The Chemistry of Heterocycles

Structure, Reactions, Syntheses, and Applications

Mohammad Jafarzadeh

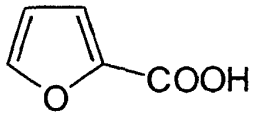
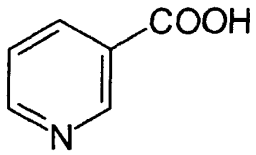
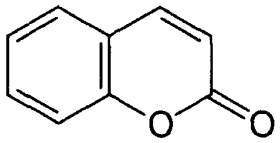
Faculty of Chemistry, Razi University

The Chemistry of Heterocycles, (Second Edition).

By Theophil Eicher and Siegfried Hauptmann, Wiley-VCH Veriag GmbH, 2003

2. Systematic Nomenclature of Heterocyclic Compounds

Many organic compounds, including heterocyclic compounds, have a *trivial name.* This usually originates from the compounds occurrence, its first preparation or its special properties.

Structure	Trivial name	Systematic name
	ethylene oxide	oxirane
	pyromucic acid	furan - 2 - carboxylic acid
	nicotinic acid	pyridine - 3 - carboxylic acid
	coumarin	2*H* - chromen - 2 - one

The derivation of the systematic name of a heterocyclic compound is based on its structure. Nomenclature rules have been drawn up by the IUPAC Commission.

IUPAC **'Blue Book**':

H.R. Panico, W.H. Powell, J.-C. Richer, A Guide to IUPAC Nomenclature of Organic Compounds, Recommendations 1993; Blackwell Scientific: Oxford, 1993.
J. Rigandy, S.P. Klesney Nomenclature of Organic Chemistry; Pergamon: Oxford, 1979.

The IUPAC rules are not given in detail here, rather instructions are given for formulating systematic names with appropriate reference to the Blue Book.

Every heterocyclic compound can be referred back to a parent ring system. These systems have only H-atoms attached to the ring atoms.

The IUPAC rules allow two nomenclatures:

1. Hantzsch- Widman nomenclature, for three- to ten-membered heterocycles

2. Replacement nomenclature, for larger ring heterocycles

2.1 Hantzsch-Widman Nomenclature

- **Type of heteroatom**

The type of heteroatom is indicated by a prefix according to Table 1. The sequence in this table also indicates the preferred order of prefixes (principle of decreasing priority).

Table 1 Prefixes to indicate heteroatoms

Element	Prefix	Element	Prefix
O	oxa	Sb	stiba
S	thia	Bi	bisma
Se	selena	Si	sila
Te	tellura	Ge	germa
N	aza	Sn	stanna
P	phospha	Pb	plumba
As	arsa	B	bora
		Hg	mercura

• Ring size

The ring size is indicated by a suffix according to Table 2. Some of the syllables are derived from Latin numerals, namely ir from tri, et from tetra, ep from hepta, oc from octa, on from nona, ec from deca.

Table 2 Stems to indicate the ring size of heterocycles

Ring Size	Unsaturated	Saturated
3	irene[a]	irane[b]
4	ete	etane[b]
5	ole	olane[b]
6A[c]	ine	ane
6B[c]	ine	inane
6C[c]	inine	inane
7	epine	epane
8	ocine	ocane
9	onine	onane
10	ecine	ecane

a The stem irine may be used for rings containing only N.

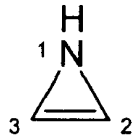
b The traditional stems 'irine', 'etidine' and 'olidine' are preferred for N-containing rings and are used for saturated heteromonocycles having three, four or five ring members, respectively.

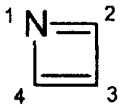
c The stem for six-membered rings depends on the least preferred heteroatom in the ring, that immediately preceding the stem. To detemine the correct stem for a structure, the set below containing this least-preferred heteroatom is selected.
6A: O, S, Se, Te, Bi, Hg; 6B: N, Si, Ge, N, Pb; 6C: B, P, As, Sb

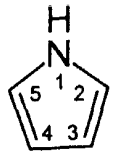
• **Monocylic systems**

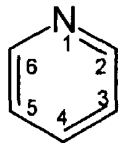
The compound with the maximum number of noncumulative double bonds is regarded as the parent compound of the monocyclic systems of a given ring size.

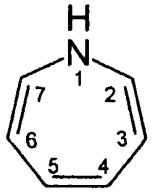
The naming is carried out by combining one or more prefixes from Table 1 with a suffix from Table 2. If two vowels succeed one another, the letter **a** is omitted from the prefix, e.g. azirine (not azairine).

