CHAPTER 10 PARENT STRUCTURES FOR NATURAL PRODUCTS AND RELATED COMPOUNDS

P-100 Introduction

This Chapter is based on the recent publication 'Revised Section F: Natural Products and Related Compounds, IUPAC Recommendations 1999' (ref. 9).

In the field of natural products, three levels of nomenclature are recognized. A new compound, isolated from a natural source, is generally given a 'trivial' name. By common usage, these trivial names are commonly related to the biological origin of the material, but frequently not in a rational way, since the available structure is not known with great detail. These trivial names are considered to be ephemeral and replaced for chemical purposes by names describing the skeleton, the characteristic groups and the organyl substituent groups.

When the full structure is known, a 'systematic name' may be generated in accordance with Rules described in Chapters 1 through 8 of these recommendations. However, this name may be too cumbersome to be continually inserted into the text of a scientific paper. To overcome this difficulty and show the close similarity to related compounds, a 'semisystematic name' can be formed.

Semisystematic names are based on specific parents, generally including the stereochemistry, that can later on be used to fully describe a compound by using the rules of systematic nomenclature. There are two general types of semisystematic parent structures used for naming natural products and related compounds:

(a) parent hydrides, i.e., structures that do not have terminal heteroatoms or functional groups and therefore consist only of skeletal atoms and hydrogen, for example, in steroid (ref. 39), terpene, carotene (ref. 40), corrinoid (ref. 41), tetrapyrrole (ref. 42), alkaloid nomenclature. and lignans and neolignans (43). This type of semisystematic parent is analogous to the parents described in Chapter 2 and is treated in the same manner to generate complete names;

(b) functional parents, that are analogous to the functional parents described in Section P-35, and used in amino acid and peptide (ref. 23), carbohydrate (ref. 22), cyclitol (ref. 44) nucleoside (ref. 45), nucleotide, and lipid (ref. 45) nomenclature; they have characteristic groups implied in their name and can be modified by specific rules and by methods used in systematic nomenclature.

In this Chapter, Section P-101 describes the rules to form trivial names and semisystematic names used as parent hydrides, and those related to their skeletal transformation and functionalization for naming alkaloids, steroids, terpenes, and some related compounds. Section 102 describes the rules for naming carbohydrates, P-103 deals with the nomenclature of amino acids and peptides, P-104 describes the nomenclature for cyclitols, P-105 and P-106 deal with nucleosides and nucleotides, and finally P-107 discusses the nomenclature of lipids. If difficulties are encountered, consultation of the full publications may be necessary, as indicated in each Section. Preferred IUPAC names are indicated when a choice is possible.

P-101 Nomenclature for natural products based on parent hydrides (alkaloids, steroids, terpenes, carotenes, corrinoids, tetrapyrroles, and similar compounds)

P-101.1 Biologically based trivial names

P-101.1.1 When a compound is isolated from a natural source and a trivial name is required, the name should be based, whenever possible, on the family or genus or species name of the biological material from which it has been isolated. If appropriate, the class or order might also be used for the name of a compound that occurs in a number of related families.

P-101.1.2 The ending ‘une’ or, for euphonic reasons, ‘iune’ is used to indicate that the trivial name it terminates describes a compound of unknown structure.

P-101.2 Semisystematic nomenclature for natural products based on parent hydrides

As soon as the structure of a new natural product has been fully determined, the trivial name should be abandoned in favor of a systematic name formed by the Rules prescribed in Chapters 1 to 9 for systematic nomenclature of organic compounds. For a more complicated structure, an existing semisystematic name listed in P-101.2.7 is used to fully name the compound. If a

previously known parent structure cannot be found, a new parent structure is formed and numbered as follows.

P-101.2.1 General guidelines for choosing a parent structure

P-101.2.1.1 A parent structure should reflect the basic skeleton (including nonterminal heteroatoms and hetero groups) that is common to most compounds of the class.

P-101.2.1.2 Parent structures should be chosen so that as many natural products as possible can be derived from each by well-defined operations and rules of the nomenclature of organic compounds.

P-101.2.1.3 A parent structure should include as much stereochemistry as possible that is common to the relevant class of natural products. Such parent structures are called 'stereoparents'

P-101.2.2 Structural features allowed for parent structures

The following rules are applicable to new parent structures. Existing parent structure names are considered as retained names if they do not follow the new rules.

P-101.2.2.1 A parent structure should exceptionally include rings that are part of a characteristic group, such as a lactone or cyclic acetal.

P-101.2.2.2 A parent structure should not contain heteroatoms or groups.

P-101.2.2.3 A parent structure should contain acyclic hydrocarbon groups that occur in most of the compounds in the natural product class.

P-101.2.2.4 A parent structure should be as nearly fully saturated or fully unsaturated in terms of maximum number of noncumulative double bonds (mancude rings), while still representing the level of saturation (or unsaturation) of as many related compounds as possible.

P-101.2.2.5 A semisystematic name for a parent structure should be derived, as far as possible, from a trivial name formed in accordance with P-101. The endings to be used in place of 'une' or 'iune' must be assigned as follows:

(a) 'ane', if the entire stereoparent hydride is fully saturated;

(b) 'ene', if the cyclic or the main chain of the acyclic part contains the maximum number of noncumulative double bonds;

(c) 'arane', if, in an otherwise fully saturated parent structure, one or more individual mancude rings is present.

Existing names of parent structures in which endings are different from those indicated above, for example morphinan and ibogamine, are considered as retained names.

P-101.2.2.6 Indicated hydrogen, as defined in P-14.6, is used to describe isomers.

P-101.2.3 Numbering of parent structures

P-101.2.3.1 A numbering pattern established among a group of structurally related natural products is used for numbering the skeletal atoms of the stereoparent structure, providing that all skeletal atoms have been included in the numbering system.

P-101.2.3.2 If no numbering pattern has been become established among the members of a group of structurally related natural products, the stereoparent hydride is numbered according to the following guidelines:

(a) examine the skeleton to identify the senior ring or ring system system, according to P-44. The locant '1' is assigned to the atom of the preferred ring system whose locant would be '1' according to systematic numbering for that particular ring or ring system;

(b) assign all skeletal atoms of the senior ring system consecutive arabic numbers, including atoms of fusion positions in fused ring systems, beginning with the locant '1', and following the path prescribed for that particular type of ring or ring system;

(c) number acyclic substituents to skeletal atoms of ring components or connecting acyclic structures each in its entirety, including branches, in order of the increasing value of the locant of the skeletal atom to which it is attached;

(d) number skeletal atoms of acyclic connections to other ring or ring systems, if any, consecutively beginning with the atom next to the senior ring system, followed by the skeletal atoms of the other rings or ring systems as prescribed in (b) above; if two or more acyclic connections to other rings or ring systems are present, the one attached to the senior ring or ring system at the lowest numbered position is numbered first, then the ring attached to it, followed by the acyclic connector at the next lower position of the senior ring or ring system, etc.;

(e) number the larger group, in terms of the number of skeletal atoms, between two groups at a geminal disubstituted position, first; if there is still a choice, alphanumerical order is followed (Rule P-14.5) . If the two groups are then identical and attached to a stereoparent structure properly drawn (see Appendix in ref. 9), the group that is stereochemically α is numbered first; if the two groups are identical and attached to an acyclic terminal double bond, the group '*trans*' to the main chain is numbered first as described in the carotenoid recommendations (Rule 12.4 in ref. 40).

P-101.2.4 Identification of individual rings

Since locants are used to describe skeletal modifications, as indicated in P-101.3, the identification of individual rings by letters A, B, C, etc., used in the past is no longer recommended, except for the rather special case of the removal of a terminal ring (see P-101.3.6). Nevertheless, to provide continuity with the use of this system, names using letters to identify rings are given where appropriate, but are no longer recommended.

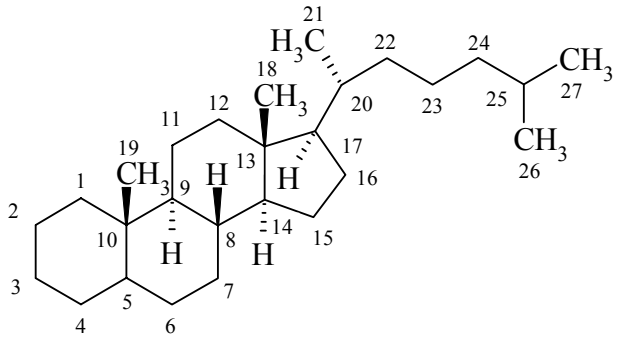
P-101.2.5 Atomic connector, terminal segment and bond connector

For nomenclature purposes, the fundamental structures are described by specific arrangements of atoms or groups of atoms called 'atomic connectors', 'terminal segments' and 'bond connectors', that must be taken into consideration in accordance with the additive or subtractive operations modifying a fundamental structure.

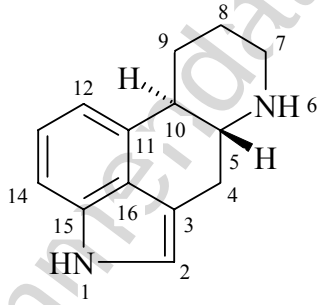
An 'atomic connector' is a chain of homogeneous skeletal atoms of the same element connecting any combination of bridgehead or ring junction atoms, rings, or ring systems (i.e. ring assemblies), substituted skeletal atoms in parent structure, or heteroatoms. A 'terminal segment' of a skeletal structure is an acyclic portion of homogeneous skeletal atoms connected at only one end by the features of structure that terminate atomic connectors. A 'bond connector' is a

connection between any combination of bridgehead or ring junction atoms, rings or ring systems (i.e. ring assemblies), substituted skeletal atoms, or heteroatoms. The structures below illustrate atomic connectors, bond connectors and terminal segments. The use of these terms is further illustrated in P-101.3.1 in relation to the removal of skeletal atoms denoted by the prefix 'nor'.

Examples:



cholestane



ergoline

Atomic connectors:

In cholestane: 1-4, 6-7, 11-12, 15-16, and 22-24.
In ergoline: 2, 4, 7-9, and 12-14

Terminal portions:

In cholestane: 18, 19, 21, 26, and 27.
In ergoline: none

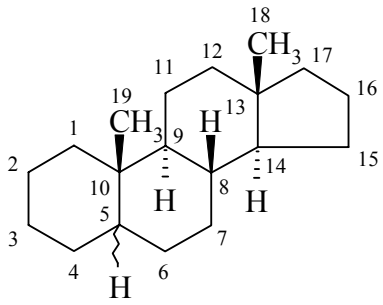
Bond connectors:

In cholestane: 5-10, 8-9, 8-14, 9-10, 13-14, 13-17, and 17-20
In ergoline: 1-15, 3-16, 5-6, 5-10, 10-11, 11-16, and 15-16

P-101.2.6 Stereochemical configuration of parent structures

The name of a parent structure usually implies (the 5α position in steroids is usually not defined), without further specification, the absolute configuration of all chiral centers and the configuration of double bonds when applicable. When a planar or quasi planar system of rings is denoted as a projection, as in this recommendation, an atom or group attached to the ring is called α if it lies below or β if it lies above the plane of the paper. Use of this system requires the orientations of structure as given in the examples used to exemplify the various rules. In the example below, the implied configurations shown define the attached hydrogen atoms and methyl groups at positions 8, 10, and 13, as β and at positions 9 and 14 as α; here, the configuration of the hydrogen atom at position 5 is not known and thus the orientation is ξ (xi), denoted by a wavy line in the formula. The stereodescriptors α, β and ξ used to describe implicit or indicated configuration are cited before the name of the fundamental parent structure without parentheses.

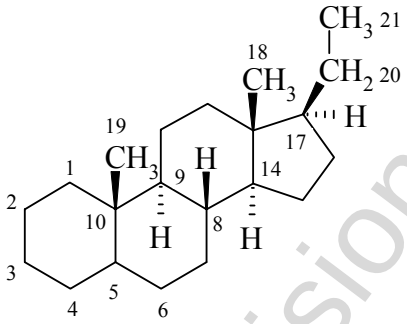
The α/β symbolism is used as defined above and extended in the following way to express different aspects of the configuration of modified fundamental parent structures.



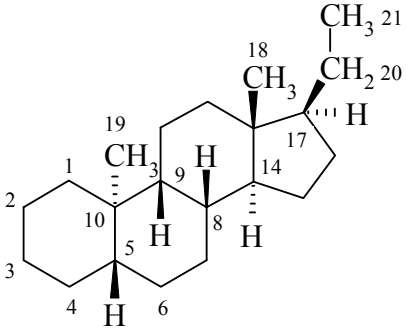
P-101.2.6.1 Configurations that are different from those in the parent structure

P-101.2.6.1.1 At chiral centers, the α/β system is used as described or in IUPAC-IUBMB recommendations for the nomenclature of steroids (ref. 39). Each chirality center is described by the stereodescriptor α, β, or ξ to indicate a configuration that must be specified and those that are inverted. The symbols α, β, or ξ, preceded by the appropriate locants, are placed immediately at the beginning of the name of the fundamental parent structure. In the following examples, configuration at C-5 must be specified; configurations at bridgeheads C-9 and C-10 are inverted when compared with those of the fundamental parent structure. This method is preferred to the alternative described in P-101.2.6.1.2 .

Examples:



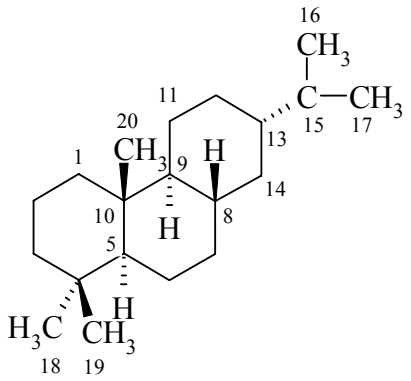
pregnane (PIN)
(fundamental parent structure)



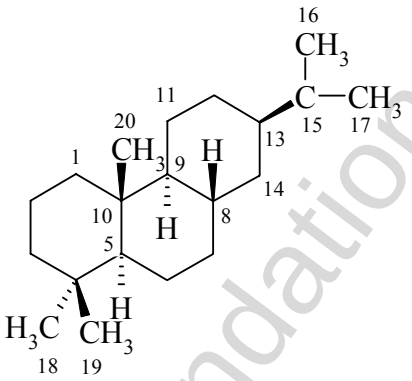
5β,9β,10α-pregnane (PIN)

A change in configuration of a nonbrigeheaded side chain that is part of the parent is denoted by the method specified for C-17 of steroids (see 3S-5.2, ref. 39), where α or β refers to the sidechain itself and not to the hydrogen atom in the same position.

Example:



abietane (PIN)
(fundamental parent structure)

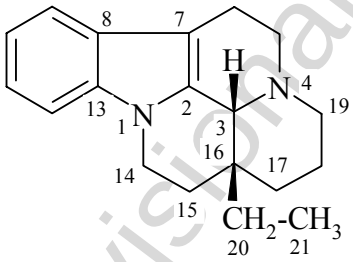


13β-abietane (PIN)

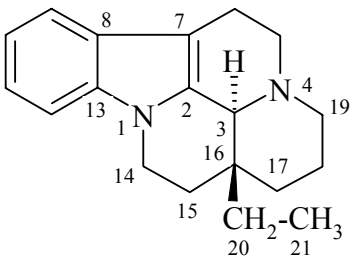
P-101.2.6.1.2 Configurational inversion at one of stereogenic centers whose configuration is implied or stated in the name of the fundamental parent structure can be indicated by the italicized prefix *epi* (derived from 'epimer') placed at the front of the name of the parent structure and prefixed by the locant of the affected atom. As method P-101.2.6.1.1 leads to preferred IUPAC names, this method may be used only in general nomenclature.

The preferred IUPAC name 13β-abietane, described above in P-101.2.6.1.1, can also be named 13-*epi*-abietane.

Example:



eburnamine (PIN)
(fundamental parent structure)



3-*epi*-eburnamine
3α-eburnamine (PIN)

P-101.2.6.1.4 The stereodescriptors *R* and *S*

The stereodescriptors '*R*' and '*S*' are used to describe the absolute configuration not specified by the α/β system described above, in accord with the CIP priority system and the rules and conventions described in Chapter 9 (see P-92.3). The stereodescriptors '*R*' and '*S*' are also used when a ring is opened and two chiral centers created, one of which may rotate, as described for vitamin D in P-101.8.4.

P-101.2.7 Semisystematic names of recommended fundamental stereoparent structures are listed in Table 10.1. All except retinal, cepham, and penam are parent hydrides.

