# How to Deduce the Structure of an Organic Compound from its Name

(Try the problems at https://cheminteractive.ie/nomen1.php; you will get feedback based on the strategy below.)

Systematic names contain all the information needed to deduce the corresponding structures, and they are constructed in a consistent way with the constituent parts always in the same order. More than one nomenclature system is in use, and non-systematic names are also common, and this simplified summary only deals with the IUPAC nomenclature system, as exemplified by the name below. The components of IUPAC names are described beneath the example, together with the strategy for deducing the corresponding structures. A basic knowledge of alkane and cycloalkane nomenclature is assumed. Two additional examples follow at the foot of the page, and there are several more on page 2.

**(2*R*,3*S*,*E*)-7-bromo-2,6-dimethylhept-5-en-1,3-diol**

**1. Draw the parent compound.**

Most IUPAC names are based on the concept that the structures are modifications of '*parent*' compounds, which are almost always unbranched chains or unbranched rings. A name can only contain one parent. The first step in constructing the structure is to draw the parent chain or ring.

As well as alkanes, and cycloalkanes, common parents include the rings listed in **Tables 1 and 2** on page 2. Some parent compounds include a 'principal functional group' (see below) as well as a chain or ring. This category includes some very common parents, e.g. acetic acid ($CH_3CO_2H$), see **Table 3** for important examples.

Some non-systematic parent names are retained for historical reasons, e.g. styrene ($C_6H_5CH{=}CH_2$).

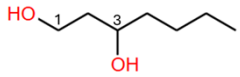
Hept means that the parent is heptane. Number the carbons from one (either) end.


**2. Add the principal functional group(s) to the parent.**

If the name ends with anything other than –ane, –ene or –yne, a '*principal functional group*' is present. Some common functional groups (FGs) that can be listed as suffixes at the end of compound names are shown in **Table 4** on page 2. Note that some FGs are named differently if the parent compound is a ring because they include a carbon that can have only one additional bond, so they must be external to the ring. If the compound is acyclic, those FGs must be at the end of the parent chain. '*Locants*' give the position(s) of the FG(s) on the parent, and '*multiplicative prefixes*' (di-, tri-, etc.) indicate more than one instance of the principal FG.

When present, a principal FG 'controls' the numbering of the parent. For rings, the carbon bearing the FG is numbered 1. For chains that include a principal FG that must be at the end of the chain, the carbon that is included in the FG is numbered 1, e.g. propanal ($CH_3CH_2CHO$) is numbered from the aldehyde carbon. In these cases, no locant is provided for the principal FG (it is position 1 by definition). The second step is to add the principal FG(s) at the correct position(s).

**1,3-diol** means that the principal FG is an alcohol, and there are two, at positions 1 and 3.



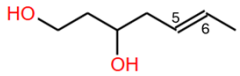

**3. Add the unsaturation to the parent.**

If the ending of the name of the parent is not –an(e), carbon-carbon double or triple bonds are present, and the next step is to add them to the structure.

If the ending is –en(e), a C=C double bond is present, and the locant is the number of the first C of the double bond (e.g. hept-**5**-en indicates a double bond between carbons 5 and 6 of the parent chain. Likewise, -yn(e) indicates the presence of a C≡C triple bond.

'*Multiplicative prefixes*' indicate that more than one double or triple bond, e.g. deca-1,3,5-**triene**-9-**yne** is a 10-carbon chain with double bonds at positions 1, 3 and 5, and a triple bond at position 9.

**5-en** means that there is one double bond (and no triple) and it is between carbons 5 and 6,




**4. Add the substituents to the parent and/or the principal FG.**

Next, add the *substituents*, additional groups that are attached to the parent chain or ring.

Names of substituents (e.g. 7-**bromo**) on the parent chain/ring, with *locants* (e.g. **7**-bromo) specifying their locations, are given as prefixes to the parent name. *Multiplicative prefixes* (e.g. **di**methyl) mean that a substituent is present more than once.

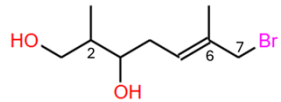
In addition to alkyl, and cycloalkyl, groups, and the substituents listed in **Tables 1 - 4**, common substituents include: vinyl ($CH_2{=}CH–$), isopropyl ($(CH_3)_2CH–$), *tert*-butyl ($(CH_3)_3C–$), phenyl ($C_6H_5–$) , and benzyl ($C_6H_5CH_2–$). Alkyloxy groups: methoxy ($CH_3O–$), ethoxy ($CH_3CH_2O–$), propoxy ($C_3H_7O–$), butoxy ($C_4H_9O–$), pentyloxy ($C_5H_{11}O–$) , etc.

Fluoro (F–), chloro (Cl–), bromo (Br–), iodo (I–), nitro ($O_2N–$).

The functional groups listed in **Table 4** on Page 2 can be present as substituents, rather than *principal* FGs, in which case they may be named quite differently; see the prefixes listed in the Table.