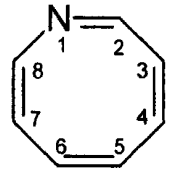

azirine


azete


pyrrole

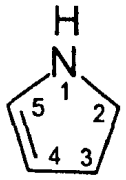

pyridine

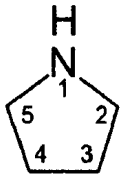

azepine

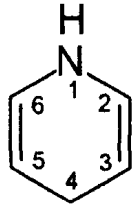

azocine

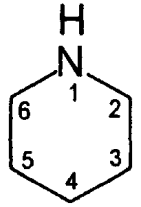
Note that trivial names are permitted for some systems, e.g. pyrrole, pyridine.

Partly or completely saturated rings are denoted by the suffixes according to Table 2. If no ending is specified the prefixes dihydro-, tetrahydro-, etc. should be used.


2,3-dihydropyrrole


pyrrolidine


1,4 - dihydropyridine


piperidine (hexahydropyridine)

• **Monocyclic systems, one heteroatom**

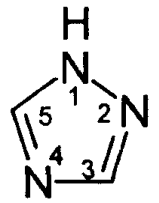
The numbering of such systems starts at the heteroatom.

• **Monocyclic systems, two or more identical heteroatoms**

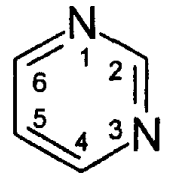
The prefixes di-, tri-, tetra-, etc., are used for two or more heteroatoms of the same kind.

When indicating the relative positions of the heteroatoms, the principle of the lowest possible numbering is used, i.e. the numbering of the system has to be carried out in such a way that the heteroatoms are given the lowest possible set of locants.

In such a numerical sequence, the earlier numbers take precedence, e.g. 1,2,5 is lower than 1,3,4.



1,2,4 - triazole (not 1,3,5 -triazole)



pyrimidine (1,3 - diazine, not 1,5 - diazine)

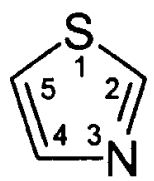
• **Monocyclic systems, two or more different heteroatoms**

For heteroatoms of different kinds, prefixes are used in the order in which they appear in Table 1, e.g. thiazole, not azathiole; dithiazine, not azadithiine.

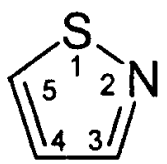
The heteroatom highest in Table 1 is allocated the 1- position in the ring. The remaining heteroatoms are assigned the smallest possible set of number locants.

Although in the first example the systematic name is 1,3-thiazole, the locants are generally omitted because, except for isothiazole (1,2-thiazole), no other structural isomers exist.

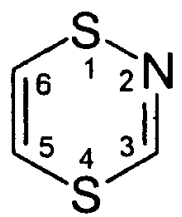
Similar rules apply to oxazole (1,3-oxazole) and isoxazole (1,2-oxazole).



thiazole
(1,3 - thiazole)



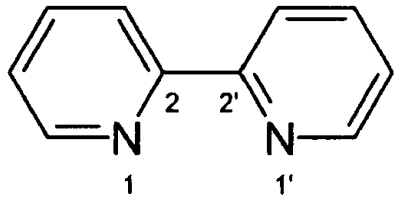
isothiazole
(1,2 - thiazole)



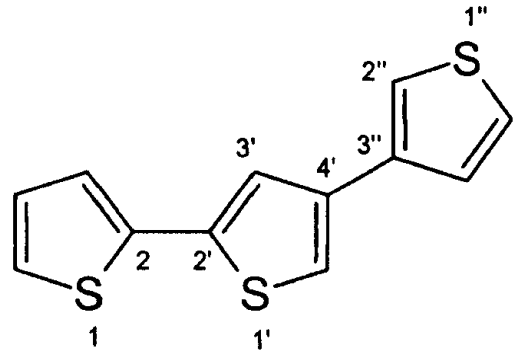
1,4,2 - dithiazine

- **Identical systems connected by a single bond**

Such compounds are defined by the prefixes bi-, tert-, quater-, etc., according to the number of systems, and the bonding is indicated as follows:



