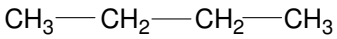
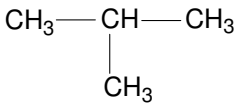
Organic Nomenclature – The Basics

There are two compounds having the molecular formula C_4H_{10}, which differ only in the way the atoms are connected to one another. Molecules which have the same molecular formula, but different connectivities among the constituent atoms, are called **structural isomers**. The structural isomers having the formula C_4H_{10} are shown below. Note that the prefix "iso" means "same", so isobutane received its name because it has the same formula as butane.

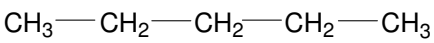


Butane

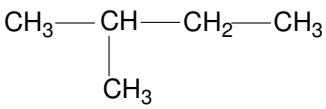


Isobutane

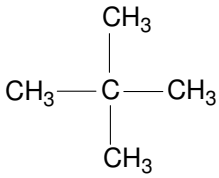
The situation become slightly more complicated when the number of carbon atoms is increased to 5. There are 3 different structural isomers having the molecular formula C_5H_{12}: pentane, isopentane, and neopentane ("neo" means "new").



Pentane

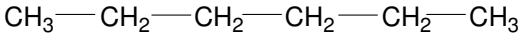


Isopentane

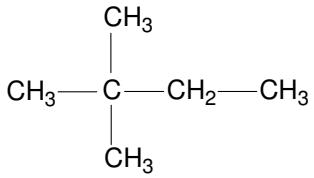


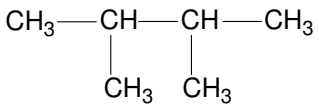
Neopentane

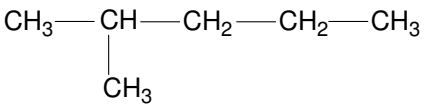
As the number of carbon atoms increases, the number of possible structural isomers increases geometrically. For example, 5 different compounds have the molecular formula C_6H_{14}. These are different compounds, and thus require different names. What prefixes should be used to distinguish one "hexane" from another?

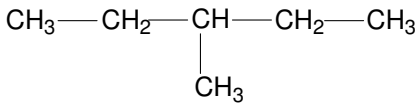


Hexane









It only gets worse. There are 75 possible structural isomers having the formula $C_{10}H_{22}$, 366,319 possible structural isomers having the formula $C_{20}H_{42}$, and 4,111,846,763 possible structural isomers having the formula $C_{30}H_{62}$! Since isomers are different compounds, they need to have different names. Furthermore, the name assigned to a given compound must be unambiguous, so that chemists all over the world draw the same structure when they see the name. As you might expect, a series of rules have been devised to make this entire process highly systematic. The rules for naming organic compounds were developed by the International Union of Pure and Applied Chemistry (IUPAC). The basic IUPAC rules for naming simple branched alkanes are outlined below. These rules form the basis for the nomenclature of all organic molecules, so it is important to become familiar with them! As we progress through the course, additions and modifications to the IUPAC rules will be introduced as necessary to accommodate the nomenclature of the various functional groups we will encounter.

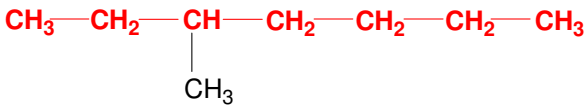
The rules are listed below in order of decreasing priority. When naming a compound, start with Rule 1, and apply each in succession until you can assign an unambiguous name to the compound. You may find that it is not necessary to use all of the rules.

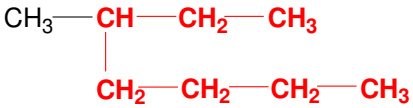
1) **Unbranched alkanes** (all carbon atoms lie in a continuous chain) are named according to the number of carbons as shown below.

Number of C Atoms	Name	Number of C Atoms	Name
1	Methane	11	Undecane
2	Ethane	12	Dodecane
3	Propane	13	Tridecane
4	Butane	14	Tetradecane
5	Pentane	15	Pentadecane
6	Hexane	16	Hexadecane
7	Heptane	17	Heptadecane
8	Octane	18	Octadecane
9	Nonane	19	Nonadecane
10	Decane	20	Icosane

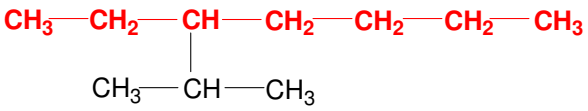
It is important to know these names (for C1 through C12), as they form the basis for naming more complicated molecules!