'*Compound substituted*' substituents are ones in which a substituent is itself substituted, e.g. chloromethyl is $ClCH_2–$. Work from right to left, drawing the substituent, numbering it from the point of attachment to the parent, then doing the same with its substituent, see ibuprofen.

**7-bromo-2,6-dimethyl** means that there is a bromo (Br) at position 7, and methyl groups ($CH_3$) at positions 2 and 6.

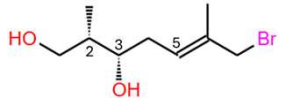



**5. Add the stereo-chemistry at chiral centres and/or alkenes.**

The final step in drawing the structure is to add the stereochemistry.

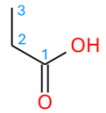
'*Stereodescriptors*' at the start of the name give the configurations at chirality centres (***R*** or ***S***) and at alkenes (***E*** or ***Z***), according to the Cahn-Ingold-Prelog (CIP) rules (see this excellent resource). *Cis* and *trans* are sometimes used to specify the stereochemistry of di-substituted alkenes, and the relative stereochemistry of di-substituted aliphatic rings. Locants (e.g. **2***R*) specify the positions on the parent chain/ring that the descriptors refer to. Check the *locants* given in the *stereodescriptors* and ensure that the correct configuration is drawn.

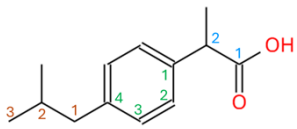
**(2*R*,3*S*,*E*)** means that the chirality centres at positions 2 and 3, and the alkene, have the *R*, *S* and *E* configurations, respectively. The structure is now complete. **(2*R*,3*S*,*E*)-7-bromo-2,6-dimethylhept-5-en-1,3-diol** is:




2-[4-(2-Methylpropyl)phenyl]propanoic acid (**ibuprofen**)

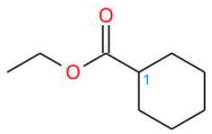
**1** Parent prop
**2** Principal FG oic acid
**3** Unsaturation an
**4** Substituent 2-[4-(2-Methylpropyl)phenyl]

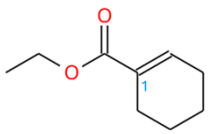


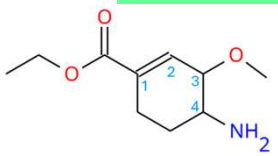



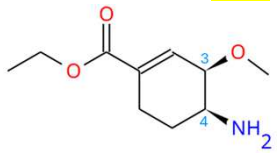

Ethyl (3*R*,4*S*)-4-amino-3-methoxycyclohex-1-enecarboxylate

**1** Parent cyclohex **2** Principal FG Ethyl –carboxylate **3** Unsaturation 1-ene **4** Substituents 4-amino-3-methoxy **5** Stereodescriptors 3*R*,4*S*



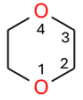
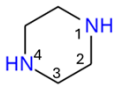



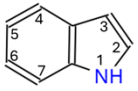
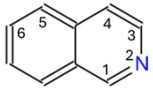
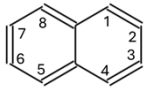
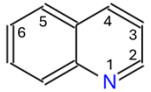
# How to Deduce the Structure of an Organic Compound from its Name continued (https://cheminteractive.ie/nomen1.php)

These Tables list names of common parent systems, and common principal functional group suffixes, and also give the prefixes when these parents/functional groups are present as substituents.

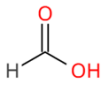
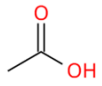
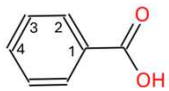
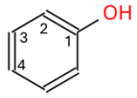
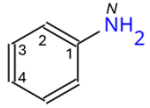
## Table 1: Common Non-Aromatic Heterocyclic Parent Rings

| Name | aziridine | 1,4-dioxane | morpholine | oxane | oxolane | oxirane | piperidine | piperazine | pyrrolidine |
|---|---|---|---|---|---|---|---|---|---|
| Structure |  |  |  |  |  |  |  |  |  |
| Substituent | aziridinyl– | 1,4-dioxanyl– | morpholinyl– | oxanyl– | oxolananyl– | oxiranyl– | piperidinyl– (or piperidyl–) | piperazinyl– | pyrrolidinyl– |

## Table 2: Common Aromatic Parent Rings

| Name | benzene | furan | imidazole | indole | isoquinoline | naphthalene | pyridine | pyrrole | quinoline | thiophene |
|---|---|---|---|---|---|---|---|---|---|---|
| Structure |  |  |  |  |  |  |  |  |  |  |
| FG prefix | phenyl– | furanyl– (or furyl–) | imidazolyl– | indolyl– | isoquinolinyl– (or isoquinolyl–) | naphthalenyl– (or naphthyl–) | pyridinyl– (or pyridyl–) | pyrrolyl– | quinolinyl– (or quinolyl–) | thiophenyl– (or thienyl–) |