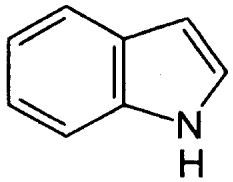
2,2' - bipyridine



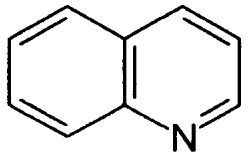
2,2' : 4',3" - terthiophene

- **Bicyclic systems with one benzene ring**

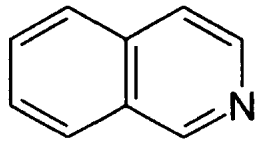
Systems in which at least two neighboring atoms are common to two or more rings are known as fused systems. For several bicyclic benzo-fused heterocycles, trivial names are permitted, e.g.:



indole

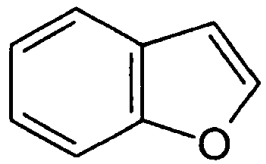


quinoline

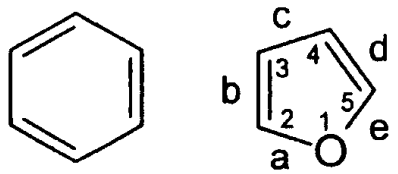


isoquinoline

If this is not the case, and only the heterocycle has a trivial name, then the systematic name is formulated from the prefix benzo- and the trivial name of the heterocyclic component as follows:



benzo [b] furan



furan

The system is dissected into its components. The heterocyclic component is regarded as the base component.

The bonds between the ring atoms are denoted according to the successive numbers of the ring atoms by the letters a, b, c, etc.

The letter **b** in brackets between benzo and the name of the base component denotes the atoms of the base component which are common to both rings.

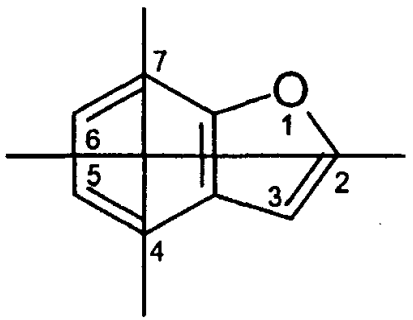
The letter must be as early as possible alphabetically and hence benzo[*d*]furan is incorrect.

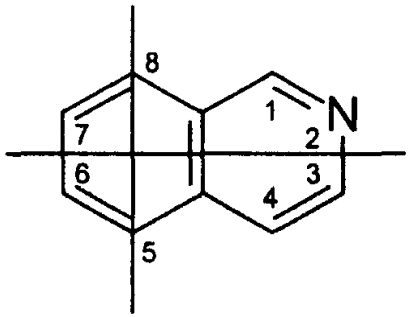
It is generally accepted that the numbering of the whole system in the case of bi- and also polycyclic systems should be done independently of the numbering of the components, and as follows:

The ring system is projected onto rectangular coordinates in such a way that
— as many rings as possible lie in a horizontal row
— a maximum number of rings are in the upper right quadrant.

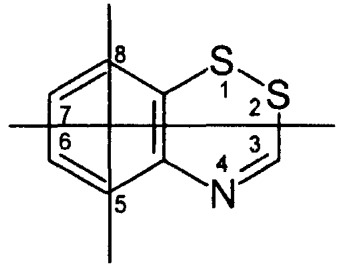
The system thus oriented is then numbered in a clockwise direction commencing with that atom which is not engaged in the ring fusion and is furthest to the left
— in the uppermost ring or
— in the ring furthest to the right in the upper row.

C-Atoms which belong to more than one ring are omitted. Heteroatoms in such positions are, however, included. If there are several possible orientations in the coordinate system, the one in which the heteroatoms bear the lowest locants is valid:

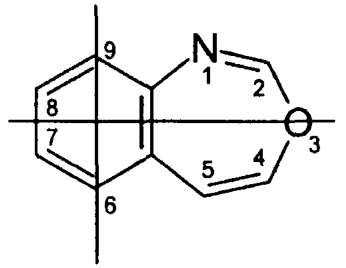




If the base component does not have a trivial name, the entire system is numbered as explained above and the resulting positions of the heteroatoms are placed before the prefix benzo:



1,2,4 - benzodithiazine



3,1 - benzoxazepine

• Bi- and polycyclic systems with two or more heterocycles

First the base component is established. To this end the criteria in the order set out below are applied, one by one, to arrive at a decision. The base component is

