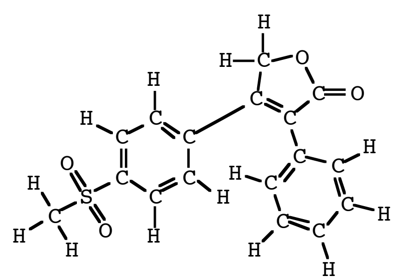
Ways of representing organic molecules

The most important, interesting and useful organic molecules tend to have lots of atoms and lots of bonds. Consider the structure of the controversial drug Vioxx:



Vioxx is a COX-2 inhibitor, that is, it reduces inflammation and relieves pain by inhibiting the action of the inflammation- and pain-producing enzyme cyclooxygenase-2 (COX-2). When it was revealed that long-term use of Vioxx increases the risk of heart attack, Merck, the manufacturer of Vioxx, voluntarily withdrew Vioxx from the market despite the drug's potential of earning $2.5 billion per year in sales. Controversy centers around whether Merck knew about the adverse side effects of Vioxx before releasing the drug. Decisions by appellate courts have tended to reduce Merck's liability by overturning monetary damage awards to patients and by wholesale dismissals of class-action lawsuits leveled against the company.

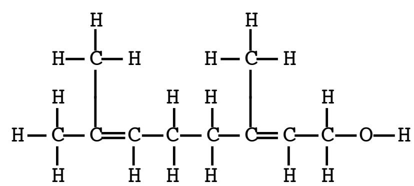
As big and complex as Vioxx looks, it is dwarfed by many organic molecules, some of which have tens of thousands of atoms (Vioxx has only 36). We need quick and compact ways of writing organic structures. We will now discuss the various notations used to represent organic molecules.

Full structure notation

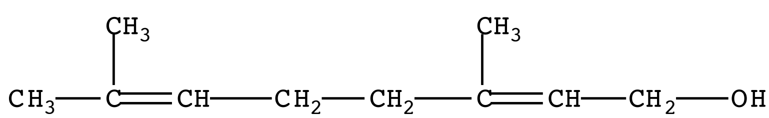
In full structure notation, every atom and every bond is explicitly shown. The representation of Vioxx given above is in full structure notation. Full structure notation is the most time-consuming, most cumbersome and least de-sirable way of representing organic molecules. You are strongly urged to wean yourself off full structure notation as quickly as possible and instead concentrate on mastering the techniques of representation discussed below, especially the bond–line notation.

Condensed structure notation

There are several levels of condensed structure notation; what they all have in common is the omission of certain kinds of bonds. At the first level of condensation, bonds to H atoms are not shown. Consider the full structure notation of geraniol (the compound responsible for the odor of roses and geraniums):



Pretty awkward, right? We now dispense with writing bonds to H and simply write the appropriate number of Hs after the atom on which they reside. We get a structure that is a little faster to write, but still not as economical as it could be:



At a more drastic level of condensation, no bonds whatsoever are shown:

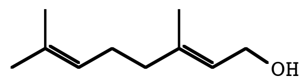
$$CH_3C(CH_3)CHCH_2CH_2C(CH_3)CHCH_2OH$$

In this type of condensed structure a group in parentheses implies that it is bonded to a preceding atom that is not in parentheses. Although this type of condensed structure is easy to type at a keyboard, it is difficult to interpret because the type of bonds (single, double, or triple) linking atoms together becomes apparent only after careful deconstruction -- not good.

Bond–line notation

The bond–line notation (also called line-segment notation) is the most efficient way of representing organic molecules. You want to concentrate on mastering this notation.

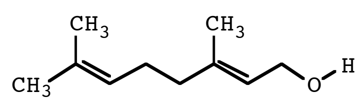
A bond–line structure portrays the carbon skeleton of a molecule. Carbon–carbon bonds are represented by line segments; no carbon–hydrogen bonds are shown. In bond–line structures, carbon atoms reside (1) at the end of line segments not terminated by other atoms and (2) where line segments meet. Usually, the points at which line segments meet look like zigzags, elbows, or crossroads. The bond–line structure of geraniol is



It is exceedingly important to remember when dealing with bond–line structures that each carbon atom still has its full complement of four bonds: if fewer than four are shown, it is assumed that there are enough carbon–hydrogen bonds present to bring the total up to four.

