl anhydrides have an O atom attached to the carbonyl group, that O atom can’t share its lone pairs very well because there’s a carbonyl group on the other side of the O atom, pulling those electrons away, so the electronegativity of O begins to dominate, making these compounds higher in energy than ketones and aldehydes. Finally, ketones are lower in energy than aldehydes because the electron-deficient carbonyl C atom can draw some electron density from the α-C–H bonds in ketones and aldehydes, and there are more α-C–H bonds in ketones than there are in aldehydes. (We’ll revisit this subject soon.)

2.9.1.1 Aldehydes and ketones.

Aldehydes and ketones often have pleasant odors. The simplest ketone, acetone (more systematically known as 2-propanone), is used in fingernail polish remover and as a solvent, and it is a natural product of metabolism of fat for energy in the body (leading to “fruity breath”). The simplest aldehyde, formaldehyde (methanal), is widely used as a preservative, although it is also extremely toxic. When you consume alcoholic beverages, your body converts the ethanol into the two-carbon aldehyde acetaldehyde (ethanal) and thence to the carboxylic acid acetic acid, which can be excreted in the urine. If you consume so much alcohol that your body can no longer convert the acetaldehyde to acetic acid, the acetaldehyde accumulates in your body, causing some of the symptoms of hangovers. Many people of East Asian descent have a mutation in the enzyme that converts acetaldehyde to acetic acid, which makes them more prone to hangovers and less prone to enjoy consuming alcoholic beverages in the first place.

\[
\begin{align*}
\text{acetone (2-propanone)} & : \quad \text{H}_3\text{C} & \quad \text{O} & \quad \text{CH}_3 \\
\text{acetaldehyde (ethanal)} & : \quad \text{H}_3\text{C} & \quad \text{O} & \quad \text{H} \\
\text{formaldehyde (methanal)} & : \quad \text{H} & \quad \text{O} & \quad \text{H}
\end{align*}
\]

Most sugars contain an aldehyde or ketone group. Glucose, for example, contains an aldehyde group, whereas its skeletal isomer fructose contains a ketone group. The vitamin pyridoxal also
contains an aldehyde. Ketones are also found in certain hormones, such as progesterone and testosterone. On the other hand, none of the amino acids coded by DNA contain ketones or aldehydes, and the presence of such groups in proteins or fatty acids is often a sign that a cell has been subjected to oxidative damage.

(By the way, note how similar progesterone and testosterone are. Most of the compounds are identical. They differ only in the functional group and number of C atoms on the upper right. This small difference leads to big developmental differences. Vive la différence!)

In condensed structures, the aldehyde group is often written as –CHO, e.g., \( \text{CH}_3\text{CH}_2\text{CH(CH}_3\text{)}\text{CH}_2\text{CHO} \). If you see CHO at the end of a condensed structure, think aldehyde.

In condensed structures, the ketone group is often written as –CO–, e.g., \( \text{CH}_3\text{CH}_2\text{CH(CH}_3\text{)}\text{COCH}_3 \). However, –CO– can occasionally represent a tertiary alcohol, as in \( (\text{CH}_3)_3\text{COH} \), or a tert-alkyl ether, \( (\text{CH}_3)_3\text{COCH}_3 \), so be careful.

2.9.1.2 Carboxylic acids and derivatives.

Carboxylic acids are one of the most common functional groups. They were among the first organic compounds to be isolated from organic matter and the first to have their formulas determined. The simplest carboxylic acid, formic acid, was first isolated from ants, which use it as a chemical weapon. Acetic acid has been know since ancient times in its dilution with water,
which we call vinegar. Wine that spoils turns to vinegar because the ethanol CH$_3$CH$_2$OH in the wine is oxidized to acetic acid CH$_3$CO$_2$H. Fatty acids such as oleic acid contain carboxylic acids as well. Fatty acids can be saturated (containing only the carboxylic acid functional group) or unsaturated (containing one or more alkene functional groups as well). Benzoic acid is an important food preservative.

![Chemical structures]

Of course, all amino acids, by definition, contain the carboxylic acid group. Two of the amino acids coded by DNA, aspartic and glutamic acids, are actually dicarboxylic acids. When these amino acids are incorporated into proteins, the CO$_2$H group further from the NH$_2$ group remains unaltered in the protein.

![Amino acid structures]

Carboxylic acids also play important roles in metabolism. For example, malonic acid is the building block for the synthesis of fatty acids, and all of the compounds in the Krebs cycle (or citric acid cycle), including pyruvic acid and malic acid, are carboxylic acids.
In condensed structures, the carboxylic acid group is often written as \( -\text{CO}_2\text{H} \) or \( -\text{COOH} \), e.g., \( \text{CH}_3\text{CH}_2\text{CH}(\text{CH}_3)\text{CH}_2\text{CO}_2\text{H} \) or \( \text{H}_2\text{NCH}_2\text{COOH} \).