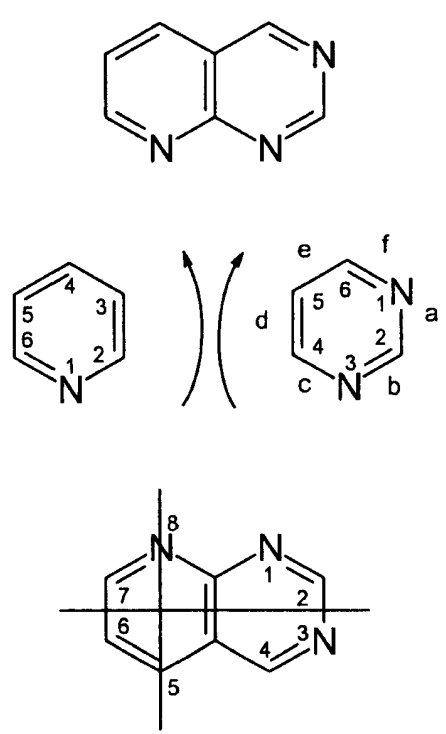
— a nitrogen-containing component
— a component with a heteroatom, other than nitrogen, which is as high as possible in Table 1
— a component with as many rings as possible (e.g. bicyclic condensed systems or polycyclic systems which have trivial names)
— the component with the largest ring
— the component with most heteroatoms
— the component with the largest number of heteroatoms of different kinds
— the component with the greatest number of heteroatoms which are highest in Table 1
— the component with heteroatoms which have the lowest locant numbers.

Two isomers are given as an example:

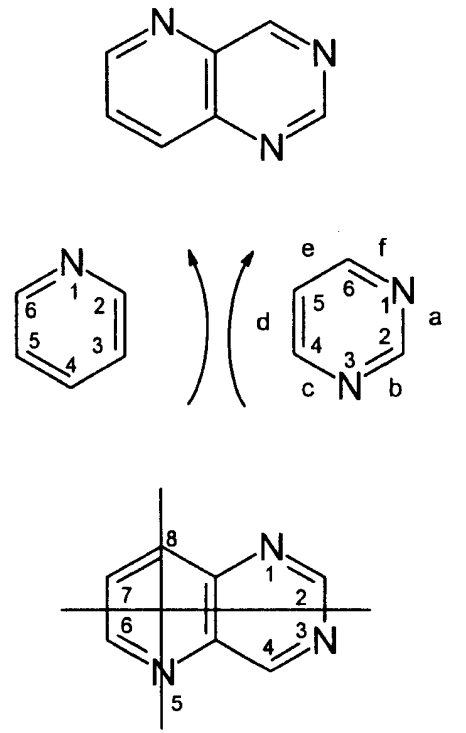
First, the system is dissected into its components. The **base component** cannot be established until the fifth criterion has been reached: pyrimidine.

The bonds between the ring atoms are marked by consecutive lettering according to the serial numbering of the base component.

In contrast to the previous example, the fused component must also be numbered, always observing the principle of assignment to the lowest possible locants.



pyrido[2,3 - d]pyrimidine



pyrido[3,2 - d]pyrimidine

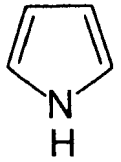
The name of the **fused component**, by the replacement of the terminal ‘e’ with 'o', is put before the name of the **base component**. The atoms common to both rings are described by numbers and letters in square brackets, wherein the sequence of the numbers must correspond to the direction of the lettering of the base component. Finally the whole system is numbered.

• **Indicated hydrogen**

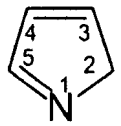
In some cases, heterocyclic systems occur as one or more structural isomers which differ only in the position of an H-atom. These isomers are designated by indicating the number corresponding to the position of the hydrogen atom in front of the name, followed by an italic capital *H*.

Such a prominent H-atom is called an indicated hydrogen and must be assigned the lowest possible locant.

The name pyrrole implies the 1-position for the H-atom.



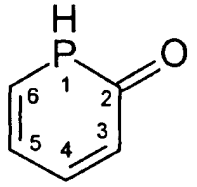
pyrrole



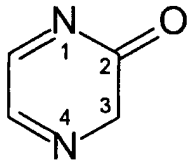
2*H* - pyrrole
(not 5*H* - pyrrole)

3,4 - dihydro - 2*H* - pyrrole
(not 4,5-dihydro-3*H*-pyrrole
or Δ^1 pyrroline)

Heterocyclic compounds in which a C-atom of the ring is part of a carbonyl group are named with the aid of indicated hydrogen as follows:



phosphinin-2-(1*H*)-one



pyrazin-2(3*H*)-one

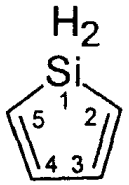
2.2 Replacement Nomenclature

- **Monocyclic systems**

The type of heteroatom is indicated by a prefix according to Table 1. As all prefixes end with the letter **a**, replacement nomenclature is also known as '**a**' nomenclature.

