The IUPAC Rules for Naming Organic Molecules

Stanislaw Skonieczny
Department of Chemistry, University of Toronto, Toronto, Ontario, M5S 3H6, Canada; *sskoniec@chem.utoronto.ca*

Textbooks used for teaching undergraduate courses provide basic rules and examples for naming monofunctional compounds. However, students of more advanced organic chemistry courses and graduate students dealing with more complex molecules need a set of additional rules and guidelines to identify polyfunctional compounds. The following summary of general principles may also be useful to instructors. The formation of the systematic name for an organic compound involves several steps, to be taken as far as they are applicable in the following order *(1–5)*.

Step 1

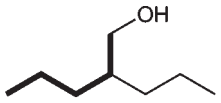
Determine the principal functional group in the compound. When a compound contains more than one group in Table 1, the principal group is that which has the highest precedence. This principal group will be cited as a suffix; all other groups are cited as prefixes.

Step 2

Determine the parent hydrocarbon (principal chain or parent ring system):

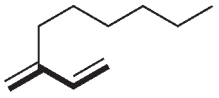
(a) If the principal group occurs in a chain, the principal chain is selected as

(i) the chain containing the functional group of the highest seniority as the parent.



2-(propan-1-yl)pentan-1-ol

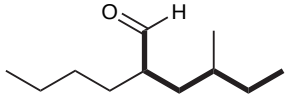
(ii) If more than one such a choice is possible, select the chain with the maximum number of multiple bonds.



2-(hexan-1-yl)buta-1,3-diene

(iii) If (i) and (ii) together are not definitive, then choose the chain with maximum length.

(iv) If two chains of the same length are possible, choose the one with maximum number of substituents.



2-(butan-1-yl)-4-methylhexanal

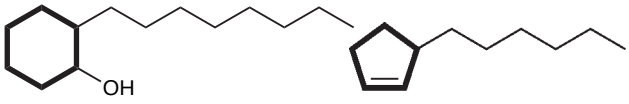
The parent chain here contains six carbon atoms.

Table 1. The Decreasing Order of Precedence for the Major Functional Groups[a]

Functional Group	Suffix if Highest Precedence	Prefix if Lower Precedence
-(C)OOH[b]	-oic acid	—
-COOH	-carboxylic acid	carboxy-
-SO_2OH	-sulfonic acid	sulfo-
-C(O)OC(O)-	-oic anhydride	—
-(C)OOR[b]	alkyl -oate	alkoxy-oxo-
-COOR	alkyl carboxylate	—
-(C)OHal[b]	-oyl halide	halo-oxo-
-COHal	-carbonyl halide	halocarbonyl-
-(C)ONH_2[b]	-amide	amino-oxo-
-$CONH_2$	-carboxamide	aminocarbonyl-
-(C)≡N[b]	-nitrile	—
-C≡N	-carbonitrile	cyano-
-(C)HO[b]	-al	oxo-
-CHO	-carbaldehyde	formyl-
>(C)=O	-one	oxo-
-OH	-ol[c]	hydroxy-
-SH	-thiol	sulfanyl-[d]
-NH_2	-azane (-amine)[d]	azanyl- (amino-)[e]
>NH	-azane (-imine)[d]	azanylidene- (imino-)[e]
>C=C<	-ene	*
-C≡C-	-yne	**
-X, -R, -OR	***	halo-, alkyl-, alkoxy-
-Ar, -NO_2,	***	aryl-, nitro-
-NO, -N_3,	***	nitroso-, azido-
=N_2, -SR	***	diazo-, (*R*)-sulfanyl