The other functional groups in the group of carbonyl compounds in which the carbonyl C atom bears one single bond to a heteroatom are often called carboxylic acid derivatives because they can all be easily prepared from or converted into carboxylic acids by replacing the OH group of the carboxylic acid with another heteroatom-bearing group, or vice versa.

Carboxylic esters, or esters for short, are derived from carboxylic acids by replacing the OH group with an alkoxy group. Ethyl and methyl esters are the most common esters, but others are common as well. Aspartame and cocaine both contain methyl ester groups; cocaine also contains a benzoate ester. Heroin contains two acetate esters, and polyester describes a group of polymers with ester functional groups in the polymer chains.
Carboxamides, or amides for short, are derived from carboxylic acids by replacing the OH group with an amino group. Like amines, amides can be 1°, 2°, or 3°, depending on the number of C atoms to which the N is bound. Nicotinamide, acetaminophen, and DEET are amides that are often encountered in everyday life. N-Acetylglucosamine is a component of chitin, the substance that arthropods use to build their shells. The diethylamide of lysergic acid is the only derivative of lysergic acid (the carboxylic acid) that has psychotropic properties.

When the amino group of the methyl ester of phenylalanine replaces the OH group of one of the carboxylic acid groups of aspartic acid, the resulting amide is Aspartame, an artificial sweetener. (Good luck trying to discern any resemblance between Aspartame and glucose!)
Aspartame is an example of a dipeptide. One can also make tripeptides, tetrapeptides, etc., by repeated application of the same process that made the dipeptide. When a polypeptide is long enough, we call it a protein. Biologists often call the N–C=O bond in peptides a “peptide bond,” but now you know what it really is: an amide bond.

Nylon 66, the first artificial fabric, is another polymer joined together by amide bonds. Each N atom of 1,6-hexanedianeine, H₂N(CH₂)₆NH₂, replaces the HO groups of a molecule of adipic acid, HO₂C(CH₂)₄CO₂H. Repeat this process indefinitely, and you have the nylon 66 polymer.
We have seen that amides are among the lowest-energy of the carbonyl groups, and it is for this reason that amides are so useful for compounds like clothing and proteins.

Acyl chlorides, acyl anhydrides, and thioesters all share the properties of being high in energy and having an electronegative atom attached to the carbonyl C atom. These two features make these compounds particularly useful in substitution reactions, in which another atom (such as C, O, or N) replaces the electronegative atom. In the laboratory, we use acyl chlorides and acyl anhydrides most often for this purpose, but organisms use acyl anhydrides and thioesters for this purpose. The most prominent example of a thioester found in organisms is acetyl-CoA.

2.9.2 Imino compounds and nitriles.

The analogs of carbonyl compounds in which the C=O bond is replaced with a C=N bond are called imino compounds. There are only two kinds of imino compounds that I want to show to you. Imines (sometimes called Schiff bases by biologists) are the N analogs of ketones and aldehydes, and guanidines are the N analogs of ureas. Arginine, one of the amino acids coded by DNA, contains a guanidine group.
Because N can make three bonds, there is a N-containing functional group that has no direct O analog. It is the carbonitrile or nitrile functional group, C≡N. A well-known nitrile is acetonitrile, CH₃CN, a common solvent. Acrylonitrile, CH₂=CHCN, can be polymerized to make polyacrylamide gels or nitrile rubber. α-Hydroxynitriles, which contain an alcohol group on the same C atom that attached to the CN group, are found in many plants and other organisms and are used a source of cyanide ion, −CN, which can be used for defense purposes.

### 2.10 Aromatic compounds.

Aromatic compounds are ring compounds in which every atom in the ring is either participating in a double bond or bears a lone pair. We will see soon that aromatic compounds are particularly common because the double bonds and lone pairs in the ring confer these compounds with extraordinarily low energy.

The most common kinds of aromatic compounds contain five-membered rings or six-membered rings. Six-membered aromatic compounds have three alternating double bonds. Benzene is the archetypical example of this kind of aromatic compound. However, one can replace one or more of the C atoms of benzene (and the associated H atoms) with N atoms and retain the three alternating double bonds. Replacing one ring C atom with N gives a pyridine, and replacing two N atoms gives either a pyridazine, pyrimidine, or pyrazine, depending on whether the N atoms are adjacent, askew, or opposite one another. Of the three C₄H₄N₂ isomers, pyrimidines are by far the most important.

- benzene
- pyridine
- pyridazine
- pyrimidine
- pyrazine
N-containing aromatic compounds often bear carbonyl groups on C atoms adjacent to one or more ring N atoms. These compounds appear at first sight not to have three alternating double bonds. However, if you draw the $\tilde{\mathbf{N}}=\tilde{\mathbf{C}}$–$\tilde{\mathbf{O}}$ resonance form, you can see that they do.