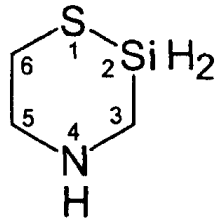
Position and prefix for each heteroatom are written in front of the name of the corresponding hydrocarbon.

Sequence and numbering of the heteroatoms follow the rules given in 2.1.

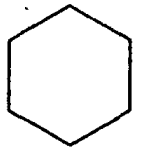


silacyclopenta-2,4-diene

cyclopentadiene



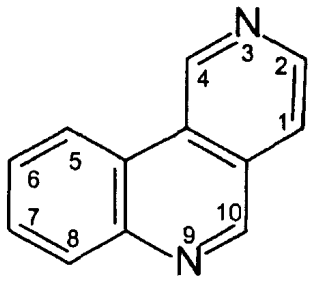
1-thia-4-aza-2-silacyclohexane



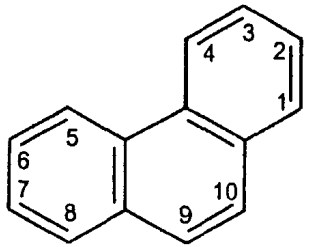
cyclohexane

• **Bi- and polycyclic systems**

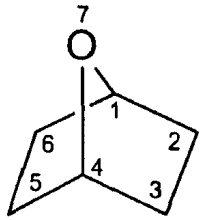
Again, position and prefix are put in front of the name of the corresponding hydrocarbon, but the numbering of the hydrocarbon is retained:



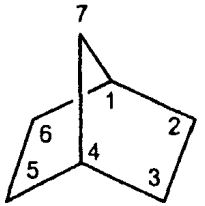
3,9 - diazaphenanthrene



phenanthrene

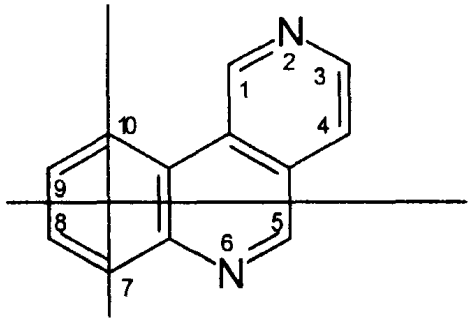


7 - oxabicyclo 2.2.1 heptane



bicyclo 2.2.1 heptane

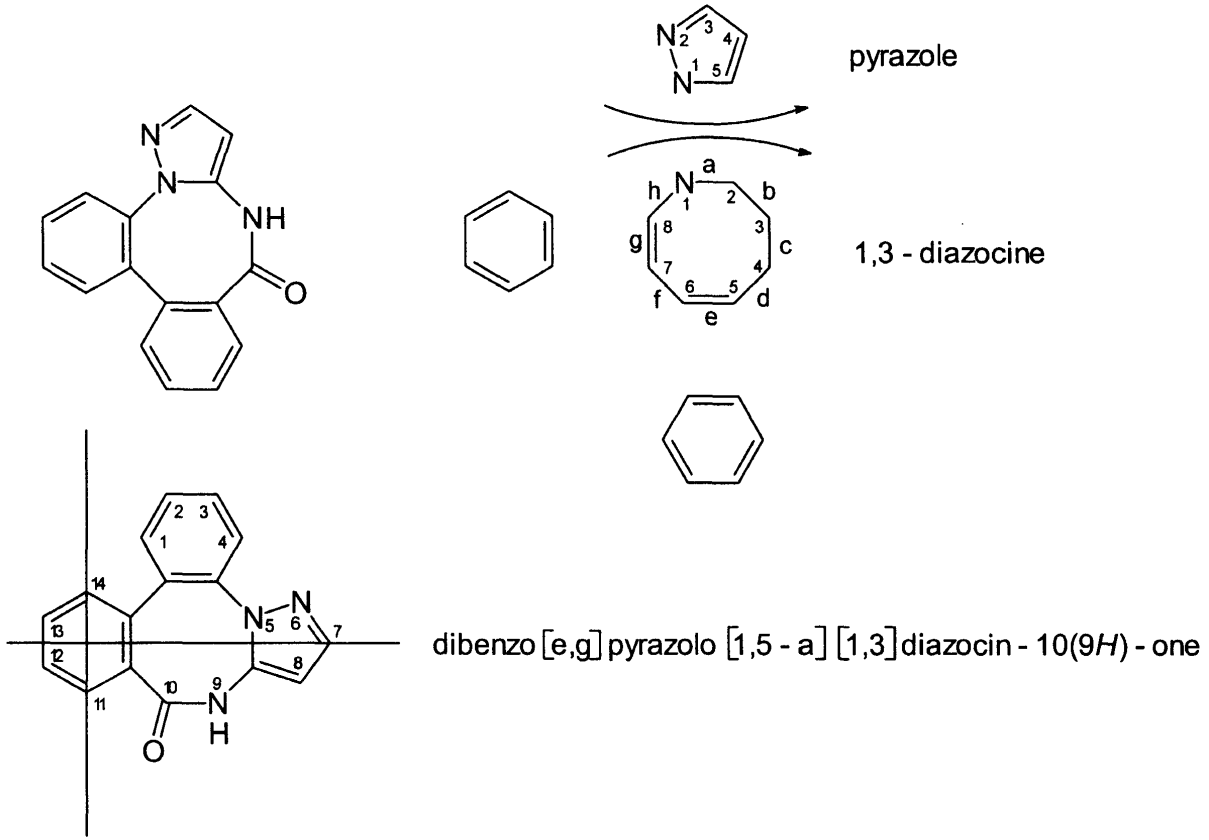
The Hantzsch-Widman nomenclature can only be applied to the first example and this then results in different numbering.