[a]Systematic and traditional trivial names for a number of common compounds representing most of the functional groups listed in the table are given in the Suplplemental Material. [b](C) designates a carbon atom included in the name of the parent hydride (it does not belong to a group designated by a suffix or a prefix). [c]Among the phenols, some trivial names are acceptable. Among them are phenol, 1-naphthol, 1-anthrol, 1-phenanthrol, and their positional isomers. Others, such as cresol, thymol, and picric acid are still accepted by IUPAC when unsubstituted, but it is best to name them by systematic nomenclature. [d]In the 1998 recommendations, the prefix 'sulfanyl-' is preferred to 'mercapto-', which was used previously. [e]IUPAC recommends -azane as an alternative to -amine and imine, but this has yet to see much usage. All text-books, catalogues, reference books, and scientific literature still use -amine (amino-) and -imine (imino-), which are acceptable. *Always used as a suffix. However, if a higher priority group is present in the molecule, the suffix is changed to -en- and it is followed by the suffix of the higher priority group. **Always used as a suffix. However, if a higher priority group is present in the molecule, the suffix is changed to -yn- and it is followed by the suffix of the higher priority group. *** Always used as a prefix.

(b) If the principal functional group occurs in a cyclic system, that system forms the parent.



2-(octan-1-yl)cyclohexanol 3-(hexan-1-yl)cyclopentene

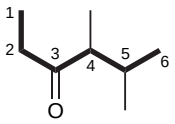
Step 3

Name the parent structure and the principal group(s). Give the parent the same name as if it were an alkane, but replace *-ane* with the suffix characteristic of the functional group of the highest priority (Table 1). Note: Multiple unsaturation in hydrocarbons is indicated by the numerical prefixes *di-*, *tri-*, etc. In such cases, however, the ending *-ane* of the parent alkane is replaced with *-adiene*, *-atriene*, etc. leaving the *a* in the root alkane whenever the first letter of the suffix is a consonant to make the name easier to pronounce. Thus we get alkadienes, alkatriens, alkadiynes, etc.

Step 4

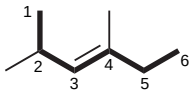
Complete the numbering.

(a) Number the carbon atoms consecutively from the end of the chain nearer the functional group of the highest priority.



4,5-dimethylhexan-3-one

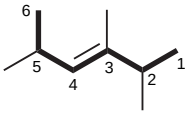
(b) If the same number for the functional group of highest priority is obtained in both directions, the correct IUPAC name is the one that contains the lowest substituent number.



2,4-dimethylhex-3-ene

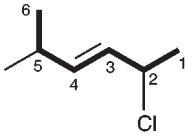
First substituent has number 2. If we numbered the chain from the other side, the first substituent would have number 3.

(c) If first substituents occur at an equal distance from each end of the chain, number from the end nearer a second substituent. If still the same, continue until first point of difference.



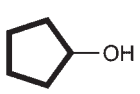
(*E*)-2,3,5-trimethylhex-3-ene

(d) If numbering from either side gives the same set of locants, number from the end with the substituent that comes first in alphabetical order.

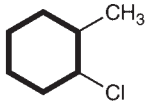


2-chloro-5-methylhex-3-ene

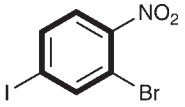
(e) In monosubstituted cycloalkanes and benzene derivatives, the carbon to which the substituent is bonded is always number 1. If there is only one substituent, the locant "1" is omitted. However, it is necessary in polysubstituted compounds.



cyclopentanol



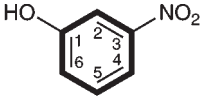
1-chloro-2-methylcyclohexane



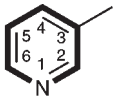
2-bromo-4-iodo-1-nitrobenzene

(f) In polysubstituted cyclic compounds, the starting point and the direction of numbering are chosen to give lowest locant to the following factors (if present), considered successively in the order listed below until a decision is reached.

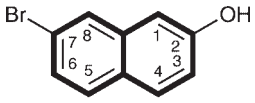
(i) The principal functional group named as suffix is always given lowest possible number. When the numbering is predetermined by the nature of the parent hydrocarbon, as in polycyclic hydrocarbons and heterocyclic compounds, lowest locants are still the rule.