Five-membered aromatic compounds have four atoms that participate in double bonds and a fifth atom that has a lone pair of electrons. The fifth atom is most often N, as in pyrrole, but it can also be O or S, as in furan and thiophene. As before, one can replace one or more of the double-bonded C atoms with N atoms to get new five-membered aromatic compounds such as imidazole and thiazole.

The double bonds of aromatic compounds can be shared between adjacent aromatic rings. For example, two benzene rings can share a double bond to give naphthalene, a major component of mothballs. A benzene can share either the 2,3- or 3,4-double bond of pyridine to give quinolone or isoquinolone. If a benzene ring shares a double bond with pyrrole, the result is an indole, and if a pyrimidine shares a double bond with imidazole, the result is a purine. There are a very large number of possible combinations just of two rings, and the number just keeps skyrocketing if one adds a third ring.
Because aromatic compounds are particularly low in energy, they are also particularly common. Four of the amino acids coded by DNA — phenylalanine, tyrosine, histidine, and tryptophan — contain aromatic rings.

\[
\begin{align*}
\text{phenylalanine} & \quad \text{tyrosine} \\
\text{R} = \text{H} & \quad \text{R} = \text{OH}
\end{align*}
\]

The five DNA and RNA bases are all aromatic compounds as well. You might expect organisms to have evolved to store information as important as their genes in compounds that are particularly low in energy!

Many enzyme cofactors also contain aromatic groups. Pyridoxal and nicotinamide both contain pyridine rings, and thiamine contains both a pyrimidine and a thiazole.
2.11 Degrees of unsaturation of heteroatom-containing compounds.

We’ve seen that an acyclic hydrocarbon with only single bonds has the formula $C_nH_{2n+2}$. We can replace one or more H atoms with halogen atoms without changing the number of degrees of unsaturation. Conversely, a compound with halogen atoms has the same number of degrees of unsaturation as the compound we would obtain if we replaced all the halogen atoms with H atoms. As a result, a compound with the formula $C_mH_nX_p$ (where X is a halogen atom) has the same number of degrees of unsaturation as the corresponding hydrocarbon with the formula $C_mH_{n+p}$. For example, a compound with the formula $C_6H_{10}BrCl$ would have the same number of degrees of unsaturation as a compound with the formula $C_6H_{12}$ (one degree).

How do we deal with compounds that contain O atoms? Any O atom in a compound can be replaced with a CH$_2$ group without changing the number of degrees of unsaturation of that compound. So one approach is to say that any compound with the formula $C_mH_nO_p$ has the same number of degrees of unsaturation as the corresponding hydrocarbon with the formula $C_{m+p}H_{n+2p}$. However, it is also true that adding or removing CH$_2$ groups from a structure does not change the number of degrees of unsaturation of the compound. (Compare CH$_3$CH$_2$CH$_2$CH=CH$_2$ to CH$_3$CH$_2$CH=CH$_2$ and CH$_3$CH$_2$CH$_2$CH$_2$CH=CH$_2$.) So any compound with the formula $C_mH_nO_p$ has the same number of degrees of unsaturation as the corresponding hydrocarbon with the formula $C_mH_n$. In other words, when computing degrees of unsaturation, you can ignore any O atoms (or S atoms, for that matter) in the formula.

How do we deal with compounds that contain N atoms? Any N atom in a compound can be replaced with a CH group without changing the number of degrees of unsaturation of that compound. So one approach is to say that any compound with the formula $C_mH_nN_p$ has the same number of degrees of unsaturation as the corresponding hydrocarbon with the formula $C_{m+p}H_{n+p}$. Another approach is to say that replacing an N atom in a compound with a CH group is the same as adding a CH$_2$ group and then removing an H atom. We know that adding a CH$_2$ group has no effect on the number of degrees of unsaturation, so it follows that any compound with the formula $C_mH_nN_p$ has the same number of degrees of unsaturation as the corresponding hydrocarbon with the formula $C_{m+n}H_{n+p}$. For example, a compound with the formula $C_6H_{11}N$ would have the same number of degrees of unsaturation as a compound with the formula $C_6H_{10}$ (two degrees).

If a compound contains several different kinds of heteroatoms, you combine all these steps. Add an H atom for each halogen, subtract an H atom for each N or P atom, and ignore the O and S
atoms. For example, consider C₇H₁₀ClN₃O. It has the same number of degrees of unsaturation as C₇H₁₀⁺₁−₃ = C₇H₈, or four degrees of unsaturation.

Note: A five-membered aromatic compound has three degrees of unsaturation, and a six-membered aromatic compound, like benzene, has four degrees of unsaturation. If you have to draw a structure for a compound that has at least three degrees of unsaturation, consider drawing a five- or six-membered aromatic compound.