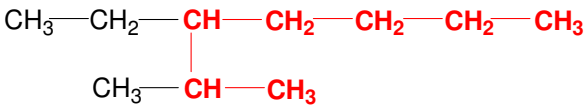
2) For branched alkanes, determine the **principal chain**. In a **branched alkane**, all of the carbon atoms do not lie in a continuous chain. A useful way to identify the principal chain is to start at one end of the molecule, and trace along with your finger. See how many carbon atoms you can incorporate into a continous chain without lifting your finger from the page. The principal chain will be the one that incorporates as many of the carbon atoms as possible. Be careful! The principal chain will not necessarily be obvious at first glance. For example, be sure that you understand why the two molecules shown below are actually the same molecule, just drawn differently! The principal chain (7 C atoms) is highlighted in each structure.



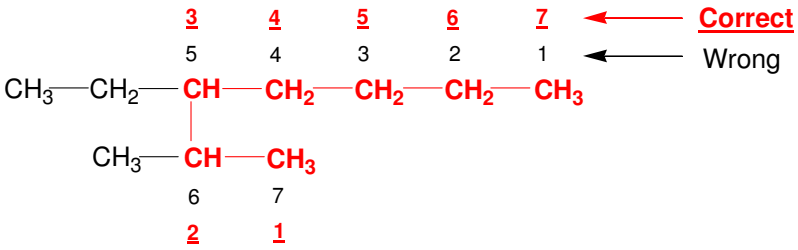


3) If two or more chains within a structure have the same length, choose as the principal chain the one with the greater number of branches. In the molecule shown below, there are two chains containing 7 C atoms. The structure on the left has 1 branch, while the structure on the right has 2 branches. The principal chain is thus identified correctly in the structure on the right.

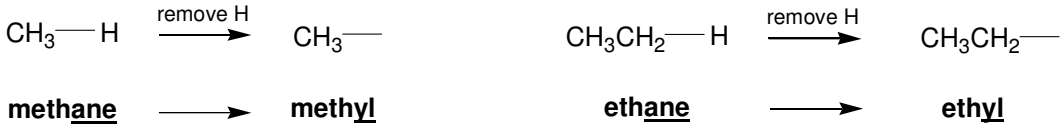




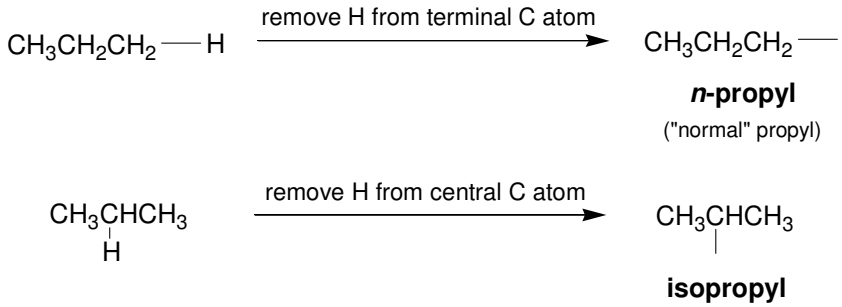
4) Number the carbon atoms of the principal chain consecutively from one end to the other in the direction that gives the lower number to the first branch. When there are substituent groups at more than one carbon of the principal chain, alternative numbering schemes are compared number by number, and the one is chosen that gives the lower number at the first point of difference. In the example below, the two possible numbering schemes are indicated. The scheme with the underlined numbers is correct, because the first branch occurs at C2 rather than C5.



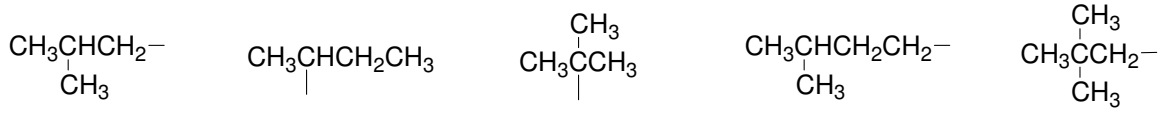
5) Name each branch and identify the carbon number of the principal chain at which it occurs. Note: Groups which are not part of the principal chain are called **substituent groups**. To name the substituent groups, you need to be able to name alkyl groups. An **alkyl group** is derived from an alkane by removing a hydrogen atom, and using the "hanging bond" left behind as the point of attachment to a larger molecule. The names of unbranched alkyl groups are derived by replacing the "-ane" ending of the alkane name with "-yl", as demonstrated below for "methyl" and "ethyl."



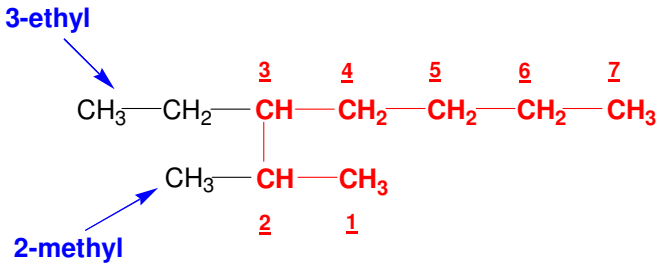
If the point of attachment is not on the terminal carbon atom, the alkyl group is a branched alkyl group. This situation can arise when an alkyl group is derived from a parent alkane containing 3 or more carbon atoms. For instance, there are two alkyl groups that can be derived from propane, as shown below.