IUPAC: 3-nitrobenzen-1-ol
Accepted: 3-nitrophenol

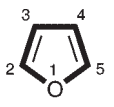


3-methylazabenzene
3-methylpyridine



7-bromonaphthalene-2-ol
7-bromo-2-naphthol

Note: When one heteroatom is present in the ring, the locant "1" is attributed to the heteroatom. If two or more kinds of heteroatom occur in the same ring, the numbering starts from the heteroatom cited first in the following descending order: O, S, Se, N, P, As, Sb, Si, Ge, Sn, Pb, B, and Hg. Other heteroatoms get other lowest possible locants.

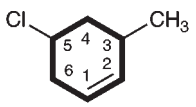


IUPAC: oxole[1]
Accepted: furan

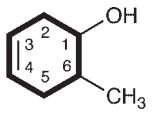
1,2-oxathiolane

1,3,5-triazine

(ii) Lowest locant for multiple bonds in cyclic compounds.



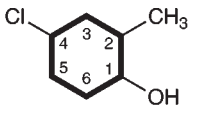
5-chloro-3-methylcyclohexene



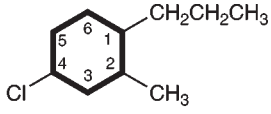
6-methylcyclohex-3-en-1-ol

Note: In naming cycloalkenes, the double bond is located between C1 and C2, and the "1" is usually omitted in the name. The ring is numbered clockwise or counterclockwise to give the first substituent the lower number. However, when a higher seniority functional group is present, the carbon bonded to such a group becomes C1 and then multiple bonds are considered for lowest possible locants.

(iii) Lowest locant for a first substituent (or first point of difference).

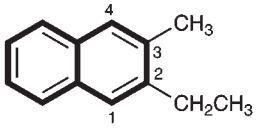


4-chloro-2-methylcyclohexan-1-ol

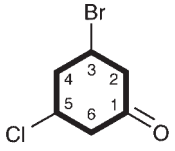


4-chloro-2-methyl-1-(propan-1-yl)cyclohexane

(iv) If the same numbers will result from numbering in either direction, give the lowest locant for that substituent which is cited first.



2-ethyl-3-methylnaphthalene



3-bromo-5-chlorocyclohexanone

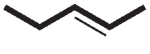
Step 5

Name alkyl groups, halides, and other substituents and determine their position on the chain by the numbering established by step 4.

Step 6

Assign the stereochemistry to chiral carbon atoms (stereocenters) and the double bond(s):

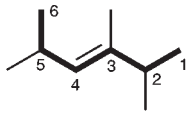
(a) Disubstituted alkenes may be named in two ways: (i) using terms *cis-* and *trans-* or (ii) using terms (*E*) or (*Z*). Note: Use of (*E*–*Z*) stereodescriptors is preferred. However, the *cis*–*trans* terms are still accepted.



(*E*)-pent-2-ene or *trans*-pent-2-ene

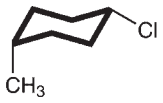
(*Z*)-hex-3-ene or *cis*-hex-3-ene

(b) Tri- and tetra-substituted alkenes may only be named using the (*E*–*Z*) system.



(*E*)-2,3,5-trimethylhex-3-ene

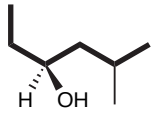
(c) Disubstituted cycloalkanes may be named using the terms *cis–trans* or (*R*–*S*).



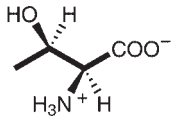
cis-1-chloro-4-methylcyclohexane
(*R*–*S*) system is not applicable here

trans-3-methylcyclopentanol
(1*R*,3*R*)-3-methylcyclopentanol

(d) Chiral carbon atoms are assigned (*R*) or (*S*) configuration.

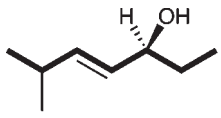


(*S*)-5-methylhexan-3-ol

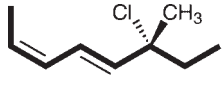


(2*S*,3*R*)-2-amino-3-hydroxybutanoic acid
trivial: (–)-theonine

(e) When both (*R*–*S*) and (*E*–*Z*) stereodescriptors are present, they are placed in parentheses followed by a hyphen; each stereodescriptor is immediately preceded by the lower or less primed locant and they are arranged in numerical order.