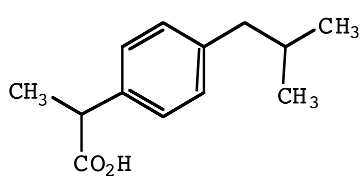
Mixed notation

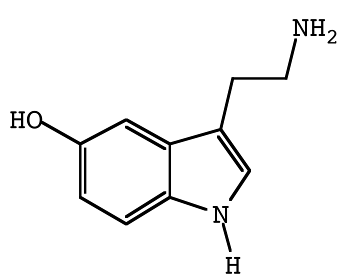
It's perfectly OK to mix features of full, condensed and bond–line notation all in the same structure. Returning to geraniol for one last time, we could write its structure as



This structure has the zigzags and crossroads typical of a bond–line notation, the CH_3 typical of a condensed structure, and an explicitly shown O–H bond typical of the full notation.

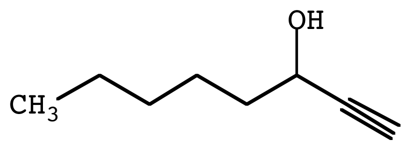
Problem Determine the molecular formula of these compounds.



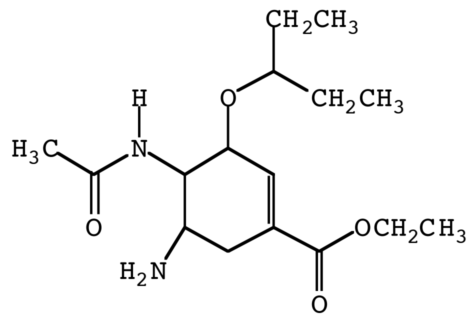


ibuprofen
analgesic

serotonin
neurotransmitter



1-octyn-3-ol
used in making the *breast-cancer drug Trichlorin A*

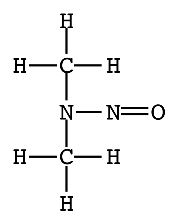


tamifluA
avian and swine flu drug

Answer ibuprofen ($C_{13}H_{18}O_2$); serotonin ($C_{10}H_{12}N_2O$); 1-octyn-3-ol ($C_8H_{14}O$); tamifluA ($C_{16}H_{28}N_2O_4$)

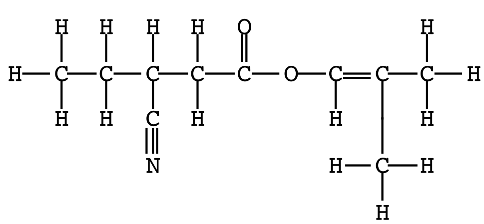
Problem Draw $N(CH_3)_2NO$ in full structure notation.

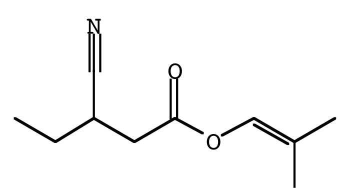
Answer



Problem Draw $CH_3CH_2CH(CN)CH_2C(O)OCHC(CH_3)CH_3$ in full structure notation and in bond–line notation.

Answer

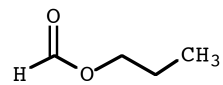




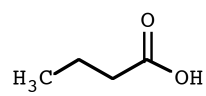
Constitutional isomers

An organic molecule's properties (e.g., molecular weight, melting point, boiling point, reactivity, color, odor, etc.) depend not only on how many and what kind of atoms make up that molecule, but also on the manner in which that molecule's constituent atoms are linked one to another.

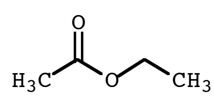
Consider the formula $C_4H_8O_2$. Four C atoms, eight H atoms, and two O atoms can combine in a vast number of ways: different arrangements of these atoms result in different compounds with different properties. Three possibilities are shown below:



• Propyl formate (boiling point 80 °C, melting point –93 °C) occurs naturally in currant, pineapple and plum. Because of its sweet smell (like rum with berry notes), it is used to make perfume.



• Butyric acid (boiling point 164 °C, melting point –5 °C) is found in vomit, rancid butter and parsnips. Nobody's making perfume out of this!

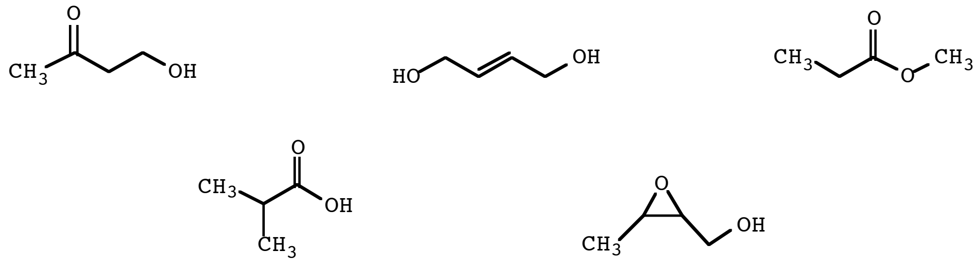


• Ethyl acetate (boiling point 77 °C, melting point –84 °C) is used by the food industry to counterfeit the presence of pineapple in ice cream and baked goods. It has a huge number of other uses as well: decaffeinating tea and coffee, nail polish and nail polish remover, solvent extraction of olive oil, etc.

