**MOLECULAR FORMULAS AND WHAT CAN BE LEARNED FROM THEM**

**Before attempting to deduce the structure of an unknown organic compound from an examination of its spectra, we can simplify the problem somewhat by examining the molecular formula of the substance.**

**The purpose of this chapter is to describe how the molecular formula of a compound is determined and how structural information may be obtained from that formula.**

**ELEMENTAL ANALYSIS AND CALCULATIONS**

**The classical procedure for determining the molecular formula of a substance involves three steps:**

**1. A qualitative elemental analysis to find out what types of atoms are present . . . C, H, N,O, S, Cl, and so on.**

**2. A quantitative elemental analysis (or microanalysis) to find out the relative numbers (percentages) of each distinct type of atom in the molecule.**

**3. A molecular mass (or molecular weight) determination.**

**The first two steps establish an empirical formula for the compound. When the results of the third procedure are known, a molecular formula is found.**

**Virtually all organic compounds contain carbon and hydrogen. In most cases, it is not necessary to determine whether these elements are present in a sample: their presence is assumed.**

**However, if it should be necessary to demonstrate that either carbon or hydrogen is present in a compound, that substance may be burned in the presence of excess oxygen.**

**If the combustion produces carbon dioxide, carbon must be present; if combustion produces water, hydrogen atoms must be present. Today, the carbon dioxide and water can be detected by gas chromatographic methods.**

**Sulfur atoms are converted to sulfur dioxide; nitrogen atoms are often chemically reduced to nitrogen gas following their combustion to nitrogen oxides.**

**Oxygen can be detected by the ignition of the compound in an atmosphere of hydrogen gas; the product is water.**

**The percentages of carbon and hydrogen in the original sample were calculated by simple stoichiometry. Table shows a sample calculation.**



Notice in this calculation that the amount of oxygen was determined by difference, a common practice. In a sample containing only C, H, and O, one needs to determine the percentages of only C and H; oxygen is assumed to be the unaccounted-for portion.

You may also apply this practice in situations involving elements other than oxygen; if all but one of the elements is determined, the last one can be determined by difference.

Table shows how to determine the empirical formula of a compound from the percentage compositions determined in an analysis.   
Remember that the empirical formula expresses the simplest whole-number ratios of the elements and may need to be multiplied by an integer to obtain the true molecular formula.



**INDEX OF HYDROGEN DEFICIENCY**

The index of hydrogen deficiency (sometimes called the unsaturation index) is the number of pi bonds and/or rings a molecule contains.  
Frequently, a great deal can be learned about an unknown substance simply from knowledge of its molecular formula. This information is based on the following general molecular formulas:



**When the molecular formula for a compound contains noncarbon or nonhydrogen elements, the ratio of carbon to hydrogen may change. Following are three simple rules that may be used to predict how this ratio will change:**

**1.) To convert the formula of an open-chain, saturated hydrocarbon to a formula containing Group V elements (N, P, As, Sb, Bi), one additional hydrogen atom must be *added* to the molecular formula for each such Group V element present.**

**In the following examples, each formula is correct for a two-carbon acyclic, saturated compound:**

**C2H6, C2H7N, C2H8N2, C2H9N3**

**2.) To convert the formula of an open-chain, saturated hydrocarbon to a formula containing Group VI elements (O, S, Se, Te), *no change* in the number of hydrogens is required.**

**In the following examples, each formula is correct for a two-carbon, acyclic, saturated compound:**

**C2H6, C2H6O, C2H6O2, C2H6O3**

**3.)To convert the formula of an open-chain, saturated hydrocarbon to a formula containing Group VII elements (F, Cl, Br, I), one hydrogen must be *subtracted* from the molecular formula for each such Group VII element present.**

**In the following examples, each formula is correct for a two-carbon, acyclic, saturated compound:**

**C2H6, C2H5F, C2H4F2, C2H3F3**

Table presents some examples that should demonstrate how these correction numbers were determined for each of the heteroatom groups.



**The index of hydrogen deficiency (sometimes called the unsaturation index) is the number of pi bonds and/or rings a molecule contains.**

**It is determined from an examination of the molecular formula of an unknown substance and from a comparison of that formula with a formula for a corresponding acyclic, saturated compound.**

**The difference in the number of hydrogens between these formulas, when divided by 2, gives the index of hydrogen deficiency.**

**The index of hydrogen deficiency can be very useful in structure determination problems.**

**A great deal of information can be obtained about a molecule before a single spectrum is examined.**

**For example, a compound with an index of one must have one double bond or one ring, but it cannot have both structural features.**

**A quick examination of the infrared spectrum could confirm the presence of a double bond. If there were no double bond, the substance would have to be acyclic and saturated.**

**A compound with an index of two could have a triple bond, or it could have two double bonds, two rings, or one of each.**

**Benzene contains one ring and three “double bonds” and thus has an index of hydrogen deficiency of four.**

**Any substance with an index of *four* or more may contain a benzenoid ring; a substance with an index less than *four* cannot contain such a ring.**

**To determine the index of hydrogen deficiency for a compound, apply the following steps:**

**1. Determine the formula for the saturated, acyclic hydrocarbon containing the same number of carbon atoms as the unknown substance.**

**2. Correct this formula for the nonhydrocarbon elements present in the unknown.**

**Add one hydrogen atom for each Group V element present and subtract one hydrogen atom for each Group VII element present.**

**3. Compare this formula with the molecular formula of the unknown. Determine the number of hydrogens by which the two formulas differ.**

**4. Divide the difference in the number of hydrogens by two to obtain the index of hydrogen deficiency.**

**This equals the number of pi bonds and/or rings in the structural formula of the unknown substance.**

**Example 1:**

The unknown substance introduced at the beginning has the molecular formula C7H14O2.

1. Using the general formula for a saturated, acyclic hydrocarbon (C*n*H2*n*+2, where *n* = 7), calculate the formula C7H16.

2. Correction for oxygens (no change in the number of hydrogens) gives the formula C7H16O2.

3. The latter formula differs from that of the unknown by two hydrogens.

4. The index of hydrogen deficiency equals one. There must be one ring or one double bond in the unknown substance.

Having this information, the chemist can proceed immediately to the double-bond regions of the infrared spectrum.

There it finds evidence for a carbon–oxygen double bond (carbonyl group).





**Example 2:**

Nicotine has the molecular formula C10H14N2.

1. The formula for a 10-carbon, saturated, acyclic hydrocarbon is C10H22.

2. Correction for the two nitrogens (add two hydrogens) gives the formula C10H24N2.

3. The latter formula differs from that of nicotine by 10 hydrogens.

4. The index of hydrogen deficiency equals five.

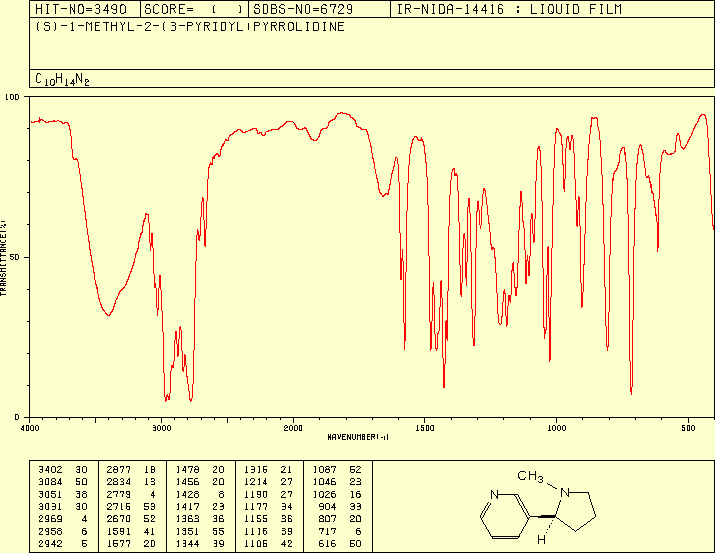
There must be some combination of five pi bonds and/or rings in the molecule.

Since the index is greater than *four,* a benzenoid ring could be included in the molecule.

Analysis of the spectrum quickly shows that a benzenoid ring is indeed present in nicotine.

More careful refinement of the spectral analysis leads to a structural formula for nicotine:





**Example 3:**

**Chloral hydrate (“knockout drops”) is found to have the molecular formula C2H3Cl3O2.**

**1. The formula for a two-carbon, saturated, acyclic hydrocarbon is C2H6.**

**2. Correction for oxygens (no additional hydrogens) gives the formula C2H6O2.**

**3. Correction for chlorines (subtract three hydrogens) gives the formula C2H3Cl3O2.**

**4. This formula and the formula of chloral hydrate correspond exactly.**

**5. The index of hydrogen deficiency equals zero. Chloral hydrate cannot contain rings or double bonds.**

**Examination of the spectral results is limited to regions that correspond to singly bonded structural features.**

**The correct structural formula for chloral hydrate follows. You can see that all of the bonds in the molecule are single bonds.**