(3*S*,4*E*)-6-methylhept-4-en-3-ol



(2*Z*,4*E*,6*R*)-6-chloro-6-methylocta-2,4-diene

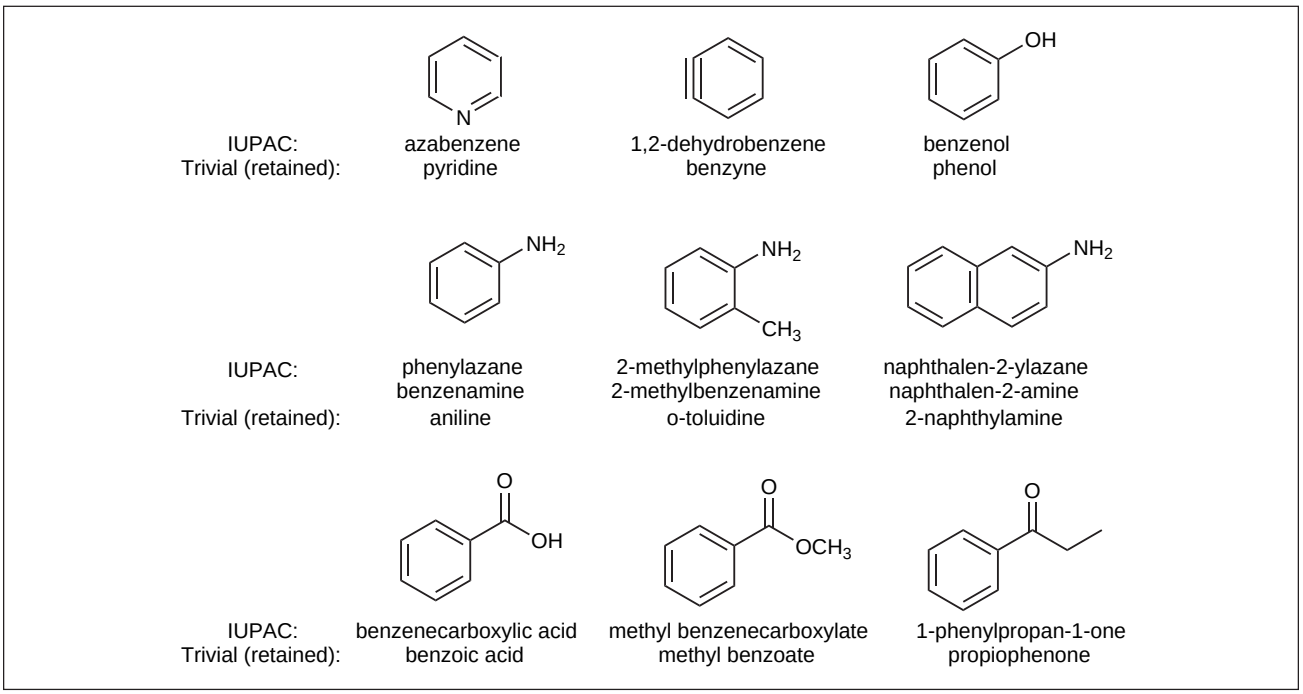
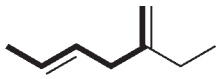


Figure 1. Some examples of IUPAC names for commonly known compounds

Step 7

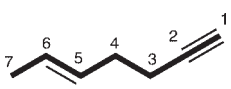
Compounds with multiple double and triple bonds have the following suffixes: *-adiene*, *-adiyne*, *-atriene*, *-atetraene*, and so on. Specify the location of each multiple bond by a locant placed between *-a* and *-diene*, *-diyne*, *-triene*, etc.



