Uploaded to VFC Website
~October 2012~

This Document has been provided to you courtesy of Veterans-For-Change!

Feel free to pass to any veteran who might be able to use this information!

For thousands more files like this and hundreds of links to useful information, and hundreds of "Frequently Asked Questions, please go to:

Veterans-For-Change

11901 Samuel Drive, Garden Grove, CA 92840-2546

Chemical formula

A **chemical formula** or **molecular formula** is a way of expressing information about the atoms that constitute a particular chemical compound.

The chemical formula identifies each constituent element by its chemical symbol and indicates the number of atoms of each element found in each discrete molecule of that compound. If a molecule contains more than one atom of a particular element, this quantity is indicated using a subscript after the chemical symbol (although 18th-century books often used superscripts) and also can be combined by more chemical elements.

Chemical formulas may be used in chemical equations to describe chemical reactions.
For ionic compounds and other non-molecular substances an empirical formula may be used, in which the subscripts indicate the ratio of the elements.

The 19th-century Swedish chemist Jöns Jakob Berzelius worked out this system for writing chemical formulas.

Molecular geometry and structural formulas

The exact connectivity among the atoms in a molecule and/or the exact spatial arrangement of those atoms is of utmost importance. Two molecules composed of the same numbers of the same types of atoms might have completely different chemical and/or physical properties if the atoms are connected differently or in different positions. Thus, sometimes one needs to use a structural formula to illustrate which atoms are bonded to which other ones and somehow describe the resulting molecular geometry of the compound.

For example methane, a simple molecule consisting of one carbon atom bonded to four hydrogen atoms, has the chemical formula:

CH_4 or H_4C

and glucose with six carbon atoms, twelve hydrogen atoms and six oxygen atoms has the chemical formula:

$C_6H_{12}O_6$.

A chemical formula supplies information about the types and spatial arrangement of bonds in the chemical, though it does not necessarily specify the exact isomer. For example ethane consists of two carbon atoms single-bonded to each other, with each carbon atom having three hydrogen atoms bonded to it. Its chemical formula can be rendered as CH_3CH_3. In ethylene there is a double bond between the carbon atoms (and thus each carbon only has two hydrogens), therefore the chemical formula may be written: CH_2CH_2, and the fact that there is a double bond between the carbons is implicit because carbon has a valence of four. However, a more explicit and correct method is to write $H_2C{=}CH_2$ or less commonly $H_2C{::}CH_2$. The two lines (or two pairs of dots) indicate that a double bond connects the atoms on either side of them.

A triple bond may be expressed with three lines or pairs of dots, and if there may be ambiguity, a single line or pair of dots may be used to indicate a single bond.

Molecules with multiple functional groups that are the same may be expressed in the following way: $(CH_3)_3CH$. However, this implies a different structure from other molecules that can be formed using the same atoms (isomers). The formula $(CH_3)_3CH$ implies a chain of three carbon atoms, with the middle carbon atom bonded to another carbon (*see image of 4 carbon "C" atoms*), and the remaining bonds on the carbons all leading to hydrogen atoms (*hydrogen atoms are not shown in image*). However, the same number of atoms (10 hydrogens and 4 carbons, or C_4H_{10}) may be used to make a straight chain: $CH_3CH_2CH_2CH_3$.

The alkene but-2-ene has two isomers which the chemical formula $CH_3CH{=}CHCH_3$ does not identify. The relative position of the two methyl groups must be indicated by additional notation denoting whether the methyl groups are on the same side of the double bond (*cis* or *Z*) or on the opposite sides from each other (*trans* or *E*).

Polymers

For polymers, parentheses are placed around the repeating unit. For example, a hydrocarbon molecule that is described as: $CH_3(CH_2)_{50}CH_3$, is a molecule with fifty repeating units. If the number of repeating units is unknown or variable, the letter *n* may be used to indicate this formula: $CH_3(CH_2)_nCH_3$.