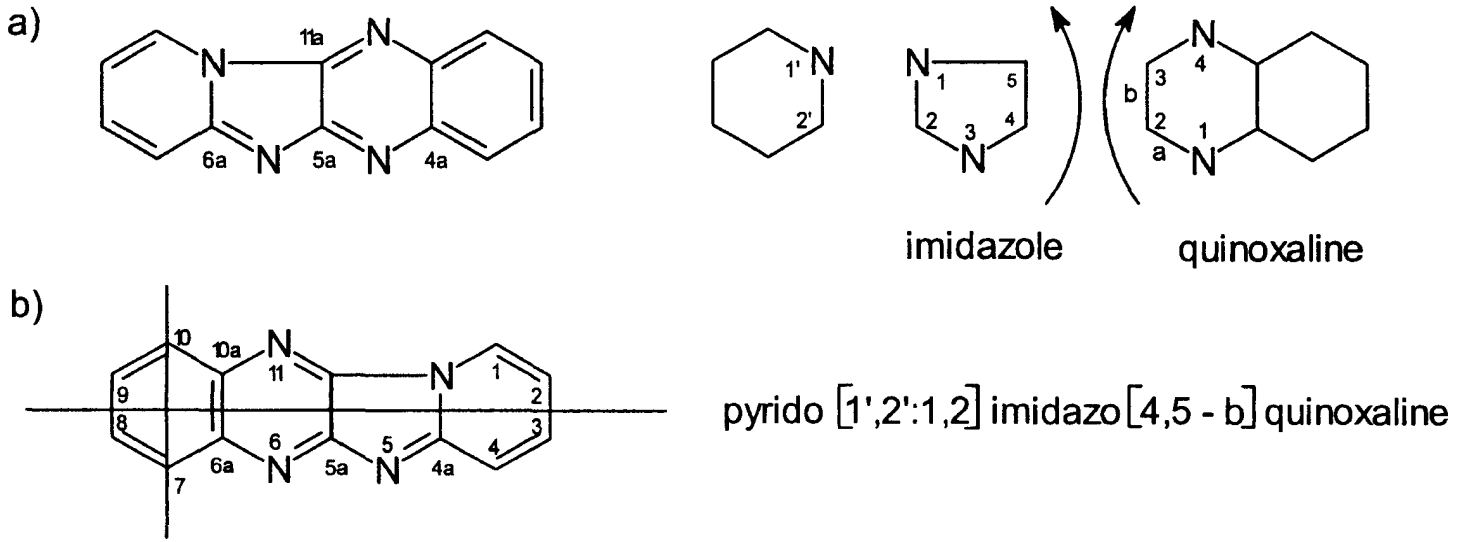
pyrido[4,3 - c]quinoline

2.3 Examples of Systematic Nomenclature

Finally, the systematic nomenclature of heterocyclic compounds will be illustrated by a few complex examples.