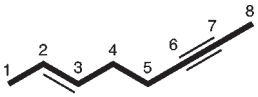
(*E*)-2-ethylhexa-1,4-diene
previously: (*E*)-2-ethyl-1,4-hexadiene

Step 8

Compounds with both double bonds and triple bonds are called enynes. Start the numbering of compounds with both double and triple bonds from the end nearer the first multiple bond, regardless of type. When a double bond is the same distance from one end as a triple bond from the other end, assign the double bond the lower number.



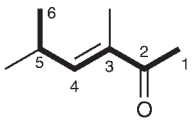
(*E*)-hept-5-en-1-yne



(*E*)-oct-2-en-6-yne

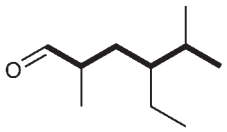
Step 9

Write the complete name of the compound as a single word with the correct locants for all substituents, which are listed in alphabetical order. The stereochemistry is indicated by placing the appropriate prefix within parentheses followed by a hyphen in front of the name.



(*E*)-3,5-dimethylhex-3-en-2-one

The prefixes *di-*, *tri-*, *tetra-*, etc. do not alter the alphabetical ordering of substituents.



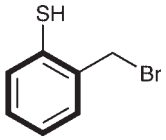
4-ethyl-2,5-dimethylhexanal

(a) The numbers indicating the locations of substituents are separated by commas.

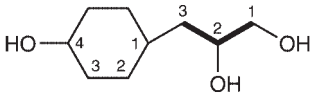
(b) The letters and numbers are joined by hyphens.

Step 10

A complex substituent is named by applying the above steps just as if the substituent were a compound itself. We begin numbering at the point of attachment to the parent hydrocarbon and we give this substituent the suffix *-yl* and put its name in parentheses when the name of the complete compound is given.



2-(bromomethyl)benzenethiol



3-(4-hydroxycyclohexyl)propane-1,2-diol

Note: There are many traditional names accepted by IUPAC and retained because of their wide use in organic, biochemical, and polymer nomenclature. A few examples are given here: toluene, styrene, phenol, aniline, acetone, acetylene, acetic acid, benzoic acid, and phthalic acid (Figure 1). A list of such compounds can be found on the Web *(6).*

Other names are retained for referring to unsubstituted compounds only. Compounds derived from them by substitution must be named systematically. For example, the names butanoic acid (systematic) and butyric acid (retained) are both approved by IUPAC for the unsubstituted acid. However, $BrCH_2CH_2CH_2COOH$ must be named systematically as 4-bromobutanoic acid. Some other examples of such names retained for unsubstituted compounds only include lactic acid, toluidine, anisole, glycerol, propionic acid, isoprene, cumene, pinacol, and many others.

[W]Supplemental Material

Examples of names of organic compounds, seniority of chains in naming, seniority of ring systems in naming, and the Hantzsch–Widman sytem of naming are available in this issue of *JCE Online*.

Note

1. The extended Hantzsch–Widman system used to name heterocycles is presented in the Supplemental Material.[W]

Literature Cited

1. Fox, R. B.; Powell, W. H. *Nomenclature of Organic Compounds, Principles and Practice,* 2nd ed.; Oxford University Press: Oxford, 2001.
2. Hellwinkel, D. *Systematic Nomenclature of Organic Chemistry, A Directory to Comprehension and Application of its Basic Principles;* Springer: Berlin, New York, London, 2001.
3. Leigh, G. J.; Favre, H. A.; Metanomski, W. V. *Principles of Chemical Nomenclature, A Guide to IUPAC Recommendations;* Blackwell Science: Oxford, 1998.
4. Panico, R.; Powell, W. H.; Richer, J.–C. *A Guide to IUPAC Nomenclature of Organic Compounds, Recommendations 1993;* Blackwell Science: Oxford, 1993.
5. Rigaudy, J.; Klesney, S. P. *Nomenclature of Organic Chemistry, Sections A, B, C, D, E, F, and H;* Pergamon Press: Oxford, 1979.
6. Trivial and Semisystematic Names Retained for Naming Organic Ccompounds. *http://www.acdlabs.com/iupac/nomenclature/93/r93_671.htm* (accessed Jul 2006).