Table 10.1 List of recommended names of stereoparent structures
(this list is nonlimiting; structures are shown in the Appendix in ref. 9)

1. Names recommended as preferred IUPAC names.

(a) alkaloids

aconitane	emetan	oxayohimban
ajmalan	ergoline	oxyacathan
akuamilan	ergotaman	pancracine
alstophyllan	erythrinan	rheadan
aporphine	evonimine	rodiasine
aspidofractidine	evonine	samandarine
aspidospermidine	formosanan	sarpagan
atidane	galanthamine	senecionan
atisine	galanthan	solanidine
berbaman	hasubanan	sparteine
berbine	hetisan	spirosolane
cephalotaxine	ibogamine	strychinidine
cevane	kopsan	tazettine
chelidonine	lunarine	tropane
cinchonan	lycopodane	tubocuraran
conanine	lycorenan	tubulosan
corynan	lythran	veratraman
corynoxan	lythrancine	vincaleucoblastine
crinan	lythranidine	vincane
curan	matridine	vobasan
daphnane	ormosanine	vobtusine
dendrobane	morphinan	yohimban
eburnamine	nupharidine	

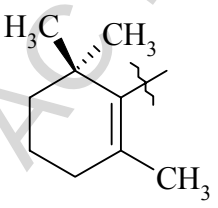
(b) steroids

androstane	cholestane	gorgostane
bufanolide	ergostane	poriferastane
campestane	estrane	pregnane
cardanolide	furostan	spirostan
cholane	gonane	stigmastane

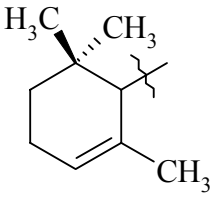
(c) terpenes (all are parent hydrides except retinal)

abietane	eremophilane	neolignan
ambrosane	eudesmane	oleanane
aristolane	gammacerane	ophiobolane
atisane	germacrane	picrasane
beyerane	gibbane	pimarane
cadinane	grayanotoxane	podocarpane
β,φ-carotene[1]	guaiane	protostane
β,ψ-carotene[1]	himalachane	retinal
ε,κ-carotene[1]	hopane	rosane
ε,χ-carotene[1]	kaurane	taxane
caryophyllane	labdane	trichothecane
cedrane	lanostane	ursane
dammarane	lignan	
drimane	lupane	

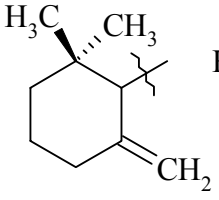
[1] Four different carotenes are exemplified; there are 28 carotene parent structures derived from all permutations of the seven following end groups:



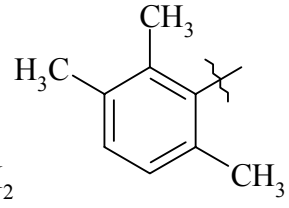
β



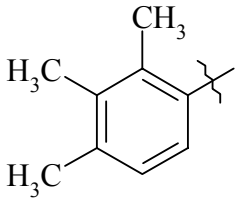
α (alpha)



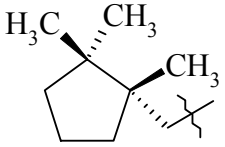
γ (gamma)



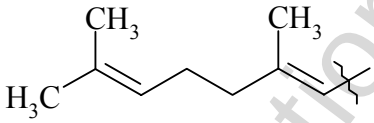
(theta)



χ (chi)



κ (kappa)



ψ (psi)

(d) Miscellaneous (all are parent hydrides except cepham and penam)

biline	prostane
cepham	porphyrin
corrin	thromboxane
penam	

(2) Names recommended for general nomenclature only

(a) Terpenes

bisabolane	humulane
bornane	menthane (*p*-isomer)
carane	pinane
fenchane	thujane

(b) Miscellaneous

flavan
isoflavan
neoflavan

P-101.3. Skeletal modifications of parent structures

P-101.3.0 Introduction

The skeleton of parent structures can be modified in many ways, contracted, expanded, or rearranged by using operations described in P-13. These operations are denoted by specific nondetachable prefixes that are added to the name of the parent structure. Changes affecting the configuration must be shown as indicated in P-101.2.6 and P-102. Preferred IUPAC names result from a narrow interpretation of the number of operations that are allowed in order to preserve the integrity of the ring system and of the absolute configuration. In natural product nomenclature and general nomenclature, the number of operations is not subject to limitations.

> This Section supersedes the provisional Section F Rules and Rules A-71 to related to terpene hydrocarbons as prescribed in the 1979 Recommendations (ref. 1).

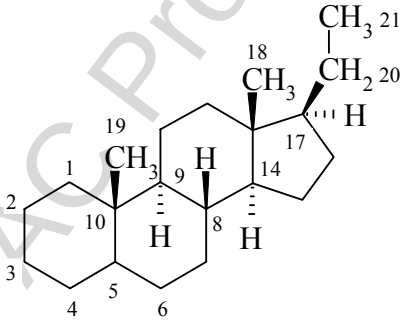
P-101.3.1 Removal of skeletal atoms without affecting the number of rings

P-101.3.1.1 The removal of an unsubstituted skeletal atom, saturated or unsaturated, from a ring or of an unsubstituted skeletal atom from an saturated acyclic portion of a parent structure with its attached hydrogen atom(s) is described by the nondetachable prefix ‘nor’; the loss of two or more such skeletal atoms are indicated by the usual numerical multiplicative prefixes ‘di’, ‘tri’ added before ‘nor’.

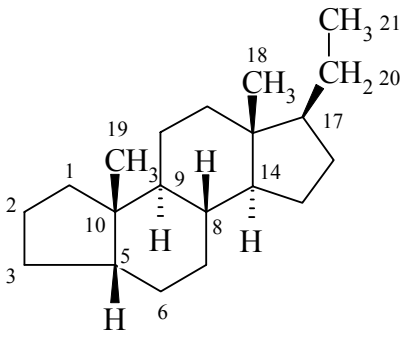
The position of the skeletal atom that is removed is denoted in all cases by its locants in the numbering of the fundamental parent structure. Although, because the locant of each skeletal atom removed is cited, an unambiguous name can be generated by the removal of any skeletal atom, carbon atom or heteroatom, it is traditional to remove skeletal atoms with the highest possible locant in an atomic connector in a cyclic portion of the skeletal structure. In carotenes, as an exception, the locant attached to ‘nor’ is the lowest possible (see Rule Carotenoid 5.1, ref, 27).

For preferred IUPAC names, modifications by the prefix ‘nor’ are limited to fundamental structures having two or more rings; ring contraction and chain shortening are limited to the removal of a maximum of two methylene groups. When the use of a fundamental structure is not allowed, preferred IUPAC names are substitutive names, systematically formed, with CIP stereodescriptors to describe configurations.

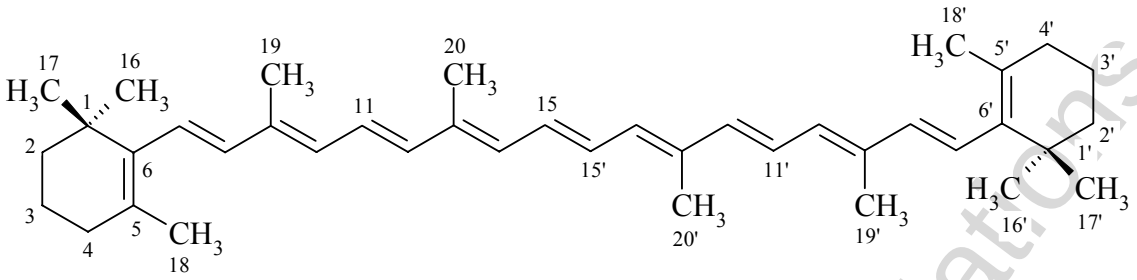
Examples:



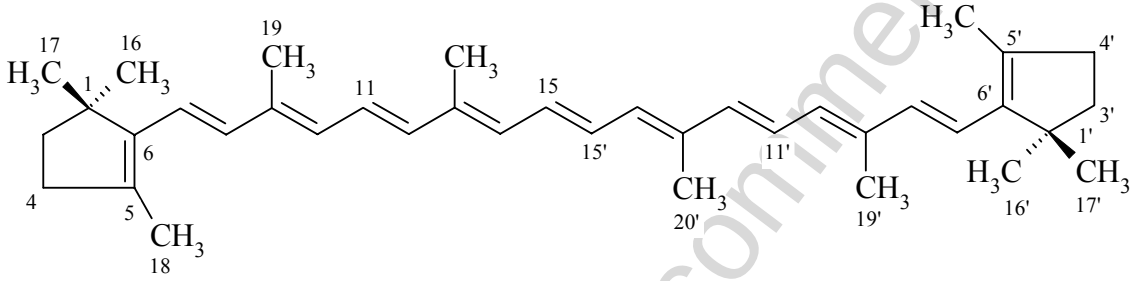
pregnane (PIN)
(fundamental parent structure)



4-nor-5β -pregnane (PIN)



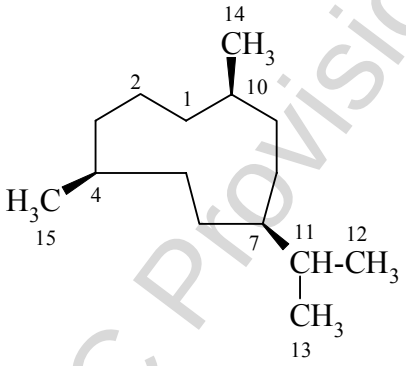
β,β-carotene (PIN)
(fundamental parent structure)



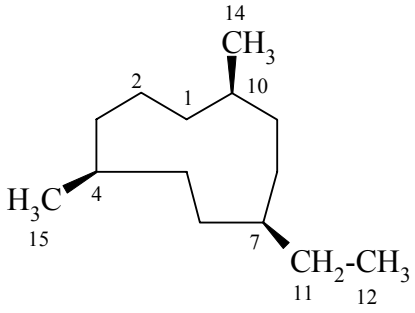
2,2′-dinor- β,β-carotene (PIN)

In an acyclic portion of a skeletal structure, the skeletal atom removed preferably is the one of an acyclic atomic connector or a terminal segment nearest to the free end of this acyclic portion. (This is done in order to maintain as far as possible the numbering of structural features of the compound and of compounds derived from it).

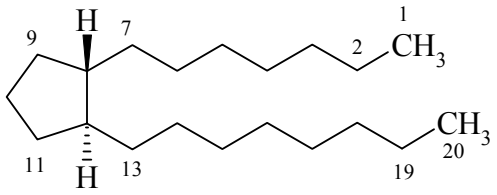
Examples:



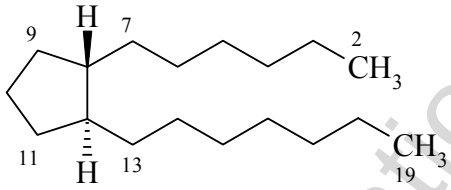
germacrane (PIN)
(fundamental parent structure)



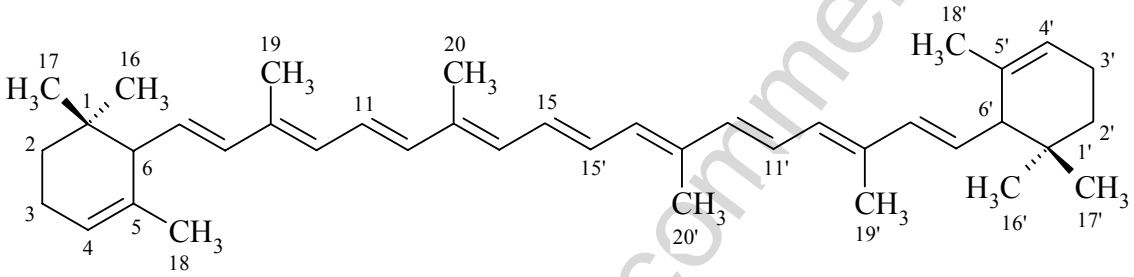
13-norgermacrane
(1*s*,4*R*,8*S*)-1-ethyl-4,8-dimethylcyclodecane (PIN)
(the prefix ‘nor’ is not permitted for monocyclic stereoparents)



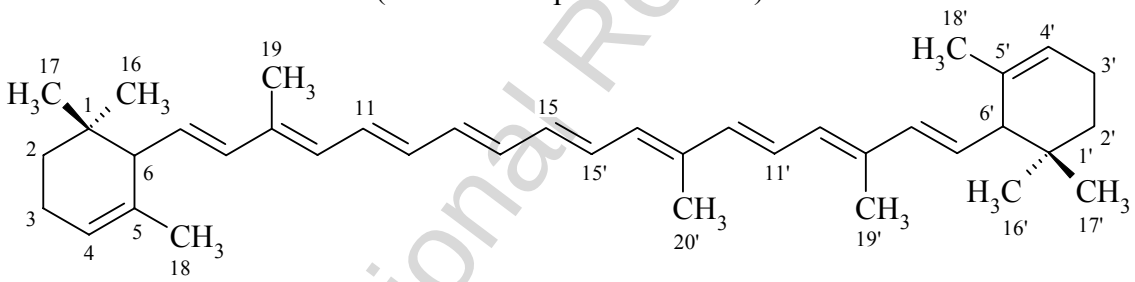
prostane (PIN)
(fundamental parent structure)



1,20-dinorprostane (PIN)



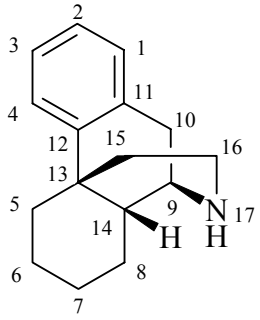
ε,ε-carotene (PIN)
(fundamental parent structure)



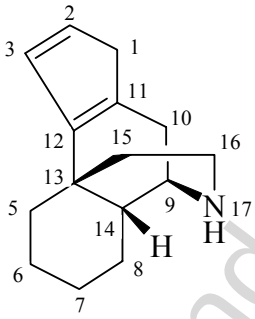
20-nor-ε,ε-carotene (PIN)

P-101.3.1.2 When the removal of an unsaturated skeletal atom from a mancude ring (containing the maximum number of noncumulative double bonds) in the parent structure results in the creation of a saturated ring position, this position is described by indicated hydrogen (see P-14.6). In names, the symbol *H*, denoted by the appropriate locant, is cited at the beginning of the name modified by the nondetachable prefix.

Example:



morphinane
(fundamental parent structure)



1*H*-4-normorphinane

P-101.3.2 Addition of skeletal atoms without affecting the number of rings

P-101.3.2.1 The addition of a methylene (–CH_2–) group between two skeletal atoms of a parent structure is described by the nondetachable prefix 'homo'; the addition of two or more methylene groups is indicated by the numerical multiplicative prefixes 'di', 'tri', etc. Positions of inserted methylene groups in the modified fundamental parent structure are indicated by the locants of the added methylene groups which are cited in front of the prefix 'homo', preceded by multiplicative prefixes when required.

The assignment of the locants to an added methylene group depends on whether it is considered to be inserted into an atomic connector or terminal acyclic portion or into a bond connector.

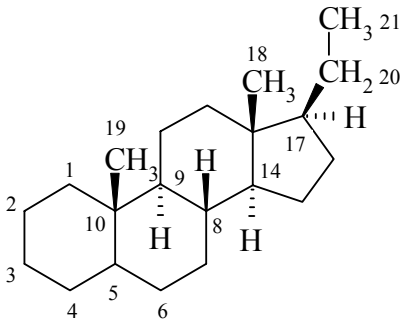
For preferred IUPAC names, modifications by the prefix 'nor' are limited to fundamental structures having two or more rings, and ring enlargement and chain lengthening are limited to the addition of a maximum of two methylene groups. When the use of a fundamental structure is not allowed, preferred IUPAC names are substitutive names systematically formed, with CIP stereodescriptors to describe configurations.

P-101.3.2.2 Numbering of additional skeletal atoms

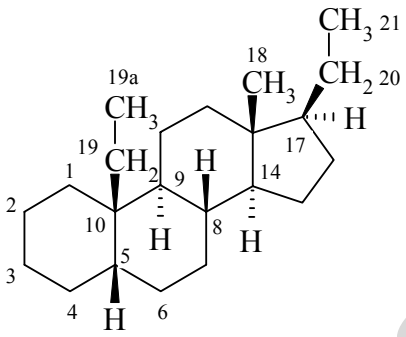
P-101.3.2.2.1 Methylene groups inserted into an atomic connector or into a terminal segment are identified by adding a letter 'a', 'b', etc., to the locant of the highest numbered skeletal atom of the atomic connector or terminal portion consistent with the location of double bonds remaining in the structure. If there are equivalent atomic connectors, the highest atomic connector is chosen, and the methylene group is inserted after the highest number skeletal atom in that connector.

Addition of acyclic side chains or extension of terminal portions of a side chain already attached to the stereoparent hydride may also be done by principles of substitutive nomenclature. The added substituent(s) are numbered as described above for 'homo' atoms.

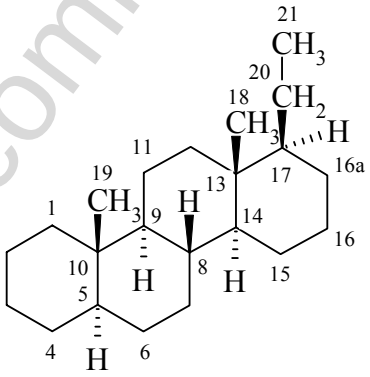
Examples:



pregnane (PIN)
(fundamental parent structure)



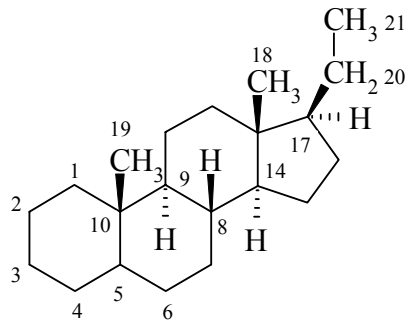
19a-homo-5β-pregnane (PIN)
(not 19-methyl-5α-pregnane;
extension of side chain not
allowed)



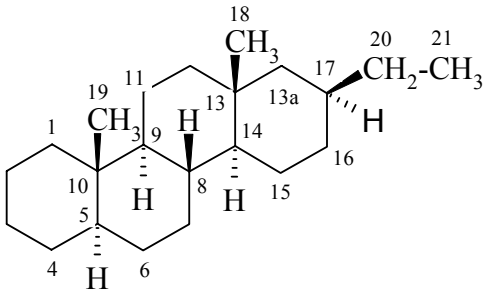
16a-homo-5α-pregnane (PIN)

P-101.3.2.2.2 Methylene groups inserted into a bond connector are identified by citing both locants of the skeletal atoms terminating the bond connector enclosing the second (higher) number in parentheses, followed by a letter ‘a’, ‘b’, etc. according to the number of methylene groups.

