

---

# 1

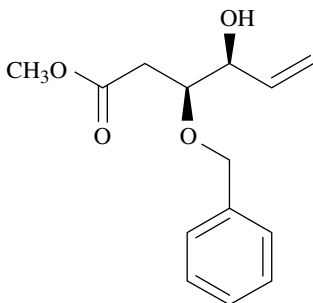
---

## READING NOMENCLATURE

Organic chemistry is understood in terms of molecular structures as represented pictorially. Cataloging, writing, and speaking about these structures require a nomenclature system, the basics of which you have studied in your introductory course. To go further with the subject, you must begin reading journals, and this requires understanding of the nomenclature of complex molecules. This chapter presents a selection of compounds to illustrate the translation of names to structural representations. The more difficult task of naming complex structures is not covered here because each person's needs will be specialized and can be found in nomenclature guides [1–5]. Most of the nomenclature rules are used to eliminate alternative names and arrive at a unique (or nearly so) name for a particular structure; thus, when beginning with names, you will need to know only a small selection of the rules in order to simply read the names and provide a structure. Although the subject of nomenclature is vast, these selections will enable you to understand many names in current journals.

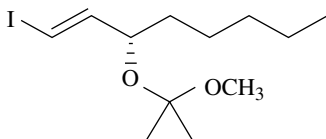
## 1.1 ACYCLIC POLYFUNCTIONAL MOLECULES

Methyl (3*S*,4*S*)-4-hydroxy-3-(phenylmethoxy)hex-5-enoate



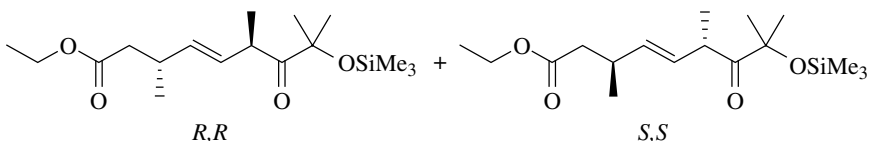
The space after methyl and the “ate” ending tells you this is a methyl ester. The acid from which the ester derives is a six-carbon chain with a double bond between carbons 5 and 6. There is an alcohol function on carbon 4. There is a methoxy group on carbon 3 and a phenyl group on the carbon of the methoxy group. Carbons 3 and 4 are stereogenic atoms each with *S* configuration as designated.

3-(*S*)-*trans*-1-Iodo-1-octen-3-yl methoxyisopropyl ether



This is an example of a derivative name, that is, the first word is the complete name of an alcohol and the other two words describe a derivatization where the alcohol is converted to an ether (ketal). Such a name would be useful in discussing a compound that has the ketal present as a temporary entity, for example, as a protecting group.

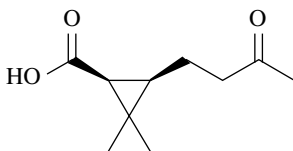
Ethyl (*E*,3*R*\*,6*R*\*)-3,6,8-trimethyl-8-[(trimethylsilyl)oxy]-7-oxo-4-nonenoate



This is the ethyl ester of a nine-carbon unsaturated acid with substituents. The *oxo* indicates that there is a keto function on carbon 7. Be careful to distinguish this from the prefix *oxa-*, which has a different meaning; see Section 1.6. The asterisks indicate that the configuration designation is not absolute but rather represents that stereoisomer and/or the enantiomer thereof. Thus this name represents the *R,R* and/or the *S,S* isomers, but not *R,S* or *S,R*. This designation excludes diastereomers and is a common way to indicate a racemate.

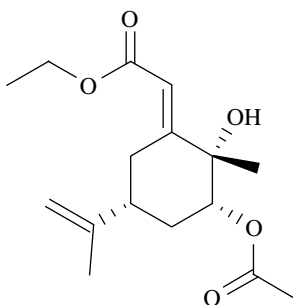
## 1.2 MONOCYCLIC ALIPHATIC COMPOUNDS

(1*S*,3*R*)-2,2-Dimethyl-3-(3-oxobutyl)cyclopropanecarboxylic acid



The ring is placed in the plane of the paper. Numbering of the ring starts at the location of the highest priority substitution, the carboxylic acid in this case. The butyl substituent on the third carbon of the ring has a keto function on the third carbon of the butyl chain.

[2*S*-(1*E*,2 $\alpha$ ,3 $\alpha$ ,5 $\alpha$ )]-[3-(Acetyloxy)-2-hydroxy-2-methyl-5-(methylethenyl)cyclohexylidene]acetic acid ethyl ester

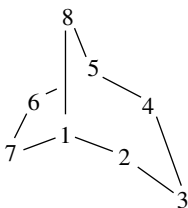


The *ylidene* indicates that the cyclohexyl is attached to the acetic acid by a double bond and the ethyl ester is indicated at the end for simplicity. The double-bonded ring atom is carbon 1 and the substituents on the ring are placed on the ring according to their locant numbers. The *E* indicates

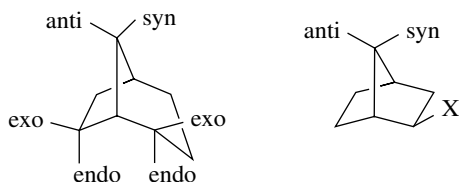
the geometry of the double bond. All the  $\alpha$  substituents reside on one face of the ring, cis to each other. Any  $\beta$  substituents would reside on the opposite face of the ring, trans to the  $\alpha$  substituents. Where two substituents are on the same ring atom, as on carbon 2 in this case, the Greek letter indicates the position of the higher-priority substituent. Here the hydroxy, acetyloxy, and methylethenyl are all cis to each other on the ring.

### 1.3 BRIDGED POLYCYCLIC STRUCTURES

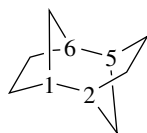
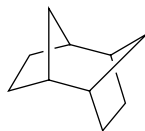
The nomenclature of bridged polycyclic systems requires additional specifications. A bicyclic system would require two bond breakings to open all the rings, a tricyclic system, three, and so on. Rather than viewing this as rings, certain carbons are designated as bridgeheads from which the bridges branch and recombine. In the system below, the first bridgehead is designated as carbon 1 and the system is numbered around the largest bridge to the second bridgehead, carbon 5. Numbering continues around the medium bridge, then the smallest bridge, as shown. The compound is named bicyclo[3.2.1]octane.



All bicyclo compounds require three numbers in brackets, tricyclo require four, and so on, and these numbers indicate the number of carbons in the bridges and are used to locate substituents, heteroatoms, and unsaturation. The name of the parent alkane includes the total number of atoms in the bridges and bridgeheads (excluding substituents) and is given after the brackets. The use of prefixes *exo*, *endo*, *syn*, and *anti* to indicate stereochemical choices is demonstrated generally as shown below.

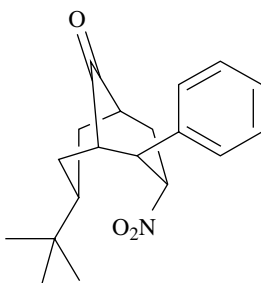


In tricyclic compounds, the relative stereochemistry among the four bridgeheads requires designation. Look at the largest possible ring in the molecule and consider the two faces of it. If there are no higher-priority substituents on the primary bridgehead atoms, the smallest bridge (but not a zero bridge) defines the  $\alpha$  face. If the smallest bridge (not zero) at the secondary bridgeheads is on the same face of the large ring as the  $\alpha$  defining one, it is also designated as  $\alpha$ ; that is, the two are *cis* to each other.

 $1\alpha, 2\beta, 5\beta, 6\alpha$  $1\alpha, 2\alpha, 5\alpha, 6\alpha$ 

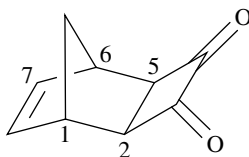
If they are *trans*, there will be two  $\alpha$ s and two  $\beta$ s as illustrated. If there is a zero bridge, the position of the bridgehead hydrogens is indicated with Greek letters.

[1*S*-(2-*exo*,3-*endo*,7-*exo*)]-7-(1,1-Dimethylethyl)-3-nitro-2-phenylbicyclo[3.3.1]nonan-9-one



This bicyclo system has bridges with three, three, and one carbons each, indicated by the bracketed numbers separated by periods. Carbon 2 carries a phenyl that projects toward the smaller neighboring one-carbon bridge rather than the larger three-carbon bridge, as indicated by 2-*exo*. The 1,1-dimethylethyl group is also *exo*. This group is commonly called *tert-butyl*, but this is a *Chemical Abstracts* name built on linear groups. The prefixes *exo* and *endo* indicate the stereochemistry.

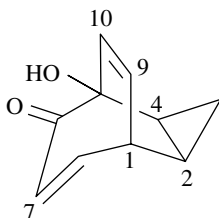
( $1\alpha, 2\beta, 5\beta, 6\alpha$ )-Tricyclo[4.2.1.0<sup>2,5</sup>]non-7-ene-3,4-dione



Starting with a pair of bridgeheads, draw the four-, two-, and one-carbon bridges. The zero bridge then connects carbons 2 and 5 as indicated by the superscripts, thus making them bridgeheads also. At bridgeheads 1 and 6, the smallest bridge is considered a substituent and given the  $\alpha$  designation at both ends. At bridgeheads 2 and 5, the  $\beta$ s indicate that the hydrogens are trans to the  $\alpha$  bridge.

Sometimes a bridgehead substituent will have a higher priority than the smallest bridge thereon. The designation for that bridgehead will indicate the position,  $\alpha$  or  $\beta$ , of that higher-priority substituent rather than the bridge as illustrated in the next example.

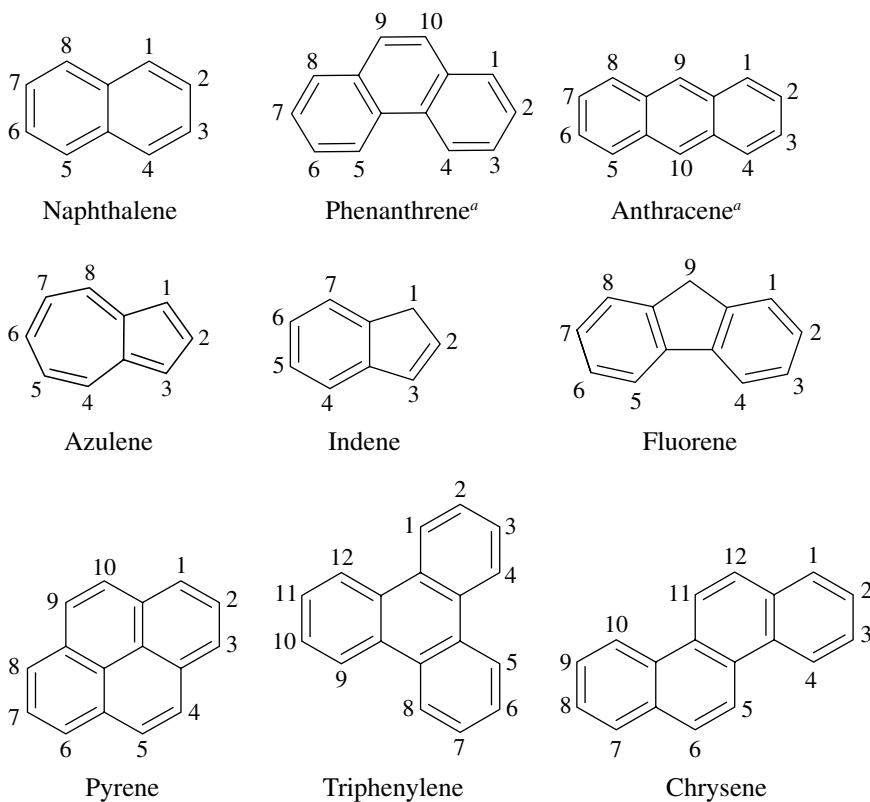
(1 $\alpha$ ,2 $\beta$ ,4 $\beta$ ,5 $\beta$ )-5-Hydroxytricyclo[3.3.2.0<sup>2,4</sup>]deca-7,9-dien-6-one



At bridgehead 1, the smallest bridge, carbons 9 and 10, is considered a substituent on the largest ring and designated  $\alpha$ . The hydrogens at carbons 2 and 4 are trans to it and marked  $\beta$ . The OH group on carbon 5 is a higher-priority substituent than the C-9 to C-10 bridge and is trans to the bridge; thus it is labeled  $\beta$ .

## 1.4 FUSED POLYCYCLIC COMPOUNDS

Fused-ring compounds have a pair or pairs of adjacent carbon atoms common to two rings. Over 35 carbocyclic examples have trivial names, some of which need to be memorized as building blocks for names of more complex examples. The names end with *ene*, indicating a maximum

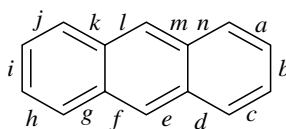
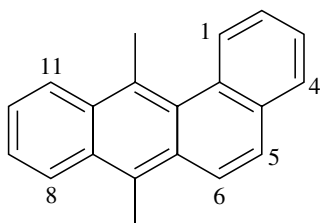
**TABLE 1.1 Trivial Names of Some Fused Polycyclic Hydrocarbons**

<sup>a</sup>Exceptions to systematic numbering.

number of alternating double bonds. A selection is illustrated in Table 1.1, showing one resonance form for each. Others can be found online [6].

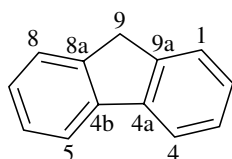
Fusing more rings onto one of these basic systems may give another one with a trivial name. If not, a name including the two rings or ring systems with bracketed locants is used, as in the following example.

#### 7,12-Dimethylbenz[*a*]anthracene

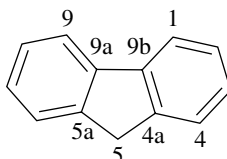


Since a side of the anthracene is shared, the sides are labeled *a*, *b*, *c*, and so on, where carbons numbered 1 and 2 constitute side *a* and 2 and 3 constitute side *b*, continuing in order for all sides. The earliest letter of the anthracene is used to indicate the side fused, and the ring fused to it appears first. The “o” ending of benzo is deleted here because it would be followed by a vowel.

The final combination is then renumbered to locate substituents, or sites of reduced unsaturation. To renumber, first orient the system so that a maximum number of fused rings are in a horizontal row. If there are two or more choices, place a maximum number of rings to the upper right. Then number clockwise starting with the carbon not involved in fusion in the most counterclockwise position in the uppermost–farthest right ring. See the numbering in the systems with trivial names above. Atoms at the fusion sites, which could carry a substituent only if they were not  $\pi$ -bonded, are given the number of the previous position with an *a*, *b*, *c*, and so on. Where there is a choice, the numbers of the fusion carbons are minimized too; for example, 4*b* < 5*a*:

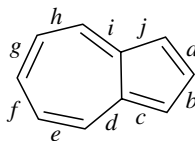
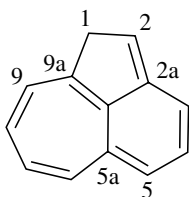


Correct



Incorrect

### 1*H*-Benz[*cd*]azulene



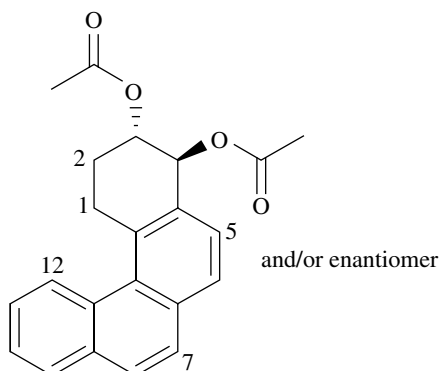
Azulene

First letter the sides of azulene. The benzo ring is fused to both the *c* and *d* faces as indicated in the brackets. Now reorient the system for numbering. The choice of which two rings go on the horizontal axis and



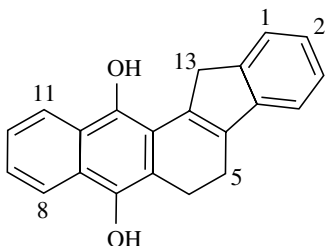
which one is in the upper position is determined by which orientation gives the smallest number to the first fusion atom—in this case 2a instead of 3a, or 4a. In molecules where one carbon remains without a double-bonding partner, it is denoted by *H*. This is called *indicated hydrogen* and is used even when an atom other than hydrogen is actually on that carbon in the molecule of interest.

*trans*-1,2,3,4-Tetrahydrobenzo[*c*]phenanthrene-3,4-diol diacetate



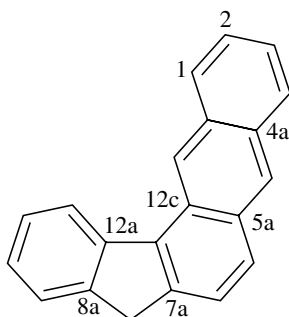
The sides of phenanthrene are lettered; the carbons numbered 3 and 4 are the *c* side. As in several systems, phenanthrene is numbered in an exceptional way. A benzene ring is fused to the *c* side, and a new systematic numbering is made. Carbons 1–4 have hydrogens added to saturate two double bonds, and then carbons 3 and 4 have hydroxy groups substituted for hydrogens in a *trans* arrangement. Finally, the compound is named as the acetate ester at both alcohol sites.

6,13-Dihydro-5*H*-indeno[2,1-*a*]anthracene-7,12-diol



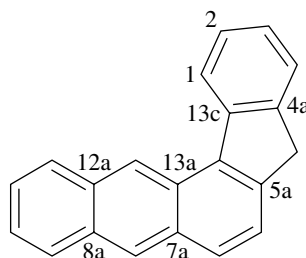
The fusion of this ring system is given in the brackets. The bracketed *a* precedes anthracene indicating that the sides of anthracene will be lettered, and the bracketed 2,1 follows the indene, thus the indene numbering indicated in Table 1.1 will be used to give the point of fusion. The order of the numbers indicates the direction of the fusion, thus the carbons of indene are fused in order 2, then 1 to the *a* side of anthracene, with the number 1 carbon of anthracene constituting the number 2 carbon of indene, and the number 2 carbon of anthracene, the number 1 carbon of indene. The united system is then renumbered according to the rules and the substituents, added hydrogen, and indicated hydrogen placed accordingly. The indicated hydrogen is assigned the lowest-numbered atom not involved in double bonding.

The other direction of fusion is in *1H*-indeno[1,2-*a*]anthracene:



Correct:

4a, 5a, 7a, 8a, 12a, 12b, 12c, 13a



Incorrect:

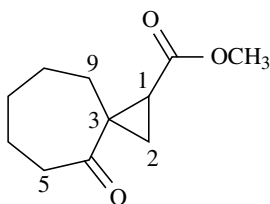
4a, 5a, 7a, 8a, 12a, 13a, 13b, 13c

Note that here a different orientation is used because it gives the lower fusion numbers.

## 1.5 SPIRO COMPOUNDS

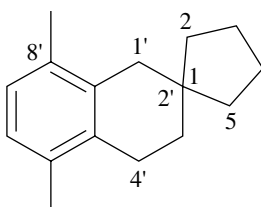
In spiro compounds, a single atom is common to two rings. There are two kinds of nomenclature for these. Where there are no fused rings present, the carbons of both rings are counted in one series as in the bicyclic nomenclature, and the hydrocarbon name includes the carbons of both rings as in the following example.

## 4-Oxospiro[2.6]nonane-1-carboxylic acid methyl ester



Of the nine carbons, one is the spiro atom, two join round to make a three-membered ring, and six finish a seven-membered ring, as indicated by the bracketed numbers. The locant numbers begin in the smaller ring at an atom adjacent to the spiro one, continue around the smaller ring including the spiro atom, and proceed around the larger ring.

When ring fusion is also present, the two ring systems that share the spiro atom are given in brackets with splicing locants as shown in the following example.

3',4'-Dihydro-5',8'-dimethylspiro[cyclopentane-1,2'(1'*H*)-naphthalene]

Two separate numbering systems are used. The unprimed number belongs to the ring nearest it in the brackets, the cyclopentane, while the primed numbers belong to their nearer neighbor in the brackets, naphthalene. The locant numbering in the fused system follows the usual pattern (Section 1.4) and is identified with primes. The -1,2' indicates that the shared atom is number 1 of the cyclopentane ring and number 2' of the naphthalene. The spiro linkage requires another naphthalene ring atom to be excluded from double bonding, in this case, the 1' as determined by the indicated hydrogen, 1'*H*.

## 1.6 MONOCYCLIC HETEROCYCLIC COMPOUNDS

Systematic and trivial names are both commonly in use for heterocyclic compounds. The systematic names consist of one or more prefixes from Table 1.2 with multipliers where needed designating the heteroatoms, followed by a suffix from Table 1.3 to give the ring size with an indication of the unsaturation. This is preceded by substituents. Thus, oxepin is a

**TABLE 1.2 Prefixes in Order of Decreasing Priority<sup>a,b</sup>**

Oxygen	ox-
Sulfur	thi-
Selenium	selen-
Nitrogen	az-
Phosphorus	phosph- (or phosphor- before -in or -ine)
Silicon	sil-
Boron	bor-

<sup>a</sup>From Ref. [7].

<sup>b</sup>An “a” is added after each prefix if followed by a consonant.

**TABLE 1.3 Suffixes Indicating Ring Size<sup>a</sup>**

Atoms in the Ring	Containing Nitrogen		Containing no Nitrogen	
	Maximally Unsaturated	Saturated	Maximally Unsaturated	Saturated
3	-irine	-iridine	-irine	-irane
4	-ete	-etidine	-ete	-etane
5	-ole	-olidine	-ole	-olane
6	-ine	— <sup>b</sup>	-in	-ane
7	-epine	— <sup>b</sup>	-epin	-epane
8	-ocine	— <sup>b</sup>	-ocin	-ocane
9	-onine	— <sup>b</sup>	-onin	-onane
10	-ecin	— <sup>b</sup>	-ecin	-ecane
>10 <sup>c</sup>	—	—	—	—

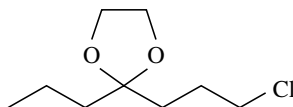
<sup>a</sup>From Ref. [7].

<sup>b</sup>Use the unsaturated name preceded by “perhydro.”

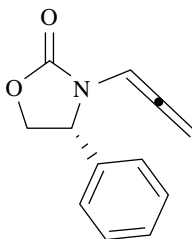
<sup>c</sup>Use the carbocyclic ring name with heteroatom replacement prefixes: oxa-, thia-, and so forth.

seven-membered ring including one oxygen and three double bonds. The ring size is explicit in some of the suffixes as *ep*, *oc*, *on*, and *ec*, which are derived from *heptane*, *octane*, *nonane*, and *decane*, respectively. Numbering begins with the element highest in Table 1.2 and continues in the direction that gives the lowest locant to the next heteroatom.

#### 2-Propyl-2-(3-chloropropyl)-1,3-dioxolane

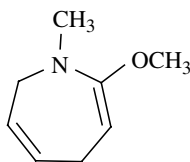


#### (*R*)-4-Phenyl-3-(1,2-propadienyl)oxazolidin-2-one



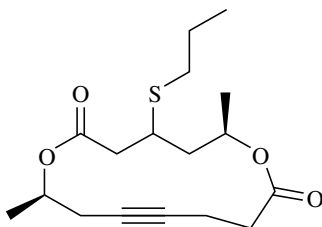
The numbering begins with oxygen for priority. The 1,2-propadienyl group is thus located at position 3, which is the nitrogen. The *-olidin-* ending specifies a saturated five-membered ring, with *-one* indicating the presence of the ketone.

#### 4,7-Dihydro-2-methoxy-1-methyl-1*H*-azepine



Azepine indicates an unsaturated seven-membered ring containing one nitrogen. Because of the odd number of carbons, one of the seven must have indicated *H*. The dihydro indicates additional hydrogens on two other carbons; therefore, there is one  $\pi$  bond less than maximally unsaturated. Notice that the indicated hydrogen is assigned the lowest-numbered atom not in double bonding (the nitrogen) and *then* replaced by the substituent.

(6*R*,14*R*)-6,14-Dimethyl-1,7-dioxa-4-(1-propylthio)cyclotetradec-11-yne-2,8-dione



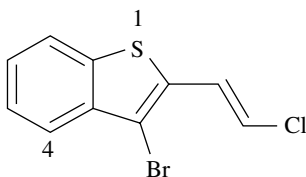
The ring is larger than 10 members; therefore, the hydrocarbon ring name cyclotetradecyne was used, modified by *1,7-dioxa*, which is a replacement of carbons 1 and 7 with oxygens. The *a* ending on *oxa* indicates replacement. The numbering begins at a heteroatom and proceeds to the other heteroatom by the shortest path. The stereochemistry at position 4 is unspecified.

Many five- and six-membered rings and fused ring systems have trivial names that are preferred over the systematic names. Table 1.4 provides a selection of the more common ones. Additional names can be found online [8].

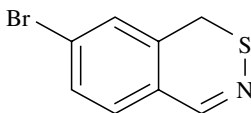
## 1.7 FUSED-RING HETEROCYCLIC COMPOUNDS

The names of the heterocycles in the previous section along with the rules for fusion in Section 1.4 are the basis for the following names.

3-Bromo-2-(2-chloroethenyl)benzo[*b*]thiophene



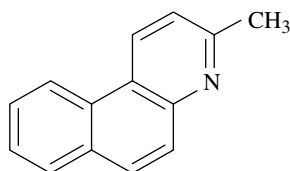
The final numbering begins at the most counterclockwise atom not involved in fusion, arranged to give the heteroatom the smallest possible number, in this case 1.

7-Bromo-1*H*-2,3-benzothiazine

Sulfur is higher-priority than nitrogen and the lowest number for it is 2. The 3 locates the nitrogen. In this case, no bracketed site of fusion is specified because the fusion must precede the atom numbered 1. This is usual when there is more than one heteroatom and the fusion is simply benzo. The presence of one divalent atom in a six-membered ring excludes another atom from double bonding, thus the indicated hydrogen.

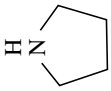
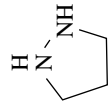
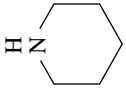
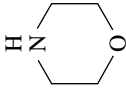

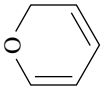
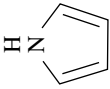



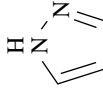
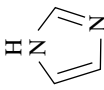
In choosing where to start numbering, and in which direction to proceed, a hierarchy of rules must be followed. Numbering always begins at a nonfused atom adjacent to a fused atom, but since there are several possible orientations for the molecule, a choice is made as follows:

1. Give the lowest possible number to the first heteroatom regardless of the priority of the atom.
2. If this allows two choices, choose the one that gives the second heteroatom the lowest number. If there are still two choices, minimize the number of the third, and so on.
3. If there are still two choices, give the lower number to the higher-priority heteroatom.
4. If all the above allow two choices, give the lowest number to the first fusion atom.
5. Finally, if there are still two choices, give the lowest numbers to the substituents.

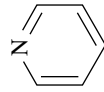
1,2-Dihydro-3-methylbenzo[*f*]quinoline

The sides of quinoline are lettered following the numbering system and the benzo is fused to side *f*. The whole system is renumbered orienting as directed

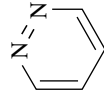
**TABLE 1.4 Trivial Names of Some Heterocycles**

Saturated			
			
Pyrrolidine	Pyrazolidine	Piperidine	Morpholine
Unsaturated			
			
Furan	2 <i>H</i> -pyran	Pyrrole	Thiophene
			
Oxazole	Isoxazole	Pyrazole	Imidazole

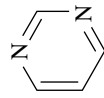




Pyridine



Pyridazine



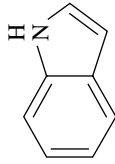
Pyrimidine



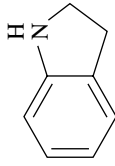
Pyrazine

---

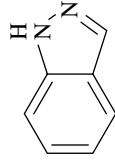
Fused



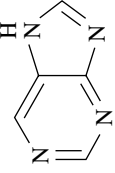
Indole



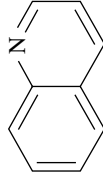
Indoline



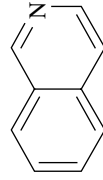
Indazole



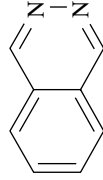
Purine



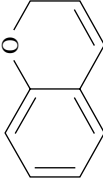
Quinoline



Isoquinoline



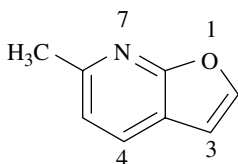
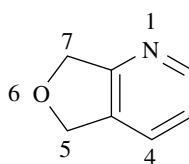
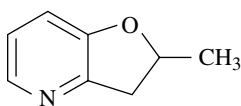
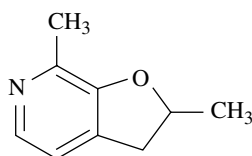
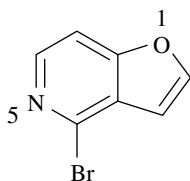
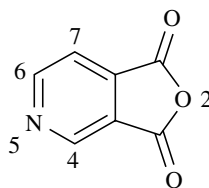
Phthalazine



Chromene

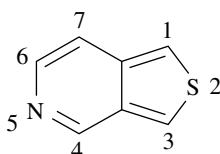
in Section 1.4, that is, maximum number of rings in a horizontal row, maximum in upper right, giving the heteroatom the lowest possible number, and numbering from the most counterclockwise nonfused atom in the upper right ring.

When two heterocyclic rings are fused, sides and direction of fusion are indicated in brackets as in the carbocyclic cases in Section 1.4. Examples of the six possible fusions between pyridine and furan are shown.

6-Methylfuro[2,3-*b*]pyridine5,7-Dihydrofuro[3,4-*b*]pyridine2,3-Dihydro-2-methylfuro[3,2-*b*]pyridine2,3-Dihydro-2,7-dimethylfuro[2,3-*c*]pyridine4-Bromofuro[3,2-*c*]pyridineFuro[3,4-*c*]pyridine-1,3-dione

Furo[3,2-*b*] indicates that carbons 3 and 2 of the furan are the 2 and 3 carbons of the pyridine, respectively. The two-ring system is renumbered following the hierarchy of rules given earlier.

Thieno[3,4-*c*]pyridine



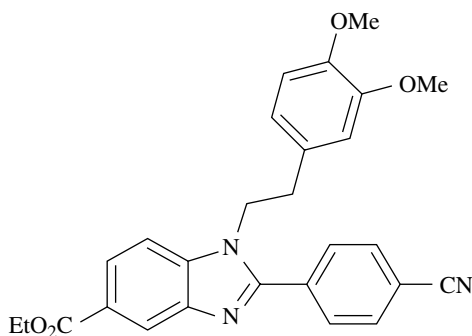
Thiophene as the first ring named in a fused system is shortened to thieno.

### 2-Methyl-2*H*-thiazolo[4,5-*e*]-1,2-oxazine



The name *thiazole*, like *oxazole*, means 1,3-thiazole. The atoms of thiazole are numbered and the sides of 1,2-oxazine are lettered. The fusion is drawn with atoms 4 and 5 of thiazole as atoms 5 and 6 of the oxazine. The system is then renumbered using the hierarchy of rules. You should flip your initial drawing about both the *x* and *y* axes to consider all four orientations to find the correct numbering.

### 2-(4-Cyanophenyl)-1-[2-(3,4-dimethoxyphenyl)-ethyl]-1*H*-benz[*d*]imidazole-5-carboxylic acid ethyl ester

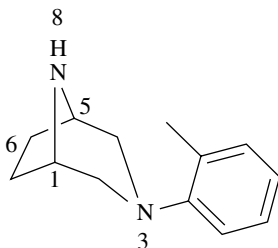


The fusion in this ring places the benzene on the *d* side of the imidazole ring. Renumbering the new ring gives the locants as shown.

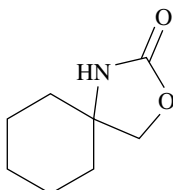
## 1.8 BRIDGED AND SPIRO HETEROCYCLIC COMPOUNDS

Bridged and spiro heterocyclic compounds are named using the replacement nomenclature; that is, the hydrocarbon name is used with oxa, aza, and so on to substitute heteroatoms for carbons as was seen in large ring monocyclic compounds in Section 1.6.

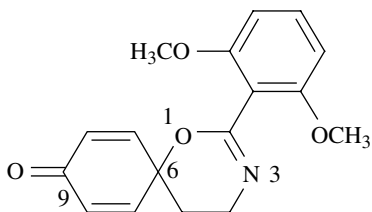
## 3-(2-Methylphenyl)-3,8-diazabicyclo[3.2.1]octane



## 3-Oxa-1-azaspiro[4.5]decan-2-one



## 2-(2',6'-Dimethoxyphenyl)-1,3-oxazaspiro[5.5]undeca-2,7,10-trien-9-one



Many of the names that you will see in journals should be understandable by analogy from the examples studied here, but certainly there are many more complicated cases beyond the scope of this chapter. The chapter references should be consulted for them.

**RESOURCES**

1. Many of the references in this chapter and the next can be helpful in finding the names of organic compounds.

2. The web site maintained by ACD Labs with the contents of the IUPAC Blue Book, <http://www.acdlabs.com/iupac/nomenclature/>, is very helpful.
3. The web site maintained by Chemical Abstracts Service, <http://www.commonchemistry.org/>, can identify synonyms from CAS Registry numbers and names.

## PROBLEMS

Draw complete structures for each of the following compounds.

- 1.1 2,4-Dimethylbenzo[*g*]quinoline
- 1.2 2-(Bromomethyl)-4,7-dimethoxyfuro[2,3-*d*]pyridazine
- 1.3 Spiro[cyclopentane-1,3'-bicyclo[4.1.0]heptane]
- 1.4 (1*R*,3*R*,5*S*)-*endo*-1,3-Dimethyl-2,9-dioxabicyclo[3.3.1]nonane
- 1.5 Spiro[5.7]trideca-1,4-dien-3-one
- 1.6 1-Benzoyl-2-phenylaziridine
- 1.7 3-Butyryl-2-(3-chloropropyl)-1-(methoxycarbonyl)-1,2-dihydropyridine
- 1.8 7-Methyl-7*H*-benzo[*c*]fluorene-7-carboxylic acid
- 1.9 (1*R*,6*S*,7*S*)-4-(*t*-Butyldimethylsiloxy)-6-(trimethylsilyl)bicyclo[5.4.0]undec-4-en-2-one
- 1.10 (1 $\alpha$ ,2 $\beta$ ,5 $\beta$ ,6 $\alpha$ )-Tricyclo[4.3.1.1<sup>2,5</sup>]undecane-11-one
- 1.11 (1 $\alpha$ ,2 $\beta$ ,3 $\alpha$ ,5 $\alpha$ )-6,6-Dimethylbicyclo[3.1.1]heptane-2,3-diol
- 1.12 [1*R*-(1 $\alpha$ ,2 $\beta$ ,6 $\alpha$ )]-4,7,7-Trimethylbicyclo[4.1.0]hept-3-en-2-ol
- 1.13 6-(Benzyloxycarbonyl)-3-cyano-4-chloro-6-azabicyclo[3.2.1]oct-3-ene
- 1.14 Thieno[3,4-*b*]pyridine
- 1.15 2*H*-3,1-Benzothiazine
- 1.16 1-Phenyl-1,4,5,6,7,8,9,10-octahydrocyclonona[*d*][1,2,3]triazole

- 1.17** 5'-Acetyl-4'-amino-1,3-dihydro-6'-methyl-1,3-dioxospiro[2H-indene-2,2'-2*H*-pyran]
- 1.18** 5-Methylbenzo[*b*]chrysene
- 1.19** Methyl 7-methylbicyclo[4.2.0]octa-1,3,5-triene-7-carboxylate
- 1.20** (*exo,syn*)-2-(1-Pyrrolidinyl)bicyclo[3.3.1]nonan-9-ol
- 1.21** 3-Amino-5,6,8,9-tetrahydro-7*H*-pyrazino[2,3-*d*]azepine-2,7-dicarboxylic acid diethyl ester
- 1.22** 2-Benzoyl-1,6,7,11*b*-tetrahydro-2*H*-pyrazino[2,1-*a*]isoquinoline-3,4-dione
- 1.23** (±)-Dimethyl 2-((5-hydroxy-4-oxo-3,5-diphenylcyclopent-2-en-1-yl)methyl) malonate
- 1.24** 8-bromo-7-chloro-2-(ethylthio)-4*H*,5*H*-pyrano[3,4-*e*]-1,3-oxazine-4,5-dione
- 1.25** [1*R*-(1α,2β,4α)]-4-chloro-2-methylcyclohexanecarboxylic acid
- 1.26** 1',2'-dihydro-4-methyl-2'-oxospiro[4-cyclohexene-1,3'-[3*H*]indole]-2,2-dicarboxylic acid diethyl ester
- 1.27** 2,4,6-Trichlorophenyl 3,4-dihydronaphthalene-2-carboxylate
- 1.28** 7-Chloro-2-methyl-1,4-dihydro-2*H*-isoquinolin-3-one
- 1.29** 4-(Methylthio)-2-phenyl-3-quinazoline
- 1.30** (2*S*,3*S*)-3-Acetyl-8-carboethoxy-2,3-dimethyl-1-oxa-8-azaspiro[4.5]decane

## REFERENCES

1. Panico, R.; Powell, W. H.; Richer, J.-C. (preparers) *A Guide to IUPAC Nomenclature of Organic Compounds (Recommendations 1993)*; Blackwell Scientific Publications: Oxford, UK, 1993.
2. Rigaudy, J.; Klesney, S. P. (preparers) *Nomenclature of Organic Chemistry*; International Union of Pure and Applied Chemistry/Pergamon Press: Oxford, 1979.
3. ACD Labs. IUPAC Nomenclature of Organic Chemistry Home Page. <http://www.acdlabs.com/iupac/nomenclature/>. (This website includes the contents of references 2 and 3 in html format.) (accessed February 17, 2015).

4. American Chemical Society. *Naming and Indexing of Chemical Substances for Chemical Abstracts*; American Chemical Society: Columbus, OH, 2008. [www.cas.org/File%20Library/Training/STN/User%20Docs/indexguideapp.pdf](http://www.cas.org/File%20Library/Training/STN/User%20Docs/indexguideapp.pdf) (accessed February 17, 2015).
5. Gupta, R. R.; Kumar, M.; Gupta, S. S. *Heterocyclic Chemistry*, Volume I; Springer: Berlin, 1998, pp. 3–38.
6. Advanced Chemistry Development Labs. Fused Polycyclic Hydrocarbons. [http://www.acdlabs.com/iupac/nomenclature/79/r79\\_63.htm](http://www.acdlabs.com/iupac/nomenclature/79/r79_63.htm) (accessed February 17, 2015).
7. Advanced Chemistry Development Labs. Heterocyclic Systems. Rule B-1. Extension of the Hantzsch-Widman System. [http://www.acdlabs.com/iupac/nomenclature/79/r79\\_702.htm](http://www.acdlabs.com/iupac/nomenclature/79/r79_702.htm) (accessed February 17, 2015).
8. Advanced Chemistry Development Labs. Heterocyclic Systems. Rule B-2. Trivial and Semi Trivial Names. [http://www.acdlabs.com/iupac/nomenclature/79/r79\\_951.htm](http://www.acdlabs.com/iupac/nomenclature/79/r79_951.htm) (accessed February 17, 2015).





---

# 2

---

## ACCESSING CHEMICAL INFORMATION

A truly vast and ever-growing body of organic chemical information is recorded in the chemical literature and in online databases. In a research library practically all of it can be searched quickly through these databases and with the use of *Chemical Abstracts*, *Beilsteins Handbook of Organic Chemistry*, and the review literature. Because print resources are still available, those will be discussed briefly, but online sources will be the focus of this chapter.

### 2.1 DATABASES

Databases gather information about compounds that can include physical constants, spectral data, and toxicity and hazard information. Much of this sort of information can be accessed through a number of familiar publications such as the CRC handbook [1] or the Merck Index [2]. The Merck Index is also available online, but full access requires a subscription.

Online databases are able to compile a great deal of information in a format that is readily searchable. PubChem [3] is maintained by the National

Center for Biotechnology Information (NCBI) and provides information on the biological activity of small molecules in three databases: Substances, Compounds, and BioAssays. As of this writing, there are about 120 million substances and 48 million compounds indexed at the site, as well as over 710 million bioassays for the compounds in the compound database. “Compounds” are identified by name or structure, while “substances” may include mixtures of compounds, such as plant extracts. It is generally more useful to search the compound index by name or structure.

ChemSpider [4] is maintained by the Royal Society of Chemistry and allows text and structure searching for over 29 million compounds. Spectra, properties, and a wealth of other information are available.

Common Chemistry [5] is provided by Chemical Abstracts Service (CAS) that links chemical names and CAS registry numbers for about 7900 chemicals. This free database is particularly useful when the name of a chemical is confusing or uncertain as it will allow searching with the CAS number. It contains links to compound entries in Wikipedia and through that to entries for compounds at ChemSpider and PubChem.

The Aldrich catalog is available online [6] and provides CAS registry numbers, Beilstein and PubChem ID numbers, in addition to physical constants and Safety Data Sheets (SDS). The catalog allows searching by name and structure.

The National Institutes of Science and Technology (NIST) maintains an extensive database of chemical information [7]. Included are thermochemical data, spectral data, and kinetic data, which are searchable by name, structure, molecular weight, and other options that are compound specific.

## 2.2 CHEMICAL LITERATURE

The chemical literature can be organized into several categories: primary literature, review literature, and abstract literature. Abstract journals such as Chemical Abstracts, as well as web sites such as the one maintained by the American Chemical Society for their journals [8] allow ready searching of the chemical literature.

The primary literature includes journals that publish original research findings. Lavoisier’s *Annales de Chimie* began publication in 1789 and is

one of the oldest chemical journals. Examples of current publications include the *Journal of the American Chemical Society*, the *Journal of Organic Chemistry*, and *Tetrahedron* to name a very few. Most journals require a subscription, and a few offer open access to their materials. Most chemical societies publish a range of journals specific to individual subfields of chemistry, as well as journals covering a broad range of chemical topics.

The review literature includes journals that publish summaries of the current knowledge on a particular topic. Examples include *Accounts of Chemical Research*, *Chemical Reviews*, and *Chemical Society Reviews*. These papers can be particularly helpful as you begin a project as they usually provide substantial lists of references into the primary literature.

An abstract is a summary of the important discoveries and conclusions in a paper and provides a way into the content of research publications. Abstract literature allows access to this information in a searchable format. Many of the early journals published annual author and subject indexes. *Chemisches Zentralblatt*, begun in 1830 and ending in 1969, is the one of the most reliable means into the very early chemical literature. A searchable database is available by subscription through InfoChem [9], which maintains a database for current literature as well.

If you are interested in the preparation and/or properties of a particular organic compound, this information may be found quickly in *Beilstein's Handbuch der Organischen Chemie*. *Beilstein* is published in German and is an organized collection of preparations and properties of organic compounds that were known before 1960. The fourth edition (*Vierte Auflage*) consists of a basic series (*Hauptwerk*) covering work up to 1909, and four supplementary series (*Erganzungswerk*) covering the literature to 1959. A fifth supplement was published in English from 1960 to 1979. In the print edition, compounds are arranged in the volumes according to the rules of the *Beilstein* system, which allows you to search directly in the volumes without using indexes, finding similar compounds located together. These rules are beyond the scope of this chapter but are available elsewhere [10]. *Beilstein's* originally contained a comprehensive survey of the literature. As of 2010, approximately 400 journals are indexed here and only selected content is included. Additional materials were added in 2013, but *Beilstein's* is not a comprehensive means for literature searches. Current access to this database is available by subscription to Reaxys [11].

The print edition of *Chemical Abstracts* appeared in weekly issues. Six months of these issues constituted a volume that was accompanied by six indexes: *General Subject Index*, *Chemical Substance Index*, *Formula Index*, *Index of Ring Systems*, *Author Index*, and *Patent Index* [12]. January to June of 1993 was volume 118. Ten volumes (5 years' coverage) were combined in the *Collective Indexes*; the 11th collective volumes covered the years from 1982 to 1986 and comprised 93 bound books occupying 17 feet of shelf space. It is easy to see why the print version is no longer published.

*Chemical Abstracts* (CA) currently covers over 10,000 periodicals, patents, and other sources and produces brief summaries of the information in each journal article along with a bibliographic heading. The print publication ceased in 2010, and information in this database is accessed online through STN (mostly librarians) or SciFinder (mostly professional chemists). *Chemical Abstracts* is selectively incorporating older abstracts into its database, in particular materials from *Chemisches Zentralblatt* and the archives of the Royal Chemical Society's abstracts.

SciFinder [13] is a powerful search engine and requires a subscription, although a limited number of free searches are available as a benefit of membership in the American Chemical Society. Searches are conducted under three search headings: References, Substances, and Reactions. A series of excellent tutorials and videos are available for training purposes [14].

Under References, searches can be conducted by author, company, or journal name, or by document identifier, patent number, or research topic. An initially broad research topic search can be narrowed using the search terms that are generated. Similarly, advanced search options can be used to narrow the results as appropriate to the application.

Under Substances, search options include chemical structures, molecular formulas, and properties. Chemical structures may be searched as exact matches, substructures, and similarity matches. Exact matches will include enantiomers, stereoisomers, radicals, salts, and the like. These results can be filtered once the search is complete. Substructure matches allow substitution at any positions not specifically blocked. A similarity search is the broadest category. If an "exact" search doesn't result in sufficient matches, a substructure search may yield results. Markush searching is also available. Markush structures are found in the patent

literature and allow chemically similar structures to be searched at the same time, using -R to avoid specification of individual groups. Once structures are found, each listing provides links to references, reactions, commercial sources, regulatory information, spectra, and experimental properties for the compound.

Under Reactions, a search of structures of either reactants or products, or both, is permitted, and advanced search options allow specification of solvents, other functionality and reaction details, and dates. Results of a reaction search can be further narrowed by yield or number of steps, for example.

### 2.3 SYNTHETIC PROCEDURES

Synthetic procedures for many organic molecules are compiled in both print and online sources. Due to the massive amount of information, print sources are gradually becoming online resources, often requiring a subscription for access. A few are discussed here in detail, and librarians and search engines will readily find more sources of such information.

One source is *Methoden der Organischen Chemie*, known as “Houben-Weyl” for the editors of the first edition. Begun in 1909 with references back into the 1800s, a series of editions exists. The fourth edition was completed in 1987 with 67 volumes and a three volume index. The fourth edition was updated with 23 supplementary volumes called the E-series, and starting in 1990 was published in English. German editions continued until 1995. Now called the *Science of Synthesis*, this resource is available in print from 2001 to 2009, and beginning in 2002, it is available online [15]. The online version requires a subscription and includes content from all the earlier editions. This is an organized, completely referenced, very detailed collection of methods of preparing essentially all classes of organic compounds plus their reactions. It includes selected experimental details and extensive tables of examples and is published by Thieme Publishing Group.

*Comprehensive Organic Chemistry* [16] is a five-volume work plus a sixth volume of indexes again spanning the whole field and providing many leading references. It was published in 1979 by Pergamon Press, Oxford.

*Organic Syntheses* began as a result of the need for research grade chemicals in the United States during World War I. The popularity of the initial set of pamphlets describing the preparation of 111 compounds inspired the first annual volume of 84 pages in 1921 [17]. In 1998, the editorial board decided to place all past and future volumes online [18]. These are searched by structure or keyword. All procedures have been checked and carefully annotated with experimental details and hazards. Citations to the original publications are often available and addenda are linked to the initial article so that changes and additions to the initial procedure are readily found.

*Organic Reactions* is a series of 86 volumes as of this printing that began publication in 1942 under the direction of Professor Roger Adams, who also initiated *Organic Syntheses*. Each volume consists of one or more chapters, with each chapter covering a reaction of particular importance to organic chemistry. The printed books are published by Wiley [19] and the individual chapters are listed on a searchable wikipedia page [20], with links to partial contents. Full access requires a subscription.

## 2.4 HEALTH AND SAFETY INFORMATION

Health and safety information for commercially available chemicals is found on Safety Data Sheets (SDS) which are available from the manufacturer. The United Nations has recently adopted the Globally Harmonized System of Classification and Labelling of Chemicals (GHS), which will be in place by December of 2015. GHS imposes worldwide consistency on classification of hazards such as toxicity and flammability.

SDS are standardized in the GHS and include the sections shown in Table 2.1 [21]. More detailed descriptions of the content of each section can be found online [21, 22].

The GHS requires pictograms on the labels of all chemicals and those pictograms and a brief description are provided in Table 2.2.

The Aldrich catalog, available online [6], offers SDS for all chemicals sold. *Organic Syntheses* [18] reviews each of its published procedures for safety. The Hazardous Substances Data Bank (HSDB) [23] has safety information for 5200 chemicals, including human, animal and environmental toxicity, emergency medical treatment, metabolic and pharmacokinetic data as well as chemical and physical information.