Propyl formate, butyric acid and ethyl acetate are constitutional isomers of each other: they have the same formula ($C_4H_8O_2$), but differ in the way their constituent atoms are linked together. As we have seen, constitutional isomers generally have very different properties.

Problem Draw three additional constitutional isomers of $C_4H_8O_2$.

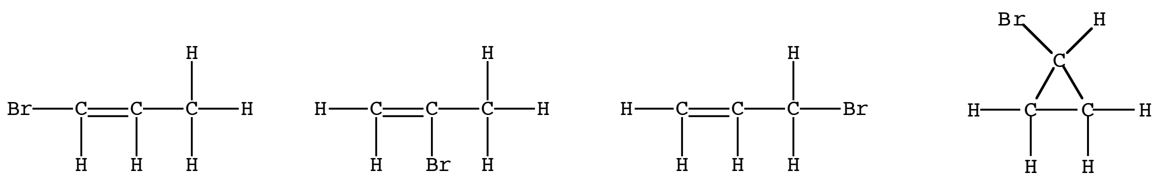
Answer Just a few of the myriad of possibilities are



Although you can let your imagination run wild to some extent when you're writing constitutional isomers, you must remember that, among the most important elements in organic chemistry, C ordinarily forms four bonds, N forms three bonds, O forms two bonds, and H and the Group VIIA elements form one bond.

Problem Draw the four constitutional isomers of C_3H_5Br.

Answer



Problem Are the two structures below constitutional isomers? Explain your answer.

Answer No: they are identical structures. We can get the structure on the right to look exactly like the structure on the left by flipping it like a pancake.

Degrees of unsaturaton

A handy method for translating a formula to a structure is to calculate the degrees of unsaturation (DOU) implied by that formula. The degrees of unsaturation are computed using

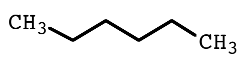
$$\text{DOU} = \frac{2C + 2 - H - X + N}{2}$$

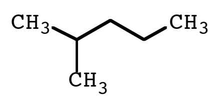
where C is the number of carbons, H the number of hydrogens, X the number of halogens, and N the number of nitrogens in the formula. The degrees of unsaturation indicate the number of rings and multiple bonds in the structure. A ring consumes one DOU, a double bond consumes one DOU, and a triple bond consumes two DOUs. Let's look at some examples.

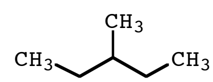
The formula C_6H_{14} has

$$\text{DOU} = \frac{2C + 2 - H - X + N}{2} = \frac{(2)(6) + 2 - 14 - 0 + 0}{2} = 0$$

Thus, C_6H_{14} has no rings, no double bonds, and no triple bonds. Some structures that fit the formula are



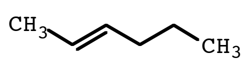


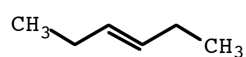


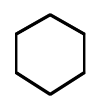
In constrast, the formula C_6H_{12} has

$$\text{DOU} = \frac{2C + 2 - H - X + N}{2} = \frac{(2)(6) + 2 - 12 - 0 + 0}{2} = 1$$

Thus, C_6H_{12} has one ring or one double bond, but it cannot have a triple bond. Some structures that fit the formula are





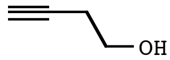
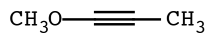
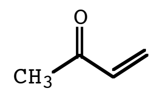
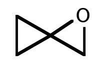
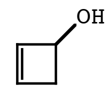


The DOU of a formula cannot tell you how those degrees of unsaturation are distributed. That is, if a formula has three DOUs, the structure may have three rings, two rings and one double bond, one ring and two double bonds, one ring and one triple bond, three double bonds, a double bond and a triple bond, etc.

Problem Calculate the DOU in C_4H_6O. Draw at least four structures that fit the formula.

Answer

$$\text{DOU} = \frac{2C + 2 - H - X + N}{2} = \frac{(2)(4) + 2 - 6 - 0 + 0}{2} = 2$$

<u>Problem</u> Calculate the DOU in C_5H_5N. Draw at least four structures that fit the formula.

<u>Answer</u>

$$\text{DOU} = \frac{2C + 2 - H - X + N}{2} = \frac{(2)(5) + 2 - 5 - 0 + 1}{2} = 4$$