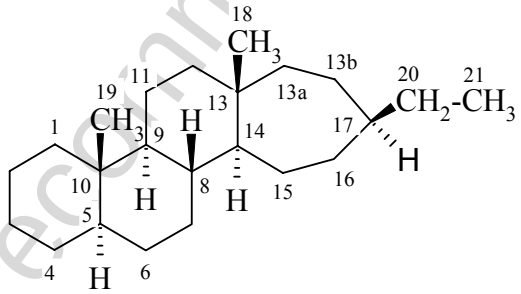
Examples:



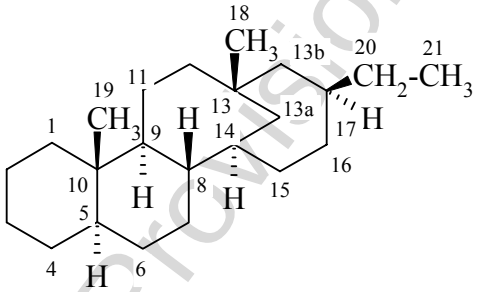
pregnane (PIN)
(fundamental parent structure)



13(17)a-homo-5α-pregnane (PIN)



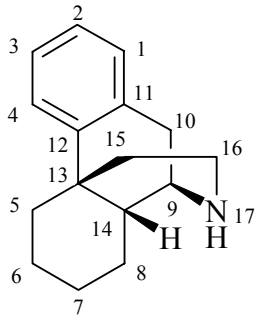
13(17)a,13(17)b-dihomo-5α-pregnane (PIN)
(has been called *D*(17a,17b)dihomo-5α-pregnane)



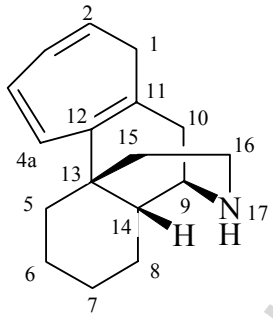
13(14)a,13(17)b-dihomo-5α-pregnane (PIN)

P-101.3.2.2.3 The insertion of a methylene group into a mancude ring or ring system (that contains the maximum number of noncumulative double bonds) or into a system of conjugated double bonds may create a saturated ring position that is described by 'indicated hydrogen' (see P-14.6). The position of the methylene group is prescribed by P-103.2.2.2, even though the saturated ring position may be elsewhere in the unsaturated ring system as denoted by the appropriate locant for the indicated hydrogen; this is a change for names of the homoporphyrins (see ref. 42, Rule TP-5.1). Two tautomeric forms, **A** and **B**, are represented below and specifically numbered and named.

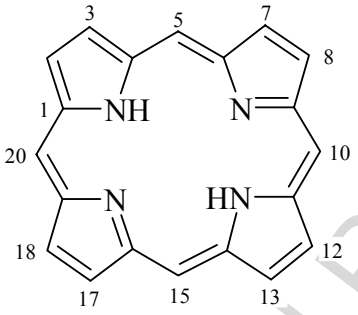
Examples:



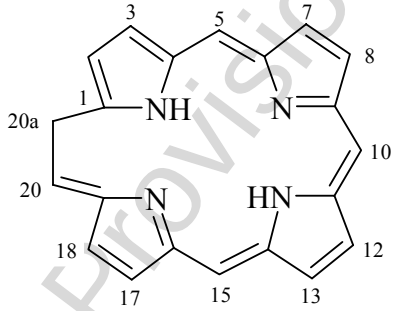
morphinan(PIN)
(fundamental parent structure)



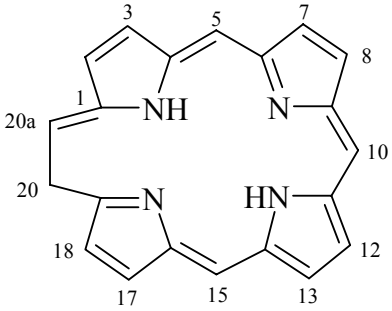
1*H*-4a-homomorphinan (PIN)



porphyrin (PIN)
(fundamental parent structure)



A 20a*H*-20a-homoporphyrin (PIN)



B 20*H*-20a-homoporphyrin (PIN)

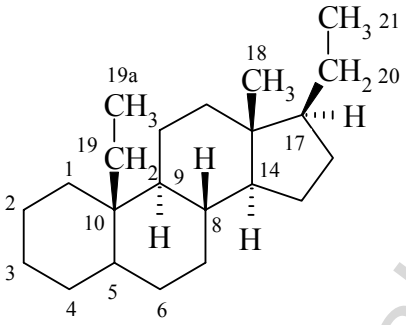
P-101.3.3 Bond formation

The creation of an additional ring by a conjunctive operation (see P-13.5.3) by means of a direct link between any two atoms of a parent structure is described by the nondetachable prefix 'cyclo' (not italicized) preceded by the locants of the skeletal atoms so connected. When necessary, the stereochemical configuration created by the new bond is denoted by α, β, or ξ descriptors in accord with P-101.2.5.1.1, or by describing the configuration of the hydrogen atom in accordance with P-101.2.5.1.2..

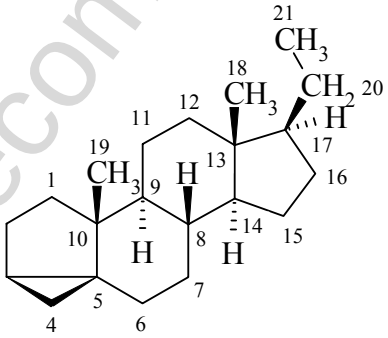
Stereochemical configurations of the parent structure are retained. New stereochemistry of the ring atoms having one hydrogen atom still present is indicated by the α/β stereodescriptors as described in P-105.5.1.2, or, if necessary, by the sequence rule method (*R*/*S*). The projection of the hydrogen atom below (α) or above (β) the reference plane of the ring system is indicated by the appropriate symbol and a capital italic letter *H* following the locant of the ring atom in the structure, all enclosed in parentheses, and cited before the appropriate prefix, in this case 'homo' (see P-103.5 for the prefix 'abeo'). This method of citation differs from that used in the Steroid Rules (ref. 39, Rule 3S-7.5).

For preferred IUPAC names, modifications by the prefix 'cyclo' are limited to fundamental structures having two or more rings, and ring formation is limited to the addition of a maximum of two more rings to a fundamental structure. When the use of a fundamental structure is not allowed, preferred IUPAC names are substitutive names systematically formed, with CIP stereodescriptors to describe configurations.

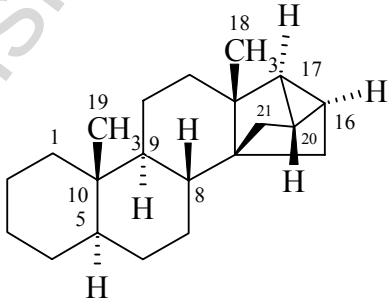
Example:



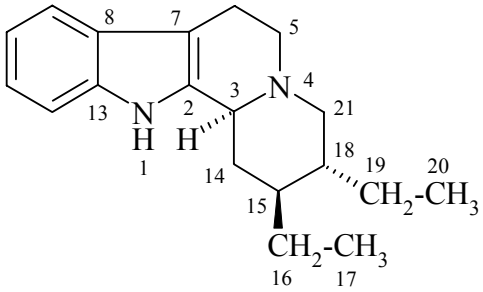
pregnane (PIN)
(fundamental parent structure)



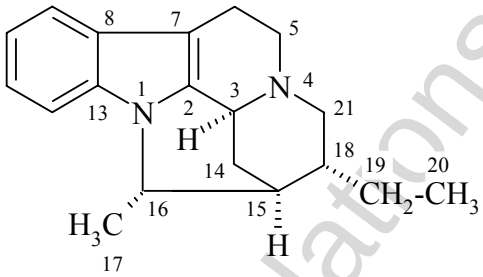
3α,5-cyclo-5α-pregnane (PIN)



(20*S*)-14,21:16β,20-dicyclo-5α,14β-pregnane (PIN)



corynan (PIN)
(fundamental parent structure)



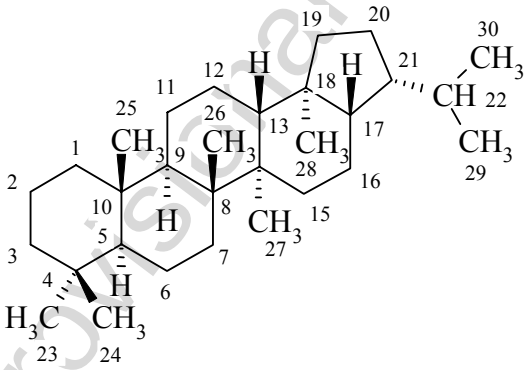
(16β*H*)-1,16-cyclocorynan (PIN)

P-101.3.4 Bond cleavage

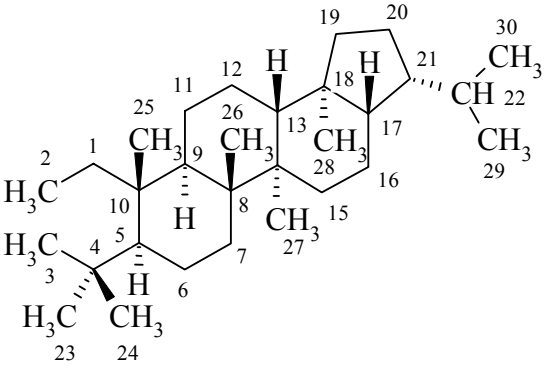
P-101.3.4.1 Cleavage of a ring bond (saturated or unsaturated) with the addition of appropriate number of hydrogen atoms at each new terminal group thus created is indicated by the prefix 'seco' (not italicized) and the locants of the cleaved bond. The original numbering is retained.

For preferred IUPAC names, modifications by the prefix 'seco' are limited to fundamental structures having three or more rings, and.ring opening is limited to one 'seco' operation. When the use of a fundamental structure is not allowed, preferred IUPAC names are substitutive names systematically formed, with CIP stereodescriptors to describe configurations.

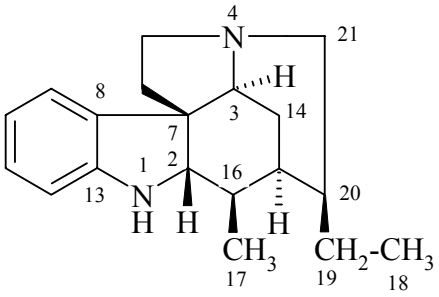
Example:



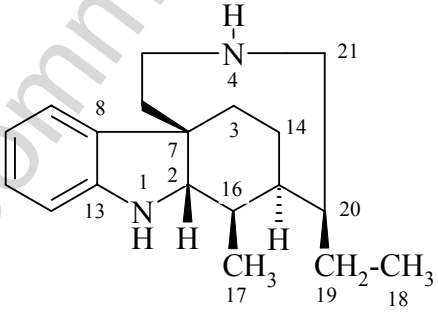
hopane (PIN)
(fundamental parent structure)



2,3-secohopane (PIN)



curan (PIN)
(fundamental parent structure)

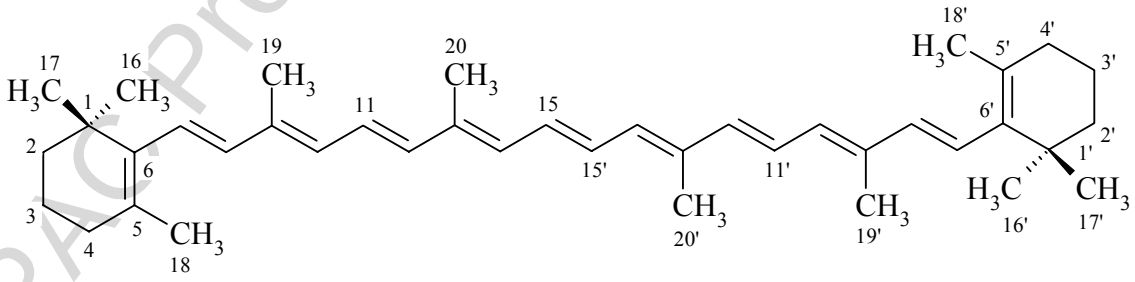


3,4-secocuran (PIN)

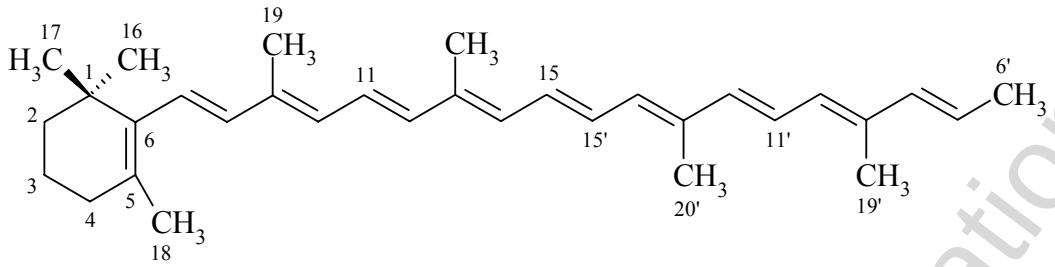
P-101.3.4.2 The unitalicized prefix 'apo' preceded by a locant is used to indicate removal of all of a side chain of a parent structure beyond the skeletal atom corresponding to that locant. Removal of more than one side chains is indicated by using the numerical multiplying prefixes 'di', 'tri', etc. preceded by the appropriate locants. Numbering of the skeletal atoms in the parent structure is retained in the resulting fragment.

This procedure is used only in carotenoid nomenclature (see ref. 40, Rule Carotenoid 10).

Example:



β-carotene (PIN)
(fundamental parent structure)



6′-apo-β-carotene (PIN)

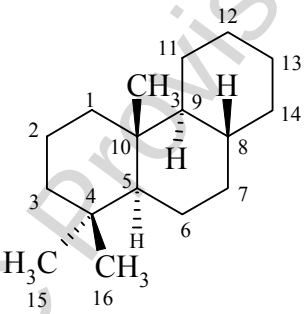
P-101.3.5 Bond migration

Parent structures that are not simple derivatives of accepted parents, but may be considered to arise from such parents by bond migration of one or more bonds, may be named by the following method. This method is described here, although it is not recommended to be used to generate preferred IUPAC names.

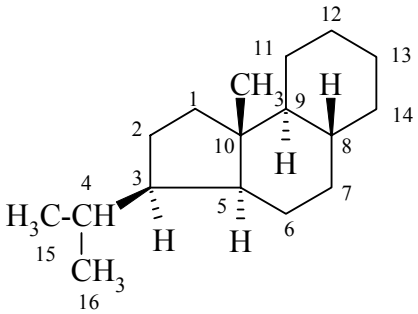
P-101.3.5.1 The nondetachable prefix '*x*(*y*→*z*)-abeo-', designates the migration from one end of a single bond from its original position in a parent structure to another position. In the prefix, '*x*' is the locant of the stationary, i.e. unchanged, end of the migrating bond; '*y*' is the locant of the position of the moving end of the migrating bond in the parent structure; and '*z*' is the locant of the position of the moving end in the final structure. The numbering of the initial structure is retained.

Previously the prefix 'abeo' was italicized (ref 1, F-4.7; ref 2, R-1.2.7.1). For consistency with the other modifying prefixes it is now recommended that a roman font be used. The 'abeo' nomenclature described in this rule is permissive, not compulsory. It is most suitable for use in discussions on reaction mechanisms and biogenesis. In general systematic names, or names assigned by the 'homo', 'nor', 'cyclo', 'seco' method, are preferred.

Example:



podocarpane (PIN)
(fundamental parent structure)



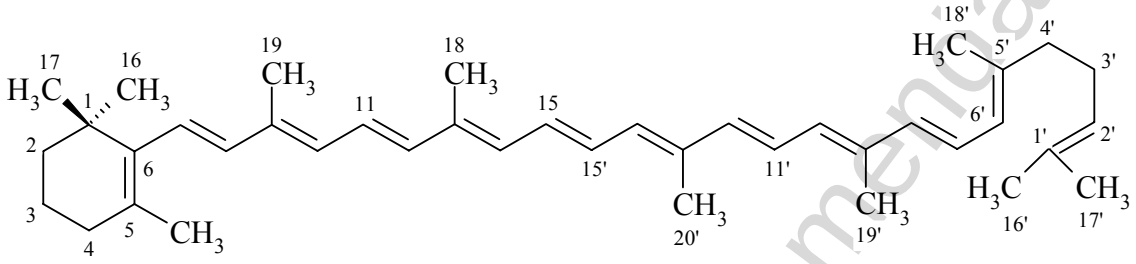
(3α*H*)-5(4→3)-abeo-podocarpane
3,5-cyclo-4,5-seco-3β-podocarpane (PIN)

P-101.3.5.2 The italic prefix '*retro*' preceded by a pair of locants is used to indicate a shift, by one position, of all single and double bonds of a conjugated polyene system delineated by the pair of locants; the conjugated polyene system cannot be part of a system of maximum number of

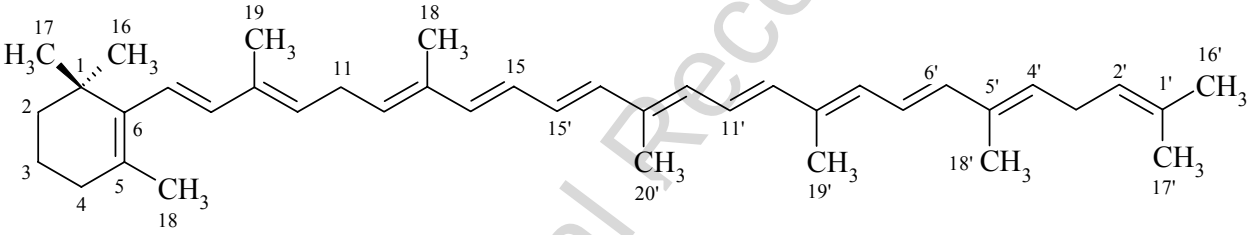
noncumulative double bonds in a ring or ring system The first locant is the skeletal atom that has lost a hydrogen atom and the second locant the one that has gained a hydrogen atom.

The descriptor '*retro*' is used in this manner only in carotenoid nomenclature (see ref 27, Carotenoid Rule 9)

Example:



β,ψ-carotene (PIN)
(fundamental parent structure)

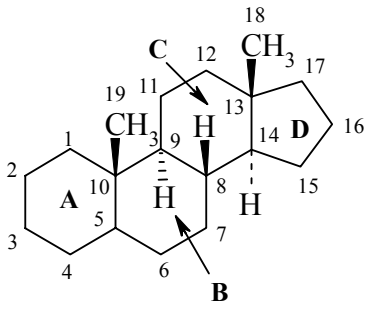


4′,11-*retro*-β,ψ-carotene (PIN)

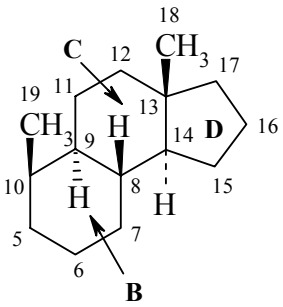
P-101.3.6 Removal of a terminal ring.

The removal of a terminal ring from a parent structure with the addition of an appropriate number of hydrogen atoms at each junction with the adjacent ring is indicated by the nondetachable prefix 'des' followed by the capital italic letter of the ring removed (see P-103.2). This is the only time that the capital letters are now used to identify rings in a parent structure. Stereochemistry implied by the name of the stereoparent structure remains the same, unless otherwise specified. Numbering of skeletal atoms of the parent structure is retained in the modified structure. This use of 'des' is restricted to steroids.

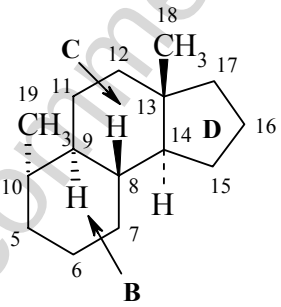
Example:



androstane (PIN)
(fundamental parent structure)



des-*A*-androstane (PIN)



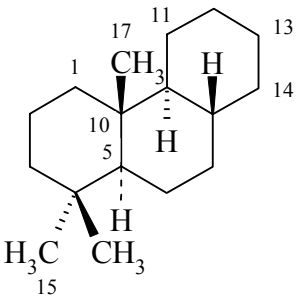
des-*A*-10α-androstane(PIN)

P-101.3.7 Combination of the prefixes ‘cyclo’, ‘seco’, ‘apo’, ‘homo’, and ‘nor’ for generating preferred IUPAC names

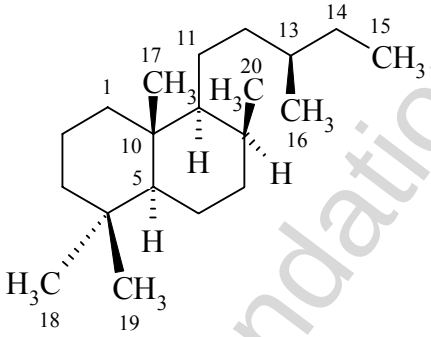
The modifications to a fundamental parent structure prescribed by the prefixes in the preceding recommendations (P-101.3.1 through P-101.3.4) may be combined to effect even more drastic changes in structure. The operation indicated by each prefix ‘cyclo’, ‘seco’, “apo’ ‘homo’, and ‘nor’ is applied to the fundamental parent structure sequentially as one ‘advances backward’, i.e. moves from right to left, from the name of the fundamental parent structure.

P-101.3.7.1 When different combinations of prefixes ‘cyclo’, ‘seco’, ‘apo’, ‘homo’, ‘nor’ can be used to effect the same transformation in structure, the combination of choice must express the fewest number of operations. Both detachable (e.g. alkyl) and nondetachable (e.g. homo or nor) prefixes are considered as modifications but detachable prefixes are preferred. Dihomo, dinor, etc., are counted as two modifications each (see ref 26, 3S-6.3). When the number of operations is the same, the combination of homo/nor is preferred to cyclo/seco; choice between other combinations expressing the same number of operations is based on earlier alphabetical order of the prefixes.

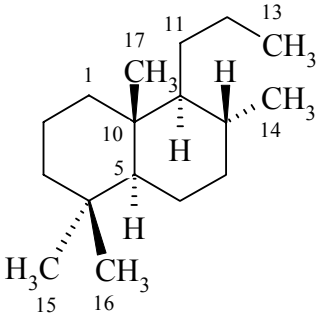
Examples:



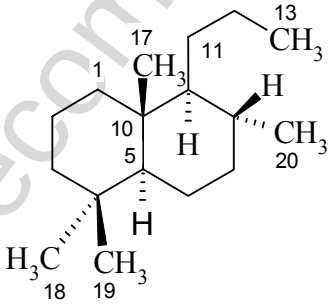
podocarpane (PIN)
(fundamental parent structure)



labdane (PIN)
(fundamental parent structure)



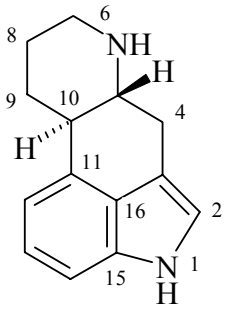
I



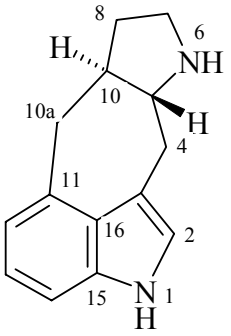
II

13,14-secopodocarpane **(I)** (PIN) not 8α-14,15,16-trinorlabdane **(II)**

Guide to name formation: Podocarpane may be used to generate the 'seco' compound by one operation; the same compound may be obtained from labdane in three operations. Thus, 13,14-secopodocarpane, the name resulting from the minimal use of structure modifying prefixes is the preferred IUPAC name. Compounds I and II are compared, irrespective of the fact that three operations are not allowed to generate preferred IUPAC names.



ergoline (PIN)
(fundamental parent structure)



10(11)a-homo-9-norergoline (PIN)
[not 5,9-cyclo-5,10-secoergoline, nor (9*H*)-5(10→9)-abeoergoline]

P-101.3.7.2 The order of citation of combinations of structure modifying prefixes must avoid improper use of the prefixes as defined above or impossible situations when the corresponding operations are carried out in the manner prescribed above.

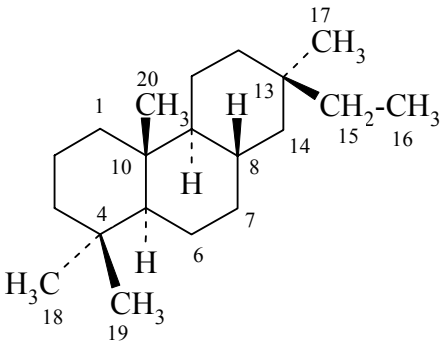
P-101.3.7.3 After satisfying P-101.3.7.1 and P-101.3.7.2, the nondetachable prefixes that indicate bond rearrangements ('cyclo' and 'seco') are cited, followed by those that indicate addition or removal of skeletal atoms ('homo' and 'nor'). If more than one of any of these operations is needed, they are cited in alphabetical order before the name of the fundamental parent structure. Multiplying prefixes denoting multiple operations of the same kind do not affect the order.

Preferred IUPAC names result from modifications by two operations only involving the prefixes 'cyclo', 'seco', 'homo', and 'nor'. In general nomenclature, more than two operations are allowed. Names are formed by citing the bond rearrangement prefixes 'cyclo' and 'seco', in that order from left to right, followed by removal/addition prefixes 'homo' and 'nor, in that order from left to right, at the front of the name of the parent structure. Schematically this order may be shown as follows:

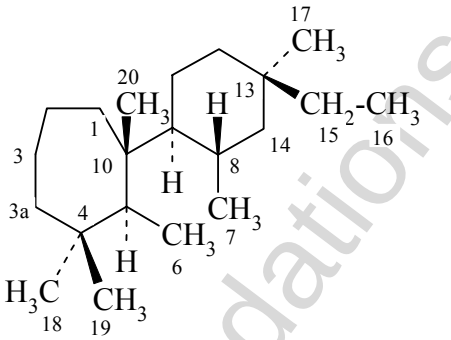
bond rearrangement	addition/removal of skeletal atoms	parent structure
cyclo, seco	apo, homo, nor	

Names in which the order of prefixes is cyclo/seco/homo/nor are preferred to those denoted by the alphabetical order apo/cyclo/homo/nor/seco for the four prefixes.

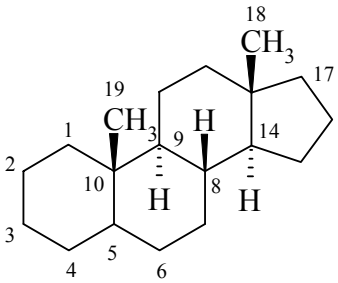
Examples:



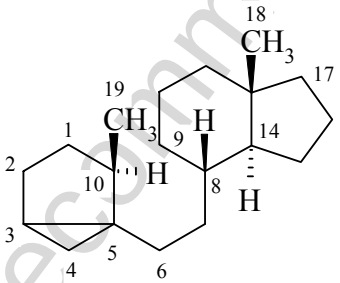
pimarane (PIN)
(fundamental parent structure)



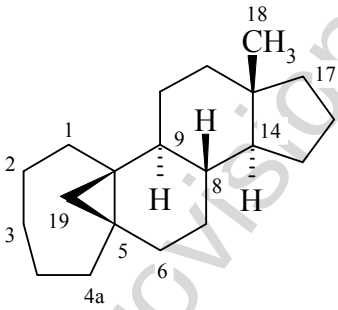
6,7-seco-3a-homopimarane (PIN)
3a-homo-6,7-secopimarane



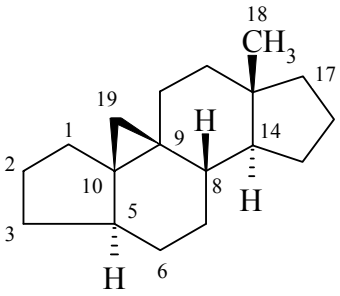
androstane (PIN)
(fundamental structure)



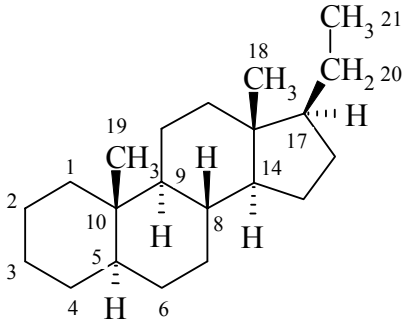
3α,5-cyclo-9,10-secoandrostane (PIN)



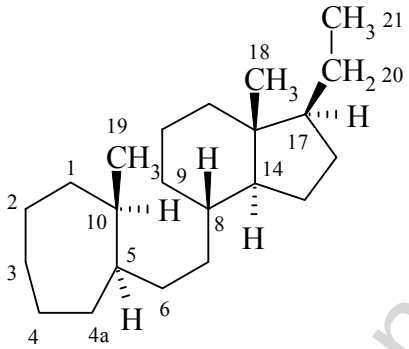
5β,19-cyclo-4a-homo-5β-androstane (PIN)



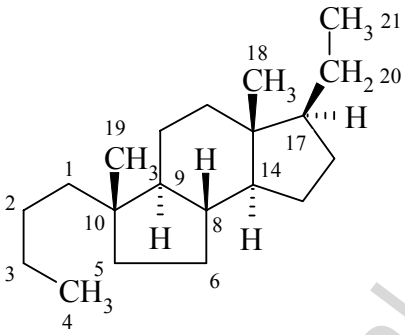
9β,19-cyclo-4-nor-5α,9β-androstane (PIN)



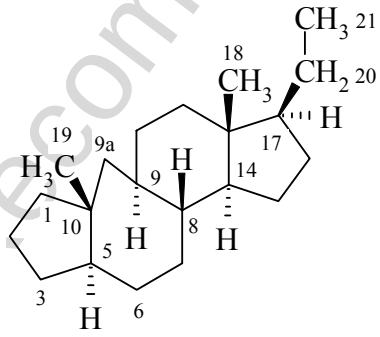
5-α-pregnane (PIN)
(fundamental structure)



9,10-seco-4a-homo-5α-pregnane (PIN)



4,5-seco-7-nor-5α-pregnane (PIN)



9a-homo-4-nor-5α-pregnane

P-101.4. Replacement of skeletal atoms

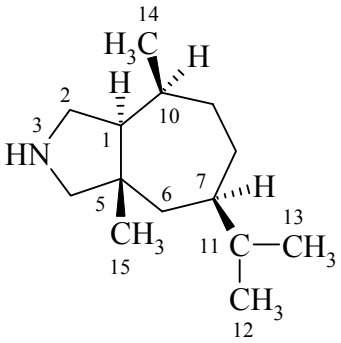
P-101.4.1 General methodology

The principles of skeletal replacement ('a') nomenclature, as described in P-15.4 to modify parent structures, are applied to replace carbon skeletal atoms by heteroatoms, such as O, S, N. Contrary to the recommended alphabetical order for citation in names in Revised Section F (ref. 9), the seniority order of the 'a' prefixes prescribed in P-15.4 is recommended for skeletal replacement. In addition to the methodology used to generate systematic names, skeletal replacement ('a') replacement is also used to replace heteroatoms in parent structures by carbon atoms and by other heteroatoms.

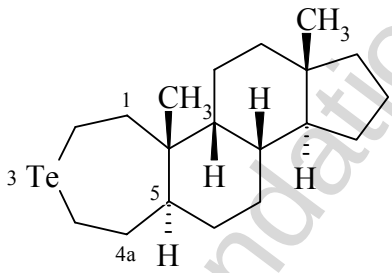
P-101.4.2 Skeletal replacement of carbon atoms by heteroatoms

Heteroatoms are denoted by 'a' prefixes that are cited before nondetachable prefixes expressing skeletal modifications in fundamental parent structures, each with a locant to indicate its position; the fixed numbering of the parent structure is maintained. Skeletal modifications, if any, must be completed before skeletal replacement ('a') nomenclature can be applied.

Examples:



3-azaambrosane (PIN)

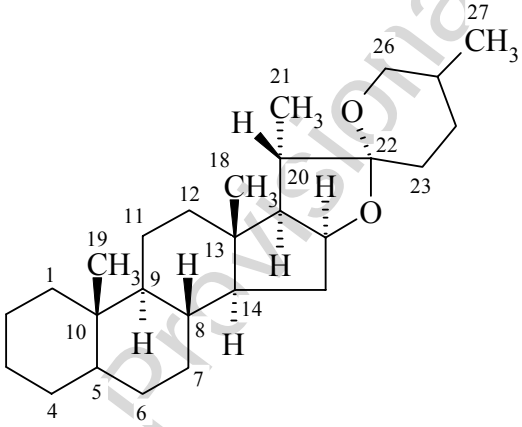


3-tellura-4a-homo-5α-androstane (PIN)

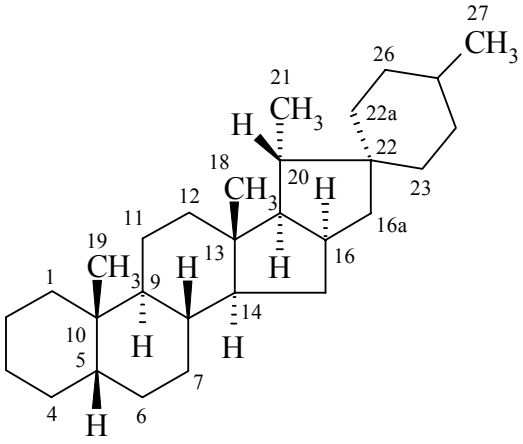
P-101.4.3 Replacement of skeletal heteroatoms by carbon atoms

Replacement of a heteroatom in a parent structure by a carbon atom is indicated by the 'a' prefix 'carba'. The original numbering is maintained. If the heteroatom is not numbered, the replacing carbon atom is numbered by affixing the letter 'a' to the locant of the immediately adjacent lower numbered skeletal atom. If the immediately adjacent lower numbered skeletal atom is a 'homo' atom, the letter 'b', 'c', etc., as appropriate, is used. Stereochemical configuration at the new carbon skeletal atom is described by methods for specifying additional stereochemistry (see P-101.2.6.1).

Examples:



spirostan (PIN)
(fundamental parent structure)

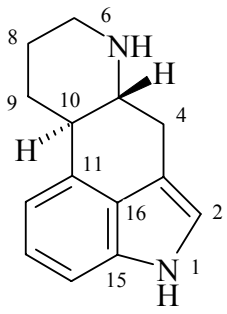


16a,22a-dicarba-5β-spirostan (PIN)

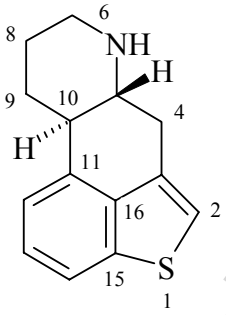
P-101.4.4 Replacement of heteroatoms by other heteroatoms

Replacement of a heteroatom in a stereoparent hydride by another heteroatom is denoted by the appropriate 'a' prefix and locant.

Example:



ergoline (PIN)
(fundamental parent structure)

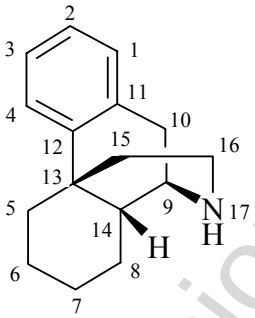


1-thiaergoline (PIN)

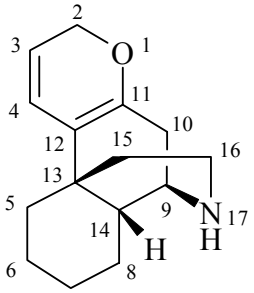
P-101.4.4 Indicated hydrogen

When the replacement of a skeletal atom in a portion of the structure of a parent structure that is mancude (contains the maximum number of noncumulative double bonds) or an extended conjugated system of double bonds results in the creation of a saturated skeletal position, that position is indicated by the symbolism of indicated hydrogen (see P-14.6).

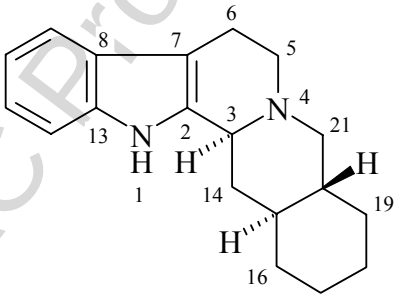
Examples:



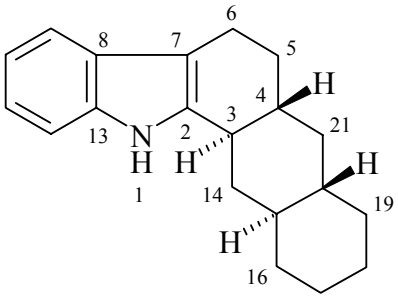
morphinan (PIN)
(fundamental parent structure)



2*H*-1-oxamorphinan (PIN)



yohimban (PIN)
(fundamental parent structure)



(4β*H*)-4-carbayohimban (PIN)

P-101.5. Additional rings and ring systems

Three types of rings and ring systems can be incorporated into parent structures:

P-101.5.1 Mancude rings and ring systems incorporated by fusion nomenclature
P-101.5.2 Rings and ring systems incorporated by bridged fused ring nomenclature
P-101.5.3 Rings and ring systems incorporated by spiro nomenclature

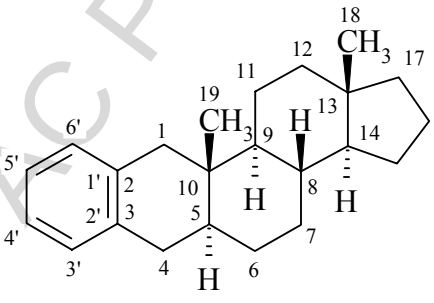
The methods, in certain cases adapted to parent structures, used for the construction of systematic names and described in Chapters 1 to 8 above, are applied.

P-101.5.1 Mancude rings and ring systems incorporated by fusion nomenclature

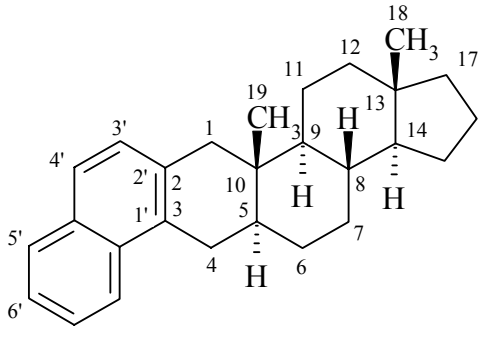
The parent structure as a component is used in fusion nomenclature in its normal state of hydrogenation. Accordingly, a double bond is not cited at the fusion site just because the other component contains the maximum of noncumulative double bonds. Furthermore, contrary to the rules prescribed in P-25, a fundamental parent structure is always chosen as the principal component and the attached component must be a mancude ring or ring system.

P-101.5.1.1 A ring or ring system considered as a mancude parent hydride in accordance with the rules prescribed in Chapter 2, carbocyclic or heterocyclic, fused to a parent structure is described by its fusion prefix name (see P-25) preceding the name of the parent structure. The skeletal atoms of the parent structure involved in the fusion are identified by plain (unprimed) locants and not by italicized letters '*a*', '*b*', etc.; the skeletal atoms of the mancude component, involved in the fusion, are identified by primed locant numbers. The position of the fusion is indicated by a fusion descriptor, including two sets of locants; the first cited set is that of the attached component, the second set relates to the principal component, the fundamental parent structure; the two sets are separated by a colon, enclosed in brackets, and cited between the two components. Where there is a choice, the locant for the mancude attached component are as low as possible and are cited in the same direction of numbering as for the parent structure. Terminal vowels, 'o' or 'a', in the name of the prefix are not elided when followed by a vowel, as prescribed for normal fusion nomenclature in P-25.3.1.2 (this is a change from previous recommendations).

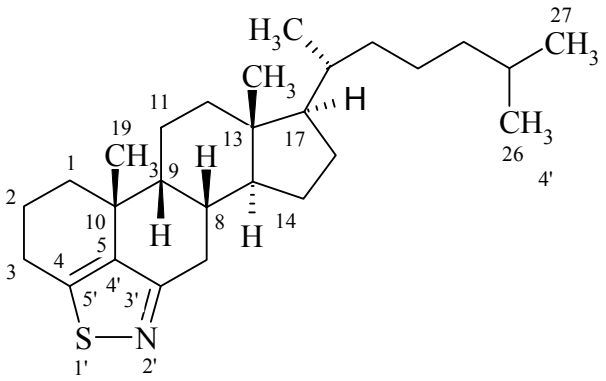
Examples:



benzo[2,3]-5α-androstane (PIN)
(locants 1′,2′ are omitted)



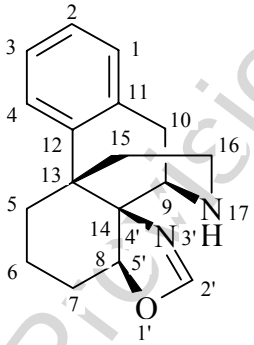
naphtho[2′,1′:2,3]-5α-androstane (PIN)



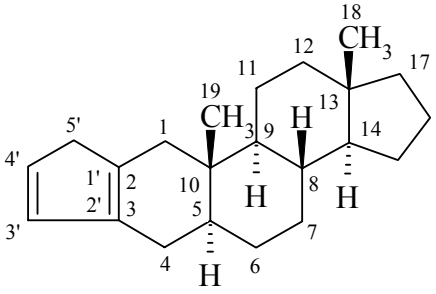
[1,2]thiazolo[5′,4′,3′:4,5,6]cholestane (PIN)
(not isothiazolo[5′,4′,3′:4,5,6]cholestane; the name isothiazole is no longer recommended as a fusion component, see P-25.3.2.1.3)

P-101.5.1.2 The attached component fused to a parent structure is a mancude compound (it contains the maximum number of noncumulative double bonds). Saturated positions on such components, including the fusion sites, that have at least one hydrogen atom are specified by an indicated hydrogen. They are also specified by a descriptor composed of the locant, followed by the configuration descriptor α or β and finally by the indicated hydrogen symbolism (see P-15.6), placed in parentheses at the front of the name as stereodescriptors are. Locants of the attached component are used to identify the position of the indicated hydrogen, but locants (unprimed) of the stereoparent hydride are used, if there is a choice between primed and unprimed locants.

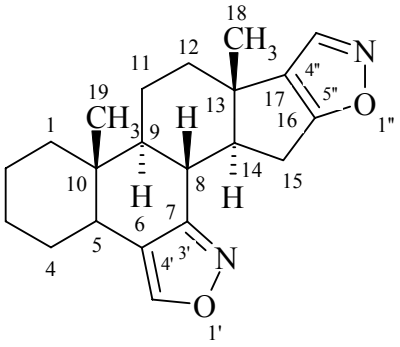
Examples:



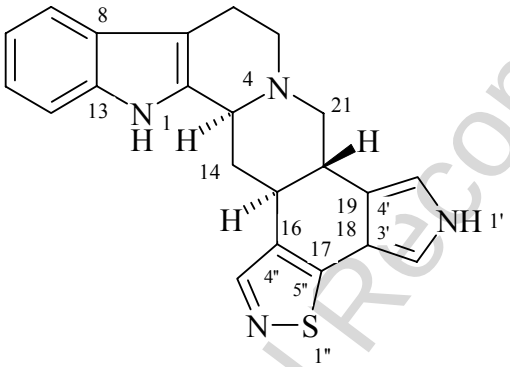
(8α*H*)-[1,3]oxazolo[5′,4′:8,14]morphinan (PIN)
(not 8α*H*-oxazolo[5′,4′:8,14]morphinan; the name oxazole without heteroatom locants is no longer recommended as a fusion component (see P-25.3.2.1.3); an indicated hydrogen atom denoted by the stereodescriptor (8α*H*) is necessary to complete the name)



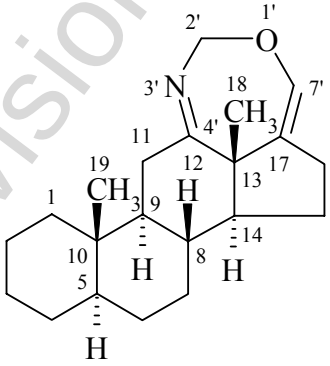
5′*H*-cyclopenta[2,3]-5α-androstane (PIN)



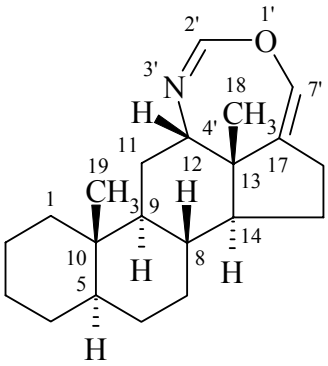
bis[1,2]oxazolo[4′,3′:6,7;5″,4″:16,17]-5α-androstane



1′*H*-pyrrolo[3′,4′:18,19][1,2]thiazolo[4″,5″:16,17]yohimban



2*H*′-[1,3]oxazepino[4′,5′,6′:12,13,17]-5α-androstane (PIN)



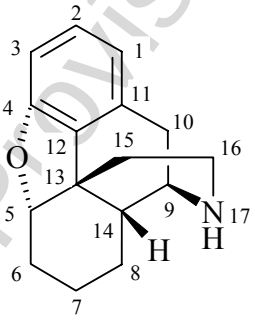
(12β*H*)-12*H*-[1,3]oxazepino[4′,5′,6′:12,13,17]-5α-androstane (PIN)

P-101.5.2 Rings and ring systems incorporated by bridges

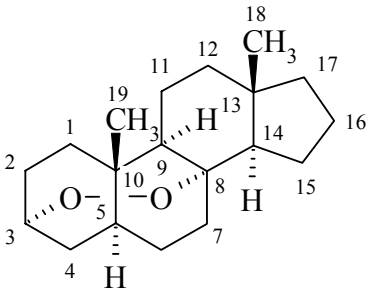
Atomic bridges added to parent structures may be described by the methods used in fusion nomenclature for bridged fused ring systems. The names of the bridges are those prescribed in P-25.4. This method is often used with heteroatom bridges. In fact, this method is often more useful than fusion procedures described in P-101.5.1 for describing certain types of heterocyclic rings fused to a parent structure, for instance 'epoxy' to denote a bridge rather than 'oxireno' to denote the ring fused as an attached component. The use of atomic bridges is preferred to fusion nomenclature to connect two nonadjacent atoms in a fundamental parent structure [epoxides and thioepoxides are exceptions as they can be named substitutively (see P-63.5)]. The prefixes used to denote bridges are nondetachable; they are cited in a name in front of the prefixes used to denote skeletal modifications, preceded by appropriate locants.

When a choice is possible, names formed by using fusion nomenclature are preferred IUPAC names.

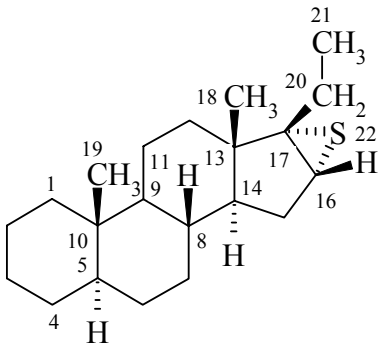
Examples:



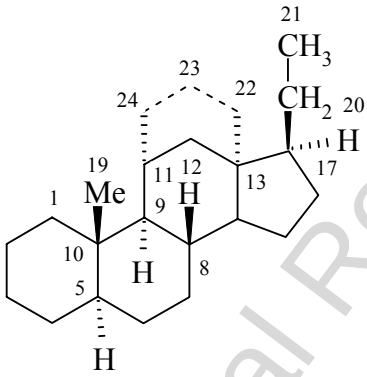
4,5α-epoxymorphinan(PIN)



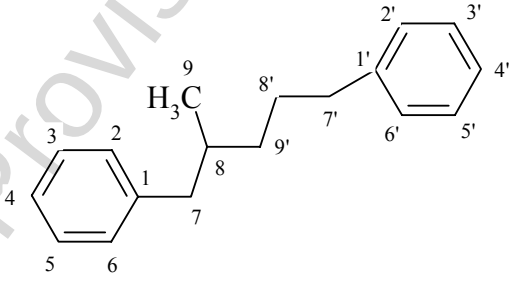
3α,8-epidioxy-5α,8α-androstane (PIN)



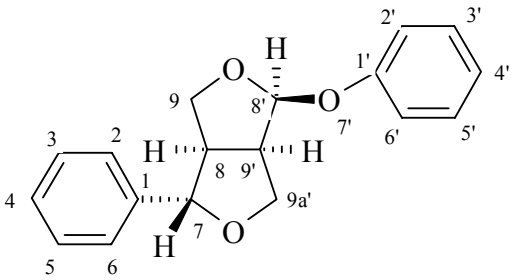
(16β*H*)-thiireno[16,17]-5α-pregnane (PIN, fusion name)
16α,17-epithio-5α-pregnane



11α,13-propano-18-nor-5α,13α-pregnane (PIN)
(not 11β,18-cyclo-12a,12b-dihomo-5α-pregnane;
nor 11α,18b-cyclo-18a,18b-dihomo-5α,13α-pregnane)



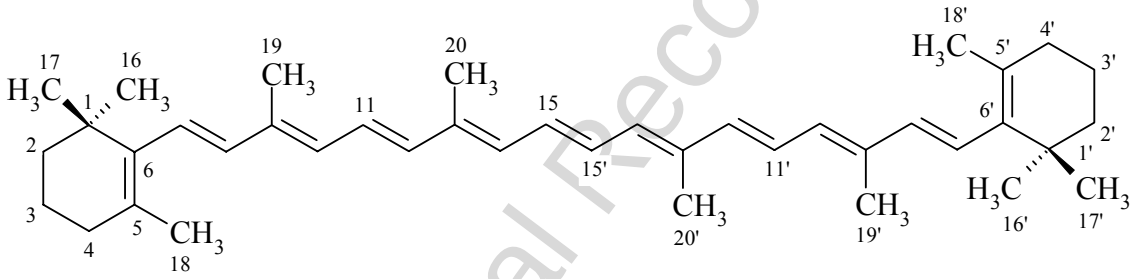
8,9′-neolignan
(fundamental parent structure)



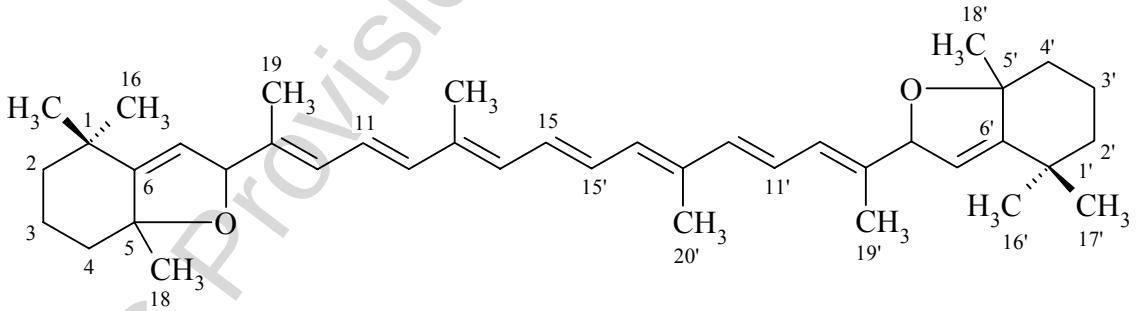
(7α,8α,8′β,9′α)-7,9a′:8′,9-diepoxy-7-oxa-9a′-homo-8,9′-neolignan

Contrary to the recommendations for systematic nomenclature of organic compounds, in carotenoid nomenclature, the bridge named 'epoxy' is considered detachable and the hydro/dehydro prefixes nondetachable. The name of the following bridged β,β-carotene is written in conformity with the rules of carotenoid nomenclature (ref. 40, Carotenoid rule 7.3).

Examples:



β,β-carotene (PIN)
(fundamental parent structure)

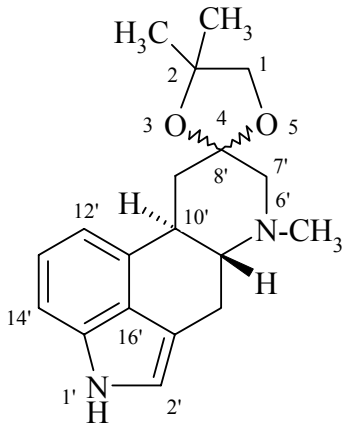


5,8:5′,8′-diepoxy-5,8,5′8′-tetrahydro-β,β-carotene (PIN)

P-101.5.3 Rings and ring systems incorporated by spiro nomenclature

Spiro compounds are named as prescribed in P-24.5 for monospiro compounds having at least one polycyclic component.

Example:



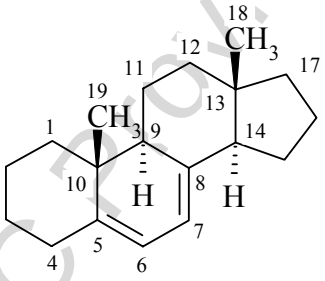
(4ξ)-2,2,6′-trimethylspiro[1,3-dioxolane-4,8′-ergoline] (PIN)

P-101.6 Modification of the degree of hydrogenation

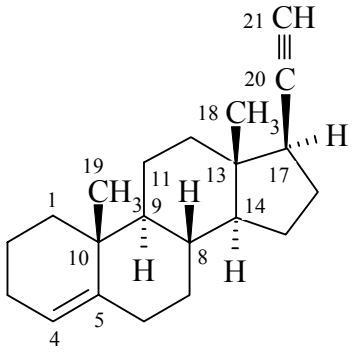
The general principles and rules for modifying the degree of hydrogenation of parent hydrides prescribed in Section P-31 are applied to parent structures. The endings 'ene' and 'yne' (see P-31.1) and the prefixes 'hydro' (see P-31.2) and 'dehydro' (see P-31.3) are used, depending on the subtractive or additive operation required.

P-101.6.1 Unsaturation in a compound whose parent structure is fully saturated or in the portion of a parent structure that is otherwise fully saturated and whose name ends in 'an', 'ane', or 'anine' is indicated by changing 'an' or 'ane' to 'ene' or 'yne' and by adding numerical mulytiplying prefixes as prescribed in P-31. Locants are placed immediately before the part of the name to which they relate.

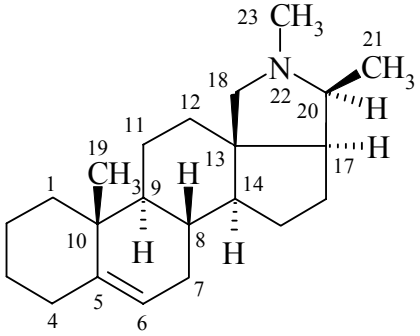
Examples



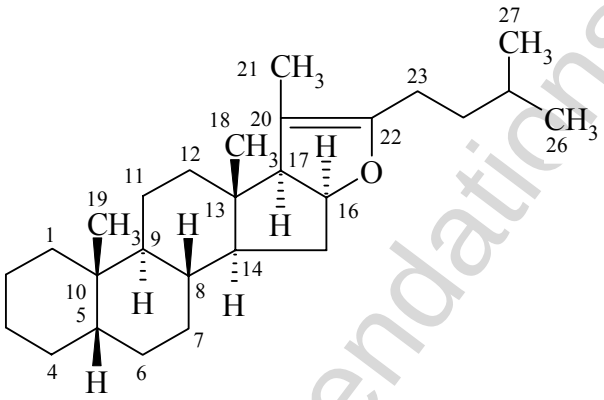
androsta-5,7-diene (PIN)



pregn-4-en-20-yne (PIN)



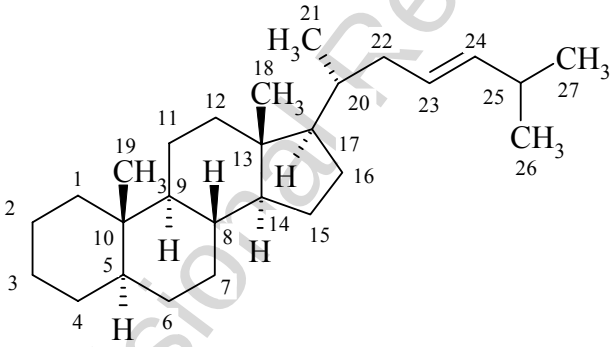
con-5-enine (PIN)



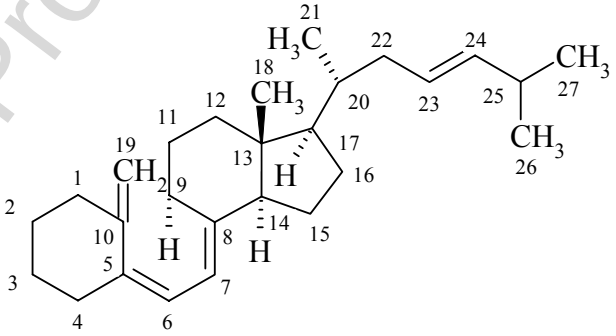
5β-furost-20(22)-ene (PIN)

P-101.6.2 The descriptors '*E*' and '*Z*', preceded by appropriate locants, are used to describe modified or additional stereochemical configurations for double bonds. The stereodescriptors '*cis*' and 'trans' are used in carotenoid nomenclature (ref. 40) and retinoid nomenclature (ref. 48).

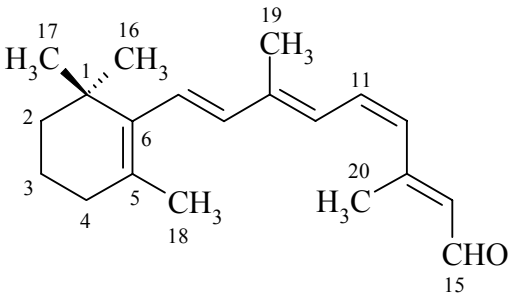
Example:



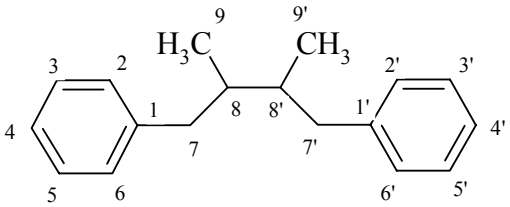
(23*E*)-5α-cholest-23-ene (PIN)



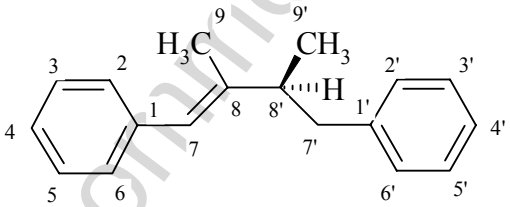
(5*Z*,7*E*)-9,10-secocholesta-5,7,10(19)-triene



11-*cis*-retinal (PIN)



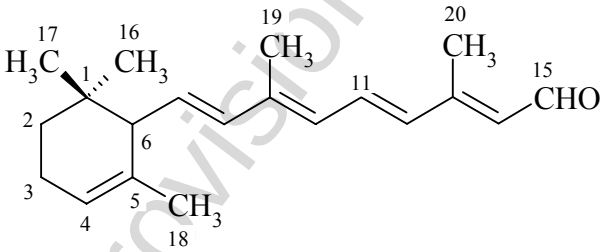
lignan (PIN)
(fundamental parent structure)



(7*E*,8′*S*)-lign-7-ene

P-101.6.3 The prefix *all* is used in front of stereodescriptors to indicate that all configurations are identical. This prefix is used only in the nomenclature of natural products, for example, *all-trans* to denote the fact that in retinal all double bonds are *trans*.

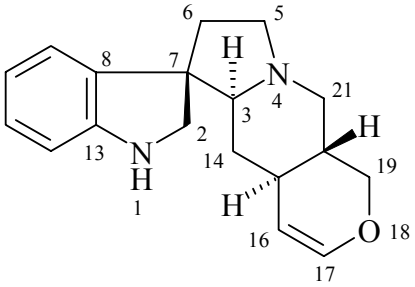
Example:



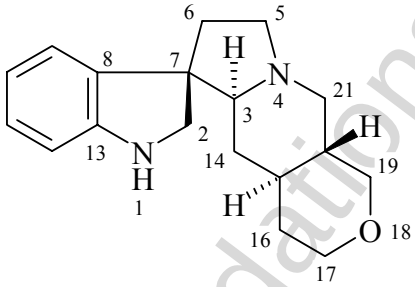
all-trans-retinal (PIN)

P-101.6.4 Saturation of double bonds in a parent structure whose name implies the presence of isolated double bonds and/or systems of conjugated double bonds is described by the prefix ‘hydro’, itself preceded by the locants of the saturated positions. The ‘hydro’ prefix is detachable and always cited immediately in front of the fundamental parent structure (see P-31.2).

Example:



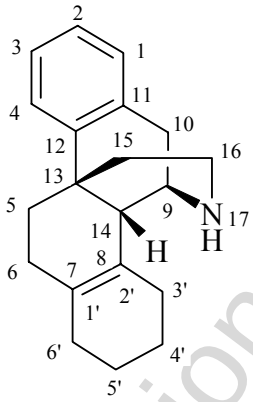
formosanan (PIN)
(fundamental parent structure)



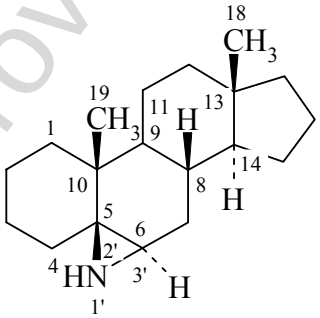
16,17-dihydroformosanan (PIN)

P-101.6.5 Saturated, or partially saturated, carbocyclic and heterocyclic ring components fused to a parent structure are named using 'hydro' prefixes. When there is a choice between primed and unprimed locants, the unprimed locants are used.

Examples:

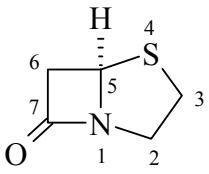


3′,4′,5′,6′-tetrahydrobenzo[7,8]morphinan (PIN)

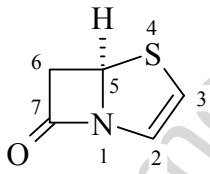


(6α*H*)-1′,6-dihydroazirino[2',3':5,6]-5β-androstane (PIN)
(for the symbol (6α*H*), see P-101.5.1.2)

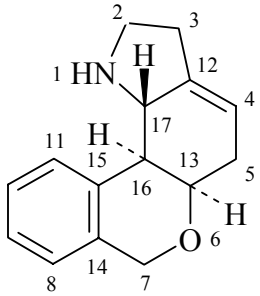
P-101.6.6 The introduction of unsaturation additional to any unsaturation implied in a parent structure whose name does not end in 'an', 'ane', or 'anine', the conversion of an implied double bond into a triple bond, and the introduction of an additional double bond with rearrangement of an implied double bond, are denoted by the prefix 'dehydro', itself prefixed by a numerical multiplying term equal to the number of hydrogen atoms removed and the appropriate locants. The 'dehydro' prefix is detachable and always cited at the front of the fundamental parent structure, after any detachable alphabetized prefixes, when present.
Examples:



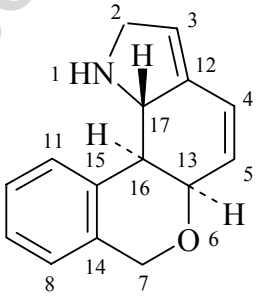
penam (PIN)
(fundamental parent structure)



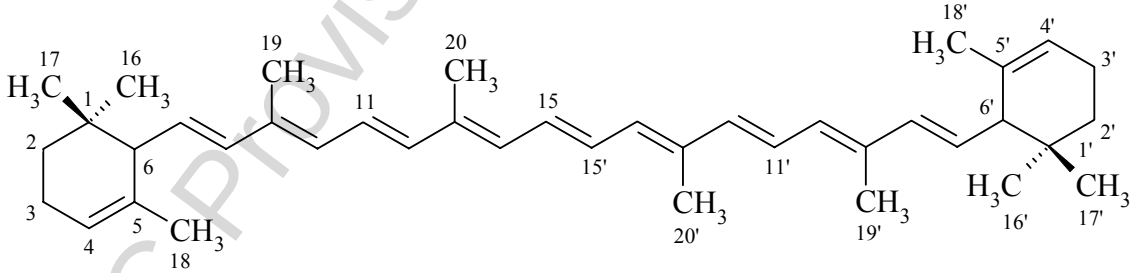
2,3-didehydropenam (PIN)



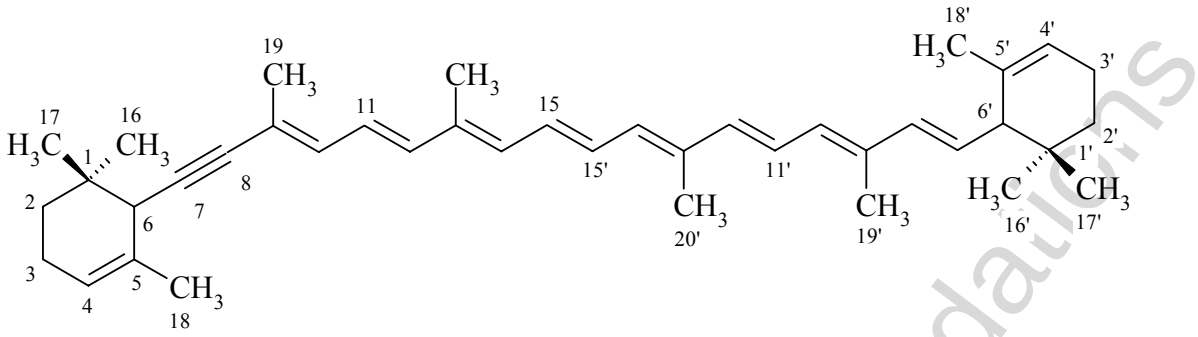
lycorenan (PIN)
(fundamental parent structure)



3,5-didehydrolycorenan (PIN)



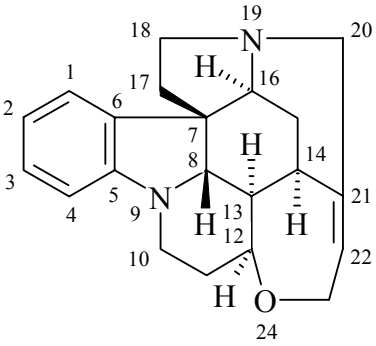
ε,ε-carotene (PIN)
(fundamental parent structure)



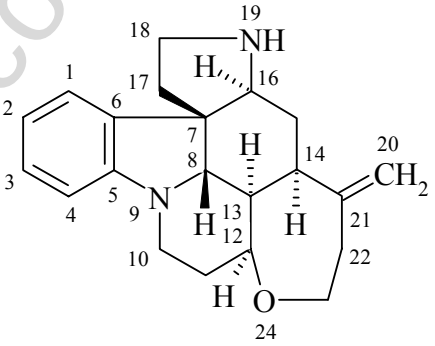
7,8-didehydro-ε,ε-carotene (PIN)

P-101.6.7 Rearrangement of double bonds may be indicated by a combination of 'hydro' and 'dehydro' prefixes. The 'dehydro' prefix is cited before the 'hydro' prefix, in accordance with the alphanumerical order.

Example:



strychnidine (PIN)
(fundamental parent structure)



20,21-didehydro-21,22-dihydro-19,20-secostrychnidine (PIN)

P-101.7 Derivatives of parent structures

Derivatives of parent structures are named according to principles, rules and conventions described in Chapters 1 to 8.

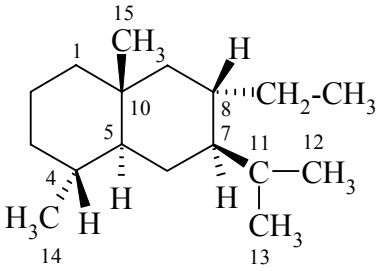
P-101.7.1 The suffixes and prefixes of the nomenclature of organic compounds are used in the prescribed manner to name atoms and groups that are considered to substitute for hydrogen atoms of parent structures. Preferred IUPAC names are based on the largest unmodified fundamental parent structures when a choice is possible. The stereodescriptors α, β and ξ are used to describe the configuration; they are cited in front of the prefix or suffix, preceded by the appropriate locant. Substitutive names so constructed are preferred to those that are formed by functional class nomenclature, except for some cyclic functional classes.

Substitution on rings and substitution on terminal segments are considered separately.

P-101.7.1.1 Substitution by alkyl groups

Organyl groups such as aryl groups and alkyl groups are introduced by substitutive nomenclature. This rule is implemented to introduce a methyl group in androstane in position 17β; the alternative method of subtracting a methylene group by using the nondetachable prefix 'nor' is not recommended.

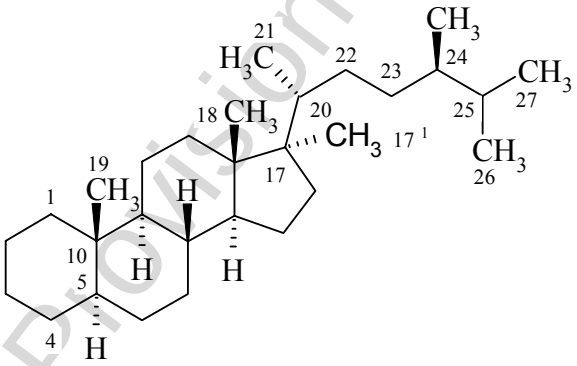
Example:



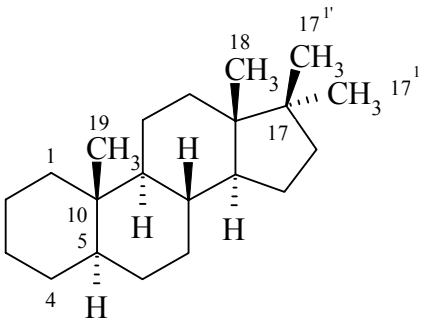
8α-ethyl-5α-eudesmane (PIN)

Rule 3S-2.7 in ref. 39 describes the methodology to name steroids with a side chain as part of the parent carbocycle and an alkyl substituent at C-17. Rule 3S-2.7 also describes the methodology to name steroids with two alkyl substituents at C-17. This methodology is applicable to any fundamental parent structure described in Section P-101.. Locants with superscript numbers are intended for the identification of the atoms, e.g. in ^{13}C-nmr assignments, not as locants for further substitution.

Examples:



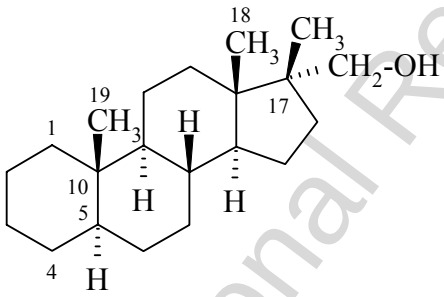
17-methyl-5α-campestane (PIN)
(the additional methyl group in position 17 is numbered 17^1;
other atoms are numbered as usual)



17,17-dimethyl-5α-androstane (PIN)
(both additional methyl groups are numbererd 17^1;
the β-methyl group is primed)

The principles, rules and conventions of substitutive nomenclature are used when a characteristic group cited as a suffix is present on an alkyl substituent group added to a fundamental parent structure.

Example:

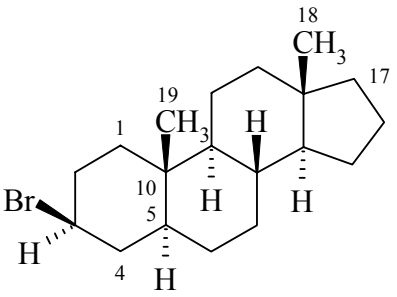


(17β-methyl-5α-androstan-17α-yl)methanol (PIN)
(not 21-nor-5α-pregnane)

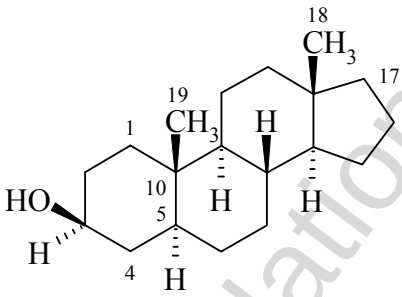
P-101.7.1.2 Substitution on rings

Suffixes are used in accordance with the seniority order of suffixes, considering the cyclic nature of the parent hydride. Detachable prefixes are cited in alphanumerical order. The endings 'ene' and 'yne' are cited in the normal way; the 'hydro-dehydro' prefixes are detachable but cited last among detachable prefixes.

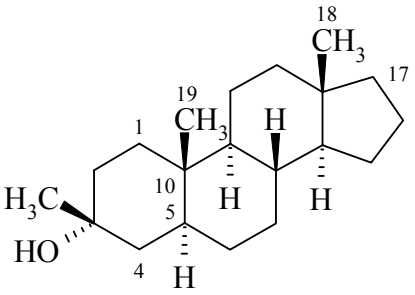
Examples:



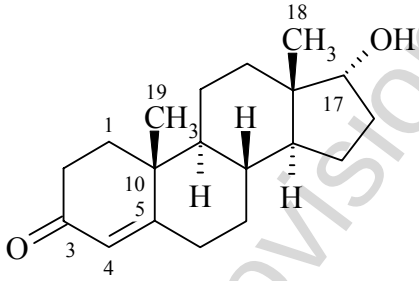
3β-bromo-5α-androstane (PIN)
5α-androstan-3β-yl bromide



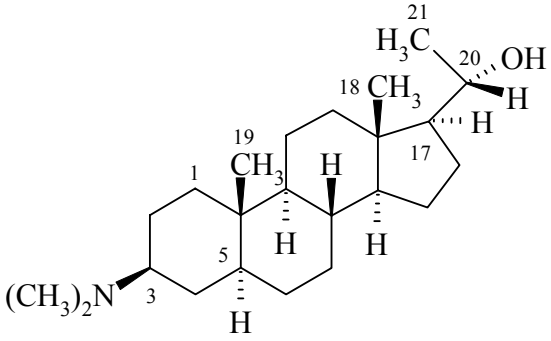
5β-androstan-3β-ol (PIN)



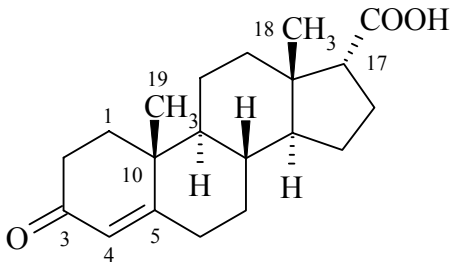
3β-methyl-5α-androstan-3α-ol (PIN)



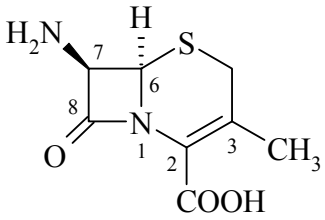
17α-hydroxyandrost-4-en-3-one (PIN)



(20*S*)-3β-(dimethylamino)-5α-pregnan-20-ol (PIN)



3-oxoandrost-4-ene-17α-carboxylic acid (PIN)
(not 21-nor-5α-pregnane-20-oic acid; the preferred name involves the fewest number of operations, see 101.3.7.1)

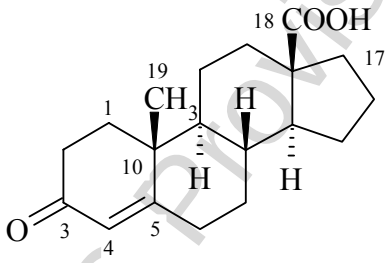


7β-amino-3-methyl-2,3-didehydrocepham-2-carboxylic acid (PIN)

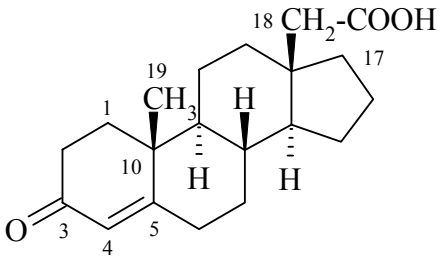
P-101.7.1.3 Substitution on terminal segments

Substitution on terminal segments by prefixes and suffixes expressing characteristic groups is recommended, even when a carbon atom included in a characteristic group is present. Lengthening a terminal segment by the addition of two methylene groups is allowed and denoted by the use of the prefix 'dihomo'. Further lengthening is possible, but alkyl groups must be used, as an exception to the rule related to seniority of the longest chain.

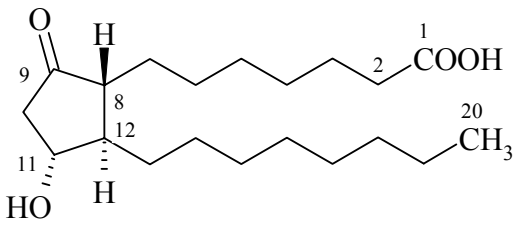
Examples:



3-oxoandrost-4-en-18-oic acid (PIN)



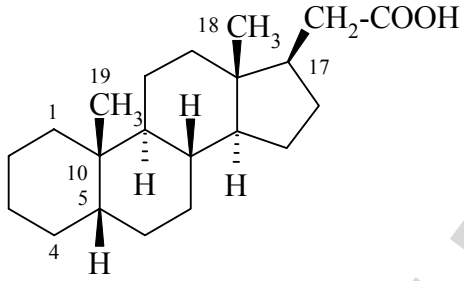
3-oxoandrost-4-ene-18-carboxylic acid (PIN)



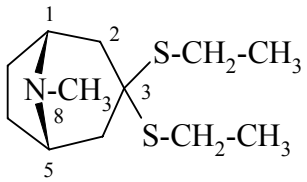
11α-hydroxy-9-oxoprostan-1-oic acid (PIN)

P-101.7.2 Modifications to principal characteristic groups such as esters (see P-65.4.3.2), acetals (see P-66.6.5), etc. are named by the usual methods described in Chapter 6. Cyclic modifications, such as lactones, cyclic acetals, etc. are named preferably as such rather than as fused or spiro ring systems, even if these names are functional class names (see also P-101.7.4).

Examples:



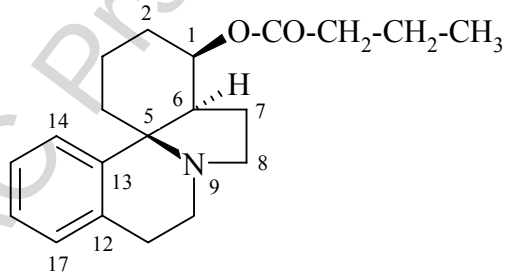
methyl 5β-androstane-17β-carboxylate (PIN)



3,3-bis(ethylsulfanyl)tropane (PIN)
tropan-3-one diethyl dithioketal

P-101.7.3 Names of substituent groups derived from parent stereoparent hydride structures are formed, by the general method described in P-29, by adding the suffixes 'yl', 'ylidene', or 'ylidyne', as appropriate, to the name of a parent, with elision of the final letter 'e', if present, before the letter 'y'.

Example:

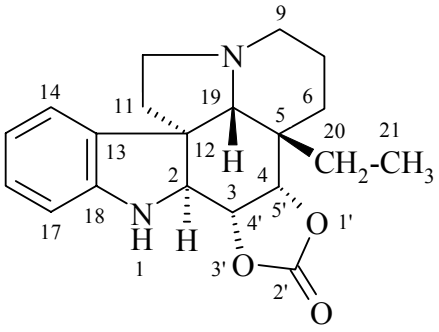


erythrinan-1β-yl butanoate (PIN)

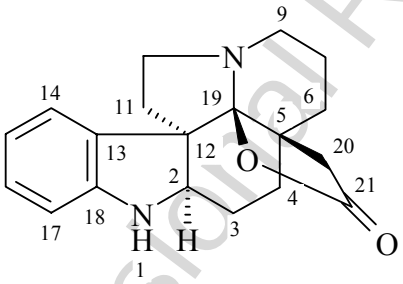
P-101.7.4 Addition of rings denoting functional groups

Rings denoting functional groups are preferably named by the usual methods described for constructing systematic names. Cyclic esters and lactones are named by the general method described for naming esters (see P-65.4.3.2). Names of acetals are formed by using the principles of functional class nomenclature (see P-66.6.5) rather than by fusion nomenclature described in P-101.5. When a choice is possible, a fusion name is preferred.

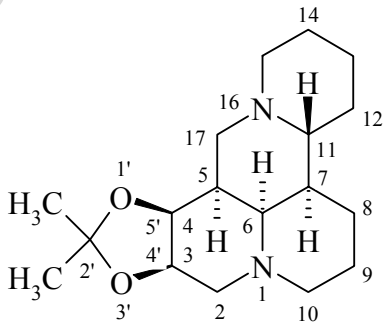
Examples:



(3β*H*,4β*H*)-3,4-dihydro[1,3]dioxolo[4′,5′-3,4]aspidospermidin-2′-one (PIN)
[not aspidospermidine-3α,4α-diyl carbonate(see P-65.4.3.2.6)]



21-noraspidospermidine-20,19-carbolactone (PIN, see P-65.4.3.4.1)
19-hydroxyaspidospermidine-21,19-lactone



(3α*H*,4α*H*)-2,2′-dimethyl-3,4-dihydro[1,3]dioxolo[4′,5′-3,4]matridine (PIN)
[not propan-2-one matridine-3β,4β-diyl acetal (see P-66.6.5)]
acetone matridine-3β,4β-diyl acetal

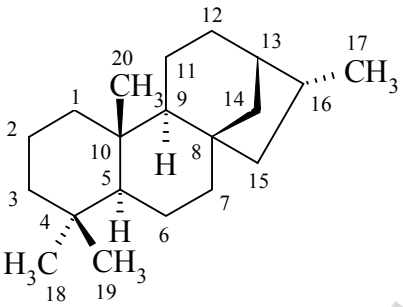
P-101.8. Further aspects of configuration specification

In addition to the specification of the absolute configuration of fundamental and modified parent structures using 'α', 'β', 'ξ', '*R*' and '*S*' stereodescriptors, many other stereochemical features have to be described. The principles, rules and conventions described in Chapter 9 are applied.

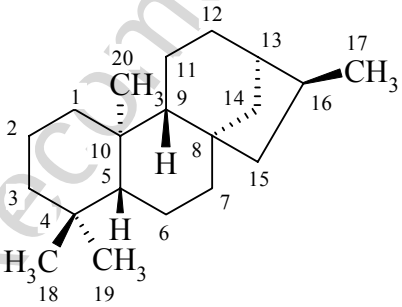
P-101.8.1 Inversion of configuration

Configurational inversion of all chirality centers is indicated by the italicized prefix '*ent*' (a contracted form for '*enantio*') placed at the front of the complete name of the compound. This prefix denotes inversion at all chirality centers (including those due to named substituents) whether these are cited separately or are implied in the name.

Example:



kaurane (PIN)



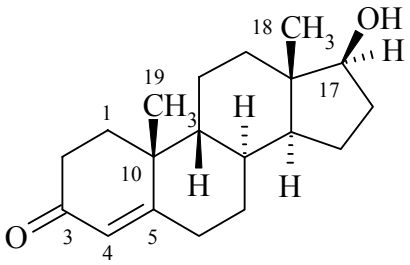
ent-kaurane (PIN)

P-101.8.2 Racemates

Racemates are named by citing the italicized stereodescriptor '*rac*' (an abbreviation for *racemo*) in front of the whole name of the compound including the prefix '*epi*', if present. In the case of a racemic compound, the enantiomeric structure drawn should be that one that shows the lowest numbered chirality center in the α-configuration. This may differ from the usual practice, which is to draw the enantiomeric structure having the same absolute configuration as the naturally occurring substance.

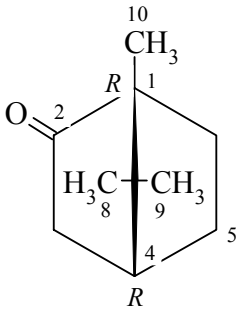
P-101.8.3 Relative configuration

When the relative, but not the absolute configurational relationships among chirality centers are known, the symbols '*R**' and/or '*S**' are used in accordance with Rule P-93.2. Alternatively, enantiomers of known relative, but unknown absolute configuration may be distinguished by the compound stereodescriptor (+)-*rel*- or (-)-*rel*-, where the plus and minus sign refer to the direction of rotation of polarized light at the sodium D line. Hence, the dextrorotatory form of the following structure would be named (+)-*rel*-17β-hydroxy-8α,9β-androst-4-en-3-one.

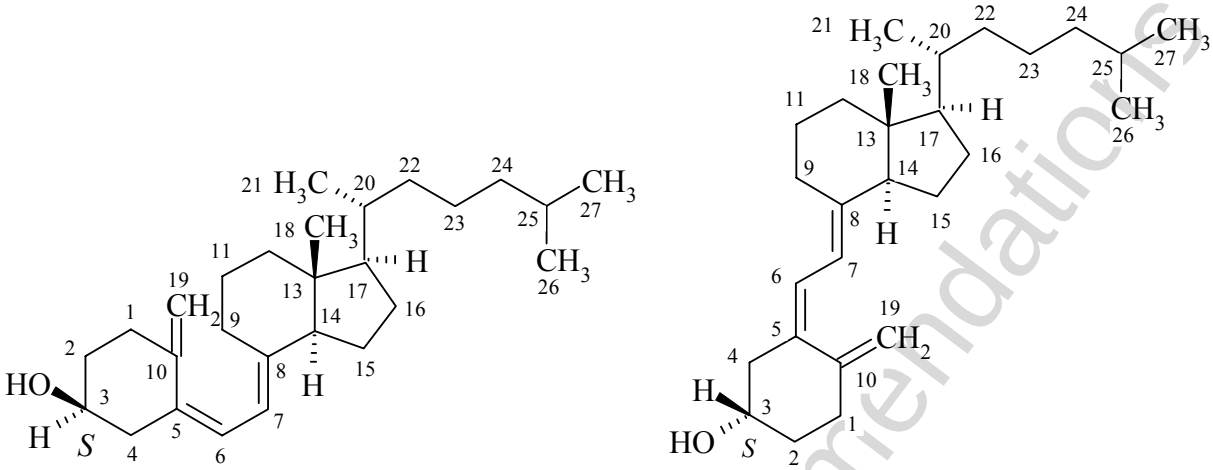


P-101.8.4 The stereodescriptors '*R*' and '*S*' are used to describe the absolute configuration of sterogenic centers for a compound whose parent structure is achiral, for example bornane. They are also used, in place of 'α', 'β', 'ξ', when a ring is opened creating two chiral portions one of which may rotate, as shown for vitamin D.

Example:



(1*R*,4*R*)-bornan-2-one
(+)-camphor
(1*R*,4*R*)-1,7,7-trimethylbicyclo[2.2.1]heptane-2-one (PIN)



(I) ≡ **(II)**

(3*S*,5*Z*,7*E*)-9,10-secocholesta-5,7,10(19)-trien-3-ol
(structures (I) and (II) are two conformations of the same 3-hydroxy derivative)

P-102 Nomenclature of carbohydrates

P-102.0 Introduction

Nomenclature of carbohydrates is based on the concept of parent monosaccharides having retained names that are preferred IUPAC names. These structures and names can be modified to indicate the nature of characteristic groups that are present, such as aldehydes, carboxylic acids, alcohols. They can also be combined to form di-, tri- and oligosaccharides.

The nomenclature has been recently revised (ref. 22). This Section describes the basic concepts of this specific type of nomenclature, in particular the extensive system of symbols and stereodescriptors to indicate the configuration of the many diastereoisomers and enantiomers.

P-102.1 Definitions

P-102.1.1 Carbohydrates
P-102.2.1 Monosaccharides
P-102.3.1 Oligosaccharides
P-102.4.1 Polysaccharides
P-102.5.1 Preferred IUPAC names

P-102.1.1 Carbohydrates

The generic term 'carbohydrates' includes monosaccharides, oligosaccharides and polysaccharides as well as substances derived from monosaccharides by reduction of the carbonyl group (alditols), by oxidation of one or more terminal groups to carboxylic acids, or by replacement of one or more hydroxy group(s) by a hydrogen atom, an amino group, a thiol group or similar heteroatomic groups. It also includes derivatives of these compounds. The term 'sugar' is frequently applied to monosaccharides and lower oligosaccharides.

Cyclitols are generally not regarded as carbohydrates. For the nomenclature of cyclitols, see P-104 and ref. 44.

P-102.1.2 Monosaccharides

Parent monosaccharides are polyhydroxy aldehydes H-[CHOH]$_n$-CHO or polyhydroxy ketones H-[CHOH]$_m$-CO[(CHOH]$_n$-H with three or more carbon atoms.

The generic term 'monosaccharide' (as opposed to oligosaccharide or polysaccharide) denotes a single unit without glycosidic connections to other such units. It includes aldoses, dialdoses, aldoketoses, ketoses, diketoses, as well as deoxy sugars and amino sugars, and their derivatives, provided that the parent compound has a (potential) carbonyl group.

P-102.1.2.1 Aldoses and ketoses

Monosaccharides with an aldehydic carbonyl or potential aldehydic carbonyl group are called aldoses; those with a ketonic carbonyl or potential carbonyl group are called ketoses.

The term 'potential aldehydic group' refers to the hemiacetal group arising from ring closure; the term 'potential ketonic group' refers to the hemiketal structure.

Cyclic hemiacetals or hemiketals of sugars with a five-membered ring (oxolane or tetrahydrofuran) ring are called 'furanoses', those with a six-membered ring (oxane or tetrahydropyran) ring 'pyranoses'.

Dialdoses are monosaccharides containing two (potential) aldehydic groups.

Diketoses are monosaccharides containing two (potential) ketonic groups.

Ketoaldoses are monosaccharides containing one (potential) aldehydic group and one potential ketonic group; this term is preferred to the terms 'aldoketoses' and 'aldosuloses'.

P-102.1.2.2 Deoxy sugars

Monosaccharides in which an alcoholic hydroxyl group has been replaced by a hydrogen atom are called 'deoxy sugars'.

P-102.1.2.3 Amino sugars

Monosaccharides in which an alcoholic hydroxyl group has been replaced by an amino group are called 'amino sugars'. When the hemiacetal group is replaced by an amino group, the compounds are called 'glycosylamines'.

P-102.1.2.4 Glycosides

Glycosides are mixed acetals formally arising by elimination of water between the hemiacetal or hemiketal hydroxyl group of a sugar and a hydroxyl group of a second compound. The bond between the two components is called a 'glycosidic bond'.

P-102.1.3 Oligosaccharides

Oligosaccharides are compounds in which monosaccharides units are joined by glycosidic linkages. According to the number of units, they are called disaccharides, trisaccharides, etc. The maximum number of units is not defined.

P-102.1.4 Polysaccharides

'Polysaccharide' (glycan) is the name given to a macromolecule consisting of a large number of monosaccharide (glycose) residues joined to each other by glycosidic linkages. The term 'poly(glycose)' is not a synonym for polysaccharide (glycan), because it includes monosaccharide residues joined to each other by nonglycoisidic linkages.

P-102.1.5 Preferred IUPAC names

Names of carbohydrates are either trivial or systematic. Many trivial names such as glucose, fructose, etc. are retained and used to describe the corresponding functional parents. This aspect of carbohydrate nomenclature is limited because it applies only to monosaccharides having four to six carbon atoms. A systematic carbohydrate nomenclature has been developed that is applicable to compounds with four or more carbon atoms, and is used extensively by carbohydrate chemists for compounds with more than six carbon atoms, and for unsaturated and branched sugars. Names generated by applying systematic nomenclature are called 'systematic carbohydrate names'. Preferred IUPAC names are either retained names, unmodified or modified, or names generated in accordance with the principles, rules and conventions of substitutive nomenclature developed in Chapters 1 to 8. These systematically formed names are called 'substitutive names' or 'systematic substitutive names' to differentiate them from 'systematic carbohydrate names. The system of preferred IUPAC names is based on the following criteria.

(1) retained names used in carbohydrate nomenclature and systematic substitutive names are selected as preferred IUPAC names. Systematic carbohydrate names may be used in general nomenclature.

(2) preferred IUPAC names for aldoses and ketoses with 4-6 carbon atoms are retained names; the following modifications also generate preferred IUPAC names:

(a) open-chain and ring forms

(b) alditols, aldonic acids, ketoaldonic acids, uronic acids, and aldaric acids

(c) *O*-substitution by alkyl, aryl, carboxylic acyl group prefixes

(d) glycosides and glycosyl substituent prefixes

(e) glycosyl halides and pseudohalides

(f) glycose phosphates and sulfates and other phosphorus or sulfur oxoacid esters

(g) deoxy compounds at the terminal CH_2OH position

(h) halo, halogenoid, and other permanent prefixes plus deoxy at the same position

(i) amino plus deoxy at the same position (*N*-substitution is allowed)

(j) replacement of oxygen atoms by other chalcogen atoms

(k) di- and oligosaccharides

(3) Preferred IUPAC names not generated according to (1) and (2) above are constructed as systematic substitutive names, for example:

(a) aldoses/ketoses having more than six carbon atoms

(b) deoxy at a chiral center (except as in (2)(i), above)

(c) *C*-substitution

(d) anhydro derivatives

(e) unsaturated derivatives

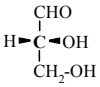
P-102.2 Parent structures and preferred IUPAC names

P-102.2.1 The bases for preferred IUPAC names are the structures of the parent monosaccharides in their acyclic form. Tables 10.2 and 10.3 give retained names for parent aldoses and ketoses with up to six carbon atoms. These retained names are used as preferred IUPAC names when the acyclic aldose or ketose has a carbon chain consisting of 4, 5 or 6 carbon atoms. Preferred IUPAC names of monosaccharides whose carbon skeleton is composed of more than 6 carbon atoms are systematic substitutive names.

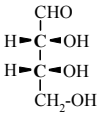
In Table 10.2 structures and retained names of the aldoses (in the aldehydic, acyclic form) with three to six carbon atoms are described. Only the D-forms are shown; the L-forms are the mirror images.

In Table 10.3 structures and retained names of the ketoses (in the ketonic, acyclic form) with three to six carbon atoms are described. Only the D-forms are shown; the L-forms are the mirror images.

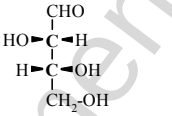
Table 10.2 Retained names (with recommended three-letter abbreviations in parentheses) and structures (in the aldehydic acyclic form) of the aldoses with three to six carbon atoms



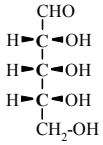
(*R*)-2,3-dihydroxypropanal (PIN)
D-glyceraldehyde
D-*glycero*



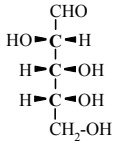
D-erythrose (PIN)
D-*erythro*



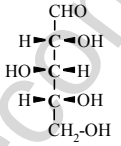
D-threose (PIN)
D-*threo*



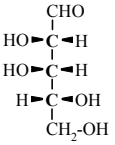
D-ribose (PIN)
D-*ribo*



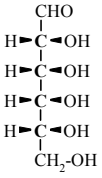
D-arabinose (PIN)
D-*arabino*



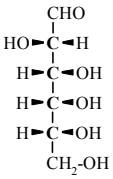
D-xylose (PIN)
D-*xylo*



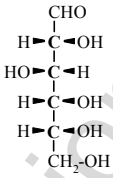
D-lyxose (PIN)
D-*lyxo*



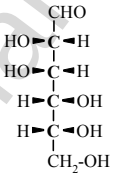
D-allose (PIN)
D-*allo*



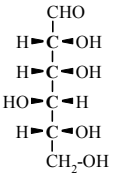
D-altrose (PIN)
D-*altro*



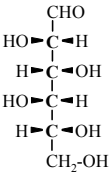
D-glucose (PIN)
D-*gluco*



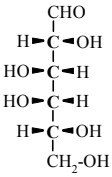
D-mannose (PIN)
D-*manno*



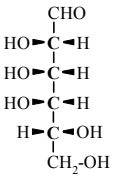
D-gulose (PIN)
D-*gulo*



D-idose (PIN)
D-*ido*

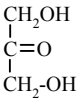


D-galactose (PIN)
D-*galacto*

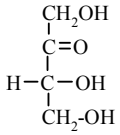


D-talose (PIN)
D-*talo*

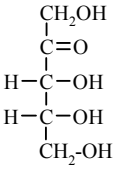
Table 10.3 Structures, with systematic and trivial names, of the 2-ketoses with three to six carbon atoms



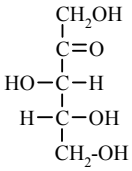
1,3-dihydroxypropan-2-one (PIN)
1,3-dihydroxyacetone
glycerone



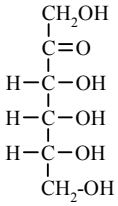
D-erythrulose (PIN)



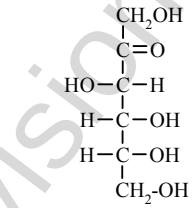
D-ribulose (PIN)



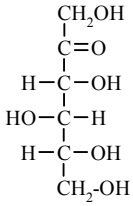
D-xylulose (PIN)



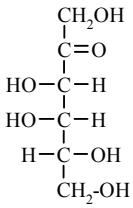
D-psicose (PIN)
D-*ribo*-hex-2-ulose



D-fructose (PIN)
D-*arabino*-hex-2-ulose



D-sorbose (PIN)
D-*xylo*-hex-2-ulose



D-tagatose (PIN)
D-*lyxo*-hex-2-ulose

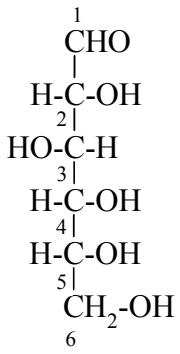
P-102.2.2 Numbering parent structures

The carbon atoms of a monosaccharide are numbered consecutively in such a way that:

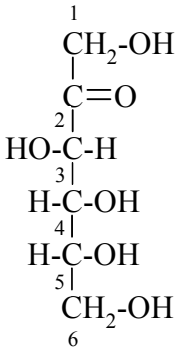
(1) a (potential) aldehyde group receives the locant 1 (even if a more senior function is present);

(2) the most senior of other functional groups expressed in the suffix receives the lowest possible locant, i.e carboxylic acid (derivatives) > (potential) ketonic carbonyl groups.

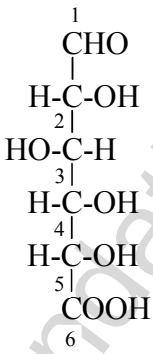
Examples:



D-glucose (PIN)



D-fructose (PIN)

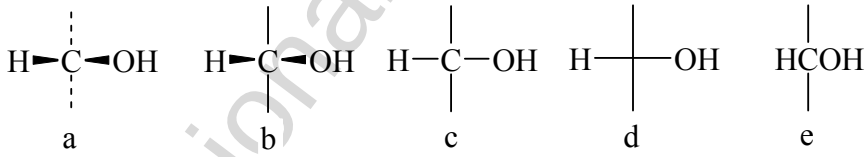


D-glucuronic acid (PIN)

P-102.3 Configurational symbolism

P-102.3.1 The Fischer projection of the acyclic form

In this representation of a monosaccharide, the carbon chain is written vertically with the lowest numbered carbon at the top, as indicated in P-102.2.2. To define the configuration, each carbon atom is considered in turn and placed in the plane of the paper. Neighboring carbon atoms are below, and the H atoms and OH groups are above the plane of the paper. Various representations 'b, c, d, e, and f' of a carbon atom in a monosaccharide in the Fischer projection are as follows (structure 'a' is a tridimensional representation; the real Fischer projection is 'd'); representation 'e' is commonly used in this Section:

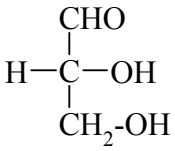


a b c d e

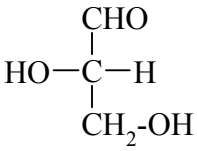
P-102.3.2 The stereodescriptors 'D' and 'L'

The simplest aldose is glyceraldehyde. It contains one center of chirality and occurs therefore in two enantiomeric forms, called D-glyceraldehyde and L-glyceraldehyde; these are represented by the Fischer projection formulas given below. It is known that these projections correspond to the absolute configurations. The configurational stereodescriptors 'D' and 'L' must be written in small capital letters and linked by a hyphen to the name of the sugar.

Preferred IUPAC names are systematic substitutive names; the configuration is described by the preferred CIP stereodescriptors '*R*' and '*S*'.



D-glyceraldehyde
(*R*)-2,3-dihydroxypropanal (PIN)

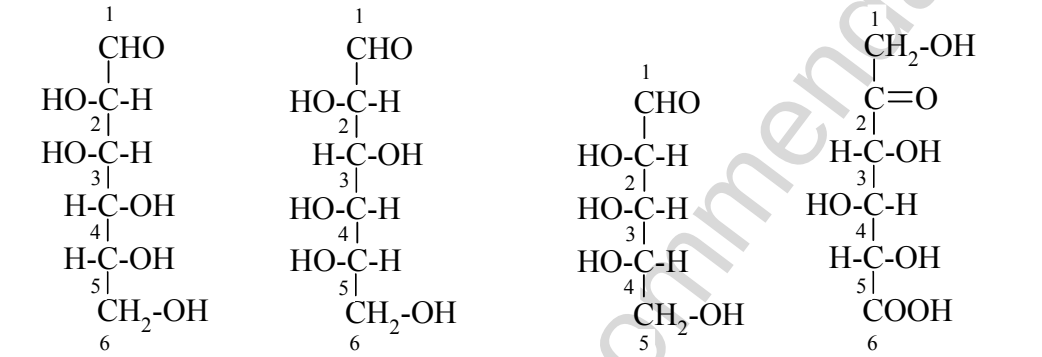


L-glyceraldehyde
(*S*)-2,3-dihydroxypropanal (PIN)

P-102.3.3 The configurational atom

A monosaccharide is assigned to the 'D' or 'L' series according to the configuration of the highest numbered center of chirality. This asymmetrically substituted carbon atom is called the 'configurational atom'. Thus, if the hydroxy group projects to the right in the Fischer projection, the sugar belongs to the D-series, and receives the 'D' stereodescriptor.

Examples:



D-mannose (PIN) L-glucose (PIN) L-ribose (PIN) D-sorbose (PIN)

P-102.3.4 Cyclic forms of monosaccharides

Most monosaccharides exist as cyclic hemiacetals or hemiketals. Two aspects of the internal cyclisation must be examined, first, the size of the ring, and secondly, the conformation of the newly created center of chirality.

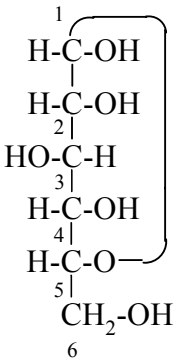
P-102.3.4.1 Ring size

Out of the various possible heterocyclic ring sizes resulting from hemiacetal or hemiketal formation, those with five and six members, including an oxygen atom, prevail and are discussed in this Section. Their names are based on those of the parent heterocycles furan and pyran, respectively. Names are formed by including the terms 'furan' and 'pyran' before the ending 'ose' in the name of a sugar. For example, D-mannose is changed to D-mannopyranose to indicate the cyclic form having a six-membered ring; furthermore, the generic term 'pyranose' includes all the sugars having a six-membered ring structure. Similarly, the sugars having a five-membered ring structure are 'furanoses'; oxiroses, oxetoses and septanoses have a three-, four- or seven-membered cyclic structure, respectively.

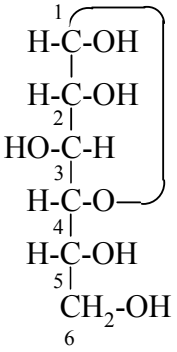
Different representations of cyclic forms are to be considered.

P-102.3.4.1.1 Hemiacetal or hemiketal formation is indicated in the Fischer projection of the cyclic form by a long bond joining the original aldehydic or ketonic group to the oxygen atom included in the ring.

Examples:



D-glucopyranose (PIN)



D-glucofuranose (PIN)

P-102.3.4.1.2 The Haworth representation

The Haworth representation is a perspective drawing. The ring is orientated almost perpendicular to the plane of the paper, but viewed from slightly above so that the edge closer to the viewer is drawn below the most distant edge, with the oxygen behind and C-1 at the right hand end. The cyclisation process is envisaged as proceeding stepwise, as exemplified for D-glucopyranose in Fig. 1, below. Two reorientations are necessary from the standard Fischer projection to prepare the acetalization or ketalization procedure; the first reorientation, step (a), consists in placing the nonterminal hydroxy groups vertically; the second one, step (c), is the reorientation of carbon C-5 to place the oxygen atom in the plane of the ring. The mode of cyclisation must be defined completely by expressing the configuration at carbon '1'.

P-102.3.4.2 Anomeric forms; use of stereodescriptors 'α' and 'β'

P-102.3.4.2.1 In the cyclic form, the configuration of the newly created center of chirality C-1 must be expressed. This center is called the 'anomeric center'. The two stereosiomers are called 'anomers'; they are designated by the stereodescriptors 'α' and 'β' according to the configurational relationship between the anomeric center and the so called 'reference center'.

P-102.3.4.2.2 Configurations 'α' and 'β' for monosaccharides

The anomeric reference center in a monosaccharide having a retained name is the configurational atom as defined in P-103.3.3. In the α-anomer, the exocyclic oxygen atom at the anomeric center is formally '*cis*', in the Fischer projection, to the oxygen atom attached to the anomeric reference atom; in the β-anomer, the relationship is '*trans*'. The reference plane for determining the configurations '*cis*' and '*trans*' is perpendicular to the Fischer projection, including all carbon atoms of the monosaccharide.

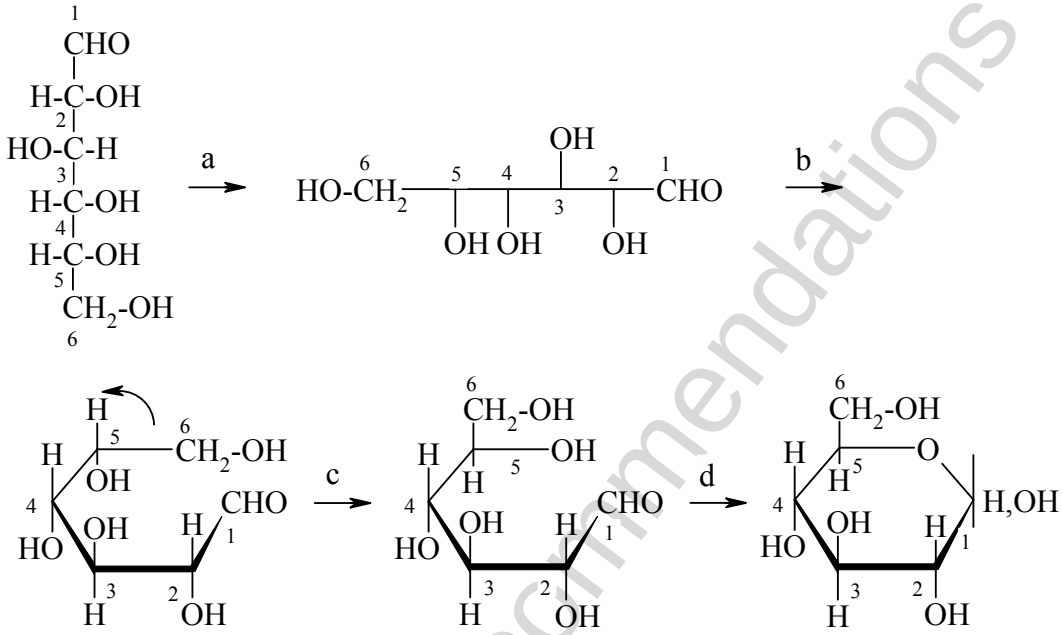