In addition to "isopropyl", there are several common branched alkyl groups that you need to know.

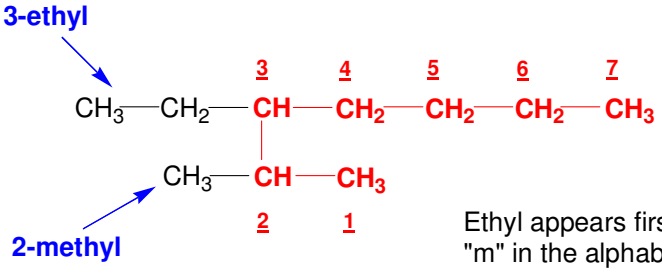


Continuing with our example from Rule 4, there is a methyl group substituent at C2 and an ethyl group substituent at C3.



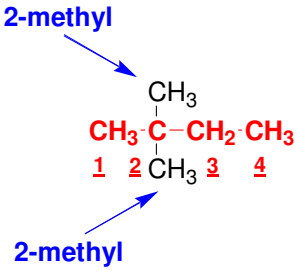
6) In writing the name, the substitutent groups are listed first. The name of each substituent is preceded by its position along the principal chain. Numbers are separated from letters by hyphens, and from other numbers by commas. The substituent groups are cited in alphabetical order regardless of their location in the principal chain. The name of the alkane corresponding to the principal chain is given after all of the substituent groups are listed. Keep in mind the following conventions. Several examples are provided below to illustrate some of the subtleties.
(a) If there are identical substitutents, the prefixes di-, tri-, tetra-, etc., are used to indicate the number.
(b) When there are multiple substituent groups on the principal chain, each substituent receives its own number, even if there are substituents located on the same carbon.

(c) The numerical prefixes di-, tri-, tetra-, etc., as well as the prefixes *sec*- and *tert*-, are ignored in alphabetizing. The prefixes iso-, neo-, and cyclo- are considered when alphabetizing substituent groups, however.



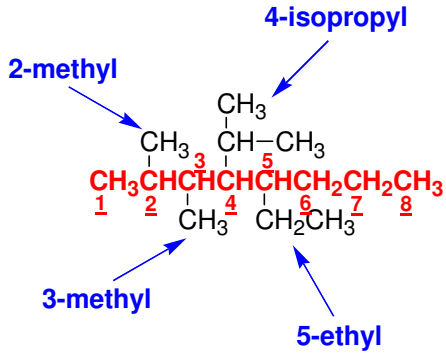
3-Ethyl-2-methylheptane

Ethyl appears first in the name because "e" precedes "m" in the alphabet, even though the methyl group appears before the ethyl group along the principal chain.



2,2-Dimethylbutane

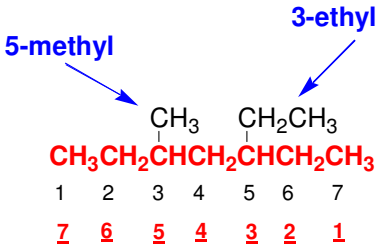
The number 2 is listed twice, because both methyl groups are on C2, and each needs to be identified by a number. Use of the modifier "di-" indicates that there are 2 methyl groups in the molecule



5-Ethyl-4-isopropyl-2,3-dimethyloctane

Even though "d" appears before "e" in the alphabet, ethyl appears first in the name because the modifier "di-" is ignored when alphabetizing. The "m" of "methyl" is used to locate the citation of the methyl groups in the name. Also note that the "i" of isopropyl is considered when alphabetizing.

7) When the numbering of different groups is not resolved by any of the other rules, the first-cited group receives the lower number.



3-Ethyl-5-methylheptane

In both possible numbering schemes, there is are substituents at C3 and C5. Since ethyl is cited before methyl alphabetically, the principal chain is numbered to give the ethyl group the lower number.

If you apply the IUPAC nomenclature rules to the five isomers having chemical formula C_6H_{14} (see above), you should come up with the following names: hexane, 2-methylpentane, 3-methylpentane, 2,2-dimethylbutane, and 2,3-dimethylbutane.

Nomenclature can be tricky when you are first learning it, but it becomes easy with practice. The only way you will learn how to apply these rules correctly is with extensive practice. As you begin to work through the nomenclature problems in Chapter 2, you may find that you need to refer to these rules frequently, and that's OK. As you practice and become familiar with the procedure, however, you'll soon find that you can name alkanes without looking at the rules. Work enough problems to reach this stage, and your understanding of nomenclature will be at the level that is expected in this course!