## Table 3: Common Functional Parents

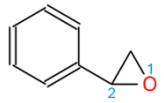
| Name | formic acid* | acetic acid* | benzoic acid* | phenol | aniline | urea |
|---|---|---|---|---|---|---|
| Structure |  |  |  |  |  |  |
| FG prefix | formyl– | acetyl– | benzoyl– | phenoxy– | anilino– | |

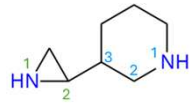
* Also, the esters (RC(=O)OR', acid halides (RC(=O)X), amide ($RC(=O)NH_2$, nitrile (RC≡N), and aldehyde (RC(=O)H

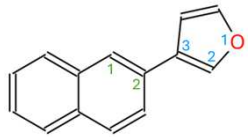
## Table 4: Common Functional Groups

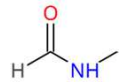
| Class | Carboxylic acids | (Carboxylic) Esters | Acid halides | Amides | Nitriles | Aldehydes | Ketones | Alcohols | Amines |
|---|---|---|---|---|---|---|---|---|---|
| FG suffix | –oic acid<br>–carboxylic acid | R –oate<br>R –carboxylate | –oyl halide<br>–carbonyl halide | –amide<br>–carboxamide | –nitrile<br>–carbonitrile | –al<br>–carbaldehyde | –one | –ol | –amine |
| Formula | –C*(=O)OH | –C*(=O)OR | –C*(=O)X | $–C*(=O)NH_2$ | –C*≡N | –C*(=O)H | =O | –OH | $–NH_2$ |
| Examples | $CH_3CO_2H$<br>Acetic acid<br>$C_3H_5CO_2H$<br>Cyclopropane-carboxylic acid | $CH_3CH_2CO_2CH_3$<br>Methyl propanoate<br>$C_5H_9CO_2CH_2CH_3$<br>Ethyl cyclopentanoate | $CH_3CH_2CH_2COCl$<br>Butanoyl chloride<br>$C_4H_7COCl$<br>Cyclobutane-carbonyl chloride | $CH_3CH_2CONH_2$<br>Propanamide<br>$C_6H_{11}CONH_2$<br>Cyclohexane-carboxamide | $CH_3CH_2CH_2C≡N$<br>Butanenitrile<br>$C_4H_7C≡N$<br>Cyclobutane-carbonitrile | $CH_3CHO$<br>Acetaldehyde (ethanal)<br>$C_3H_5C(=O)H$<br>Cyclopropane-carbaldehyde | $CH_3C(=O)CH_2CH_2CH_3$<br>Pentan-2-one | $CH_3CH(OH)CH_3$<br>Propan-2-ol | $C_4H_7–NH_2$<br>Cyclobutanamine |
| FG prefix | carboxy– | alkoxycarbonyl– | halocarbonyl– | carbamoyl– | cyano– | oxo– (chain), formyl– (ring) | oxo– | hydroxy– | amino– |

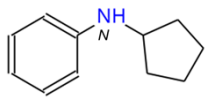
* If the parent is a chain, this C is included in the name, e.g. $CH_3CH_2CO_2H$ is propanoic acid. If the parent is a ring, this C is external to the ring and is not included in the parent name, e.g. $C_5H_9CO_2H$ is cyclopentanoic acid.

2-Phenyloxirane 

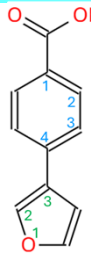
3-(Aziridin-2-yl)piperidine 

3-(Naphthalen-2-yl)furan 

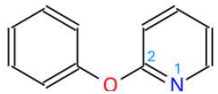
*N*-Methylformamide 

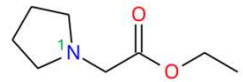
*N*-Cyclopentylaniline 

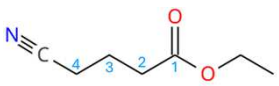
4-Acetylbenzaldehyde 

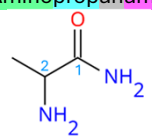
4-(Furan-3-yl)-benzoic acid 

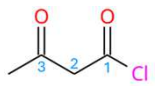
3-Bromo-4-nitrophenol 

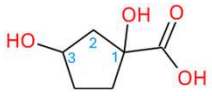
2-Phenoxypyridine 

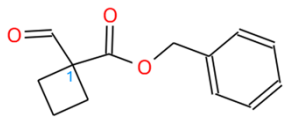
Ethyl (pyrrolidin-1-yl)acetate 

Ethyl 4-cyanobutanoate 

2-Aminopropanamide 

3-Oxobutanoyl chloride 

1,3-Dihydroxycyclopentanecarboxylic acid 

Benzyl 1-formylcyclobutanoate 

5-Hydroxyhexan-2,4-dione 