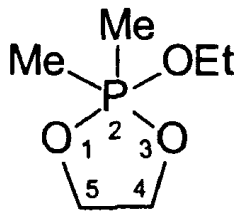
An analysis of the system reveals two benzene rings, one pyrazole ring and one 1,3-diazocine ring, the latter ring being the base component according to the fourth criterion. The square brackets [1,3] indicate that the position of the two heteroatoms is not the basis for numbering the whole system.



According to the third criterion, quinoxaline is the base component. The heterocycle imidazole, which is fused to the base component, is numbered in the usual way; the pyridine ring, however, is denoted by 1', 2', etc., and it is not necessary to mark the double bonds.

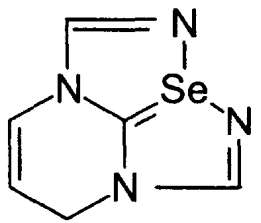
Pyrido[1',2':1,2]imidazo denotes one ring fusion, imidazo[4,5-*b*]quinoxaline the other. For numbering polycyclic systems, five-membered rings must be drawn as shown above and not as regular pentagons.

For the orientation in a system of coordinates, an additional rule has to be observed, namely that C-atoms common to two or more rings must be given the lowest possible locant. The numbering in (b) is therefore correct, while that in (a) is wrong, because 10a<11a.

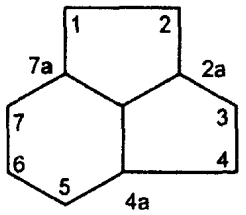


2 - ethoxy - 2,2 - dimethyl - 1,2,3 λ^5 - dioxaphospholane
(the standard bonding number of P is 3)

With ring atoms such as phosphorus, which can be tri- or pentavalent, a non-standard bonding number 5 is indicated as an exponent of the Greek letter λ after the locant. In the example, this is shown by λ^5.

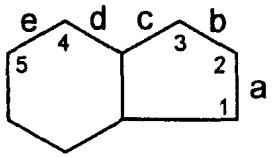


5*H* - 2a λ^4 - selena - 2,3,4a,7a - tetraazacyclopenta [c,d] indene



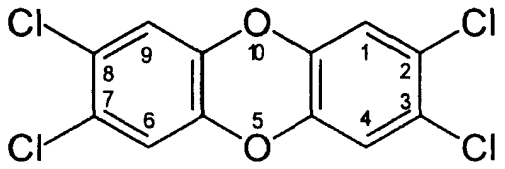
cyclopenta [c,d] indene

cyclopentadiene



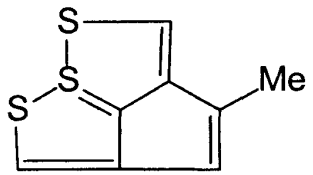
indene

The name is constructed according to replacement nomenclature. The basic hydrocarbon with the greatest number of noncumulative double bonds is cyclopenta[*c,d*]indene. Note the retention of the numbering.

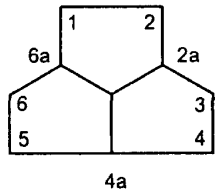


2,3,7,8 - tetrachlordibenzo [1,4] dioxin

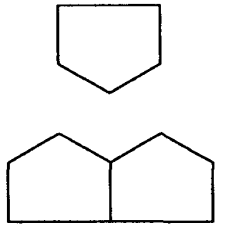
In this case, [*b,e*] is omitted after dibenzo since there is no other possibility for ring fusion. This compound is also known as TCDD or Seveso dioxin.



3 - methyl - 1,6,6aλ^4 - trithia-cyclopenta c,d pentalene

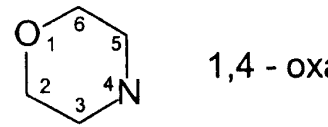


cyclopenta c,d pentalene

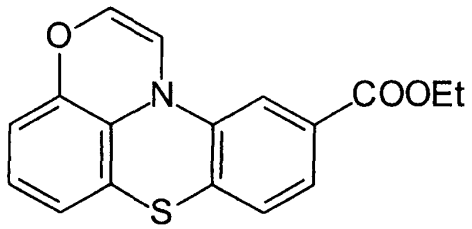


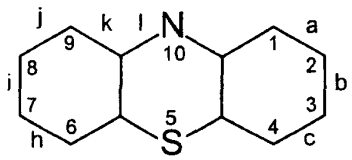
cyclopentadiene

pentalene



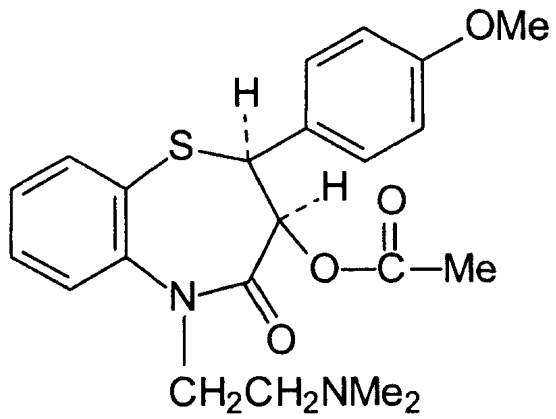
1,4 - oxazine





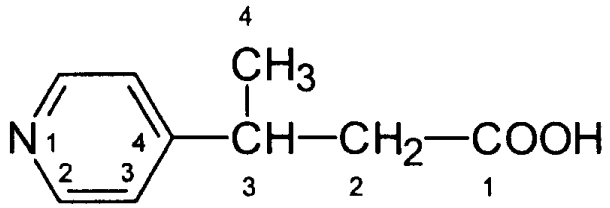
phenothiazine

ethyl [1,4] oxazino [2,3,4 - *kl*] phenothiazine - 6 - carboxylate



(2S, 3S) - 3 - acetoxy - 5 - (2 - dimethylaminoethyl) 2 - (4 - methoxyphenyl) - 2,3,4,5 - tetrahydro - 1,5 - benzothiazepin - 4 - one

So far in all the examples, the base compound has been the heterocyclic system. If this is not the case, the univalent radical of the heterocyclic system is regarded as a substituent, e.g.:



3 - (4-pyridyl) butyric acid

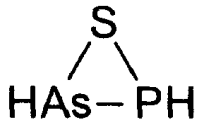
The names of some univalent heterocyclic substituent groups are to be found in the list of trivial names in the 1993 Blue Book, p 172.

The most important source of information on heterocyclic and isocyclic systems is the Ring Systems Handbook of the Chemical Abstracts Service (CAS) published by the American Chemical Society. The 1988 edition is arranged as follows:

The *Ring Systems File* is a catalogue of structural formulas and data. It lists the systems consecutively with numbering RF 1-RF 72861 on the basis of a ring analysis. The *Ring Systems File* starts with the following system:

The ring analysis shows:

1 RING: 3
AsPS



1 RING represents a monocycle, 3 denotes the ring size. The ring atoms are listed underneath in alphabetical order followed by

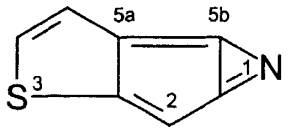
RF 1 [Ring File (RF) Number]
88212-44-6 (CAS Registry Number)
Thiaphospharsirane
AsH_2PS

the systematic name and molecular formula, and furthermore Wiswesser Notation, Chem. Abstr. reference (Chem. Abstr. volume number, abstract number), structural diagram.

An example from the *Ring Systems File* 1, p 758, is given below:

3 RINGS: 3,5,5
$C_2N–C_4S–C_5$

RF 15037 113688-14-5
Thieno[3',2':3,4]cyclopent[1,2-*b*]azirin
C_7H_3NS
T B355 CN GSJ
CA 108:112275y



The *Ring Formula Index* is a list of molecular formulas of all ring systems with ring atoms quoted in alphabetical order, H-atoms being omitted, e.g. C_6N_4: 2 RINGS, CN_4-C_6N, 1*H*-Tetrazolo[l,5-*a*]azepine [RF 9225].
With the aid of the Ring File Number RF 9225, the structural formula can be found in the Ring Systems File.

The *Ring Name Index* is an alphabetical list of the systematic names of all ring systems, e.g.: Benzo[4,5]indeno[l,2-*c*]pyrrole [RF 40064]. The Ring File Number allows access to the Ring Systems File.

2.4 Important Heterocyclic Systems

Arrangement of chapters 3-8: heteroarenes dealt with first, followed by the heterocycloalkenes and finally the heterocycloalkanes.

The guiding principle is ring size.

Heterocycles of certain ring sizes are further subdivided according to the type of heteroatoms, starting with one heteroatom, two heteroatoms, etc.

The description of each heterocyclic system is then arranged as follows:

[A] structure, physical and spectroscopic properties
[**B**] chemical properties and reactions
[**C**] syntheses
[D] important derivatives, natural products, drugs, biologically active compounds, industrial intermediates
[E] use as reagents, building-blocks or auxiliaries in organic synthesis.