Ions

For ions, the charge on a particular atom may be denoted with a right-hand superscript. For example Na^+, or Cu^{2+}. The total charge on a charged molecule or a polyatomic ion may also be shown in this way. For example: hydronium, H_3O^+ or sulfate, SO_4^{2-}.

For more complex ions, brackets [ ] are often used to enclose the ionic formula, as in $[B_{12}H_{12}]^{2-}$. Parentheses ( ) can be nested inside brackets to indicate a repeating unit, as in $[Co(NH_3)_6]^{3+}$. Here $(NH_3)_6$ indicates that the ion contains six NH_3 groups, and [ ] encloses the entire formula of the ion with charge +3.

Isotopes

Although isotopes are more relevant to nuclear chemistry or stable isotope chemistry than to conventional chemistry, different isotopes may be indicated with a left-hand superscript in a chemical formula. For example, the phosphate ion containing radioactive phosphorus-32 is $^{32}PO_4^{3-}$. Also a study involving stable isotope ratios might include the molecule $^{18}O^{16}O$.

A left-hand subscript is sometimes used redundantly to indicate the atomic number. For example, $_8O_2$ for dioxygen, and $^{16}_{8}O_2$ for the most abundant isotopic species of dioxygen. This is convenient when writing equations for nuclear reactions, in order to show the balance of charge more clearly.

Empirical formulas

Main article: Empirical formula

In chemistry, the empirical formula of a chemical is a simple expression of the relative number of each type of atom or ratio of the elements in the compound. Empirical formulas are the standard for ionic compounds, such as $CaCl_2$, and for macromolecules, such as SiO_2. An empirical formula makes no reference to isomerism, structure, or absolute number of atoms. The term **empirical** refers to the process of elemental analysis, a technique of analytical chemistry used to determine the relative percent composition of a pure chemical substance by element.

For example hexane has a molecular formula of C_6H_{14}, or structurally $CH_3CH_2CH_2CH_2CH_2CH_3$, implying that it has a chain structure of 6 carbon atoms, and 14 hydrogen atoms. However, the empirical formula for hexane is C_3H_7. Likewise the empirical formula for hydrogen peroxide, H_2O_2, is simply HO expressing the 1:1 ratio of component elements. Formaldehyde and acetic acid have the same empirical formula, CH_2O. This is the actual chemical formula for formaldehyde, but acetic acid has double the number of atoms.

Trapped atoms

The @ symbol ("at") indicates an atom or molecule trapped inside a cage but not chemically bound to it. This notation became popular in the 1990s with the discovery of fullerene cages, which can trap atoms such as La to form $La@C_{60}$ or $La@C_{82}$ for example. A non-fullerene example is $[As@Ni_{12}As_{20}]^{3-}$, an ion in which one As atom is trapped in a cage formed by the other 32 atoms.

Non-stoichiometric formulas

Main article: Non-stoichiometric compound

Chemical formulas most often use integers for each element. However, there is a whole class of compounds, called non-stoichiometric compounds, that cannot be represented by small integers. Such a formula might be written using decimal fractions, as in $Fe_{0.95}O$, or it might include a variable part represented by a letter, as in $Fe_{1-x}O$, where x is normally much less than 1.

General forms for organic compounds

Chemical formula used for a series of compounds that differ from each other by a constant unit is called **general formula**. Such a series is called the homologous series, while its members are called homologs.

Hill System

Main article: Hill system

The **Hill system** is a system of writing **chemical formulas** such that the number of carbon atoms in a molecule is indicated first, the number of hydrogen atoms next, and then the number of all other chemical elements subsequently, in alphabetical order. When the formula contains no carbon, all the elements, including hydrogen, are listed alphabetically. This deterministic system enables straightforward sorting and searching of compounds.

See also

- Empirical formula
- Structural formula
- Dictionary of chemical formulas
- Dictionary of chemical formulas/Merge
- Element symbol
- Nuclear notation
- Periodic table

References

Ralph S. Petrucci, William S. Harwood and F. Geoffrey Herring, *General Chemistry*, 8th Edition (Prentice-Hall 2002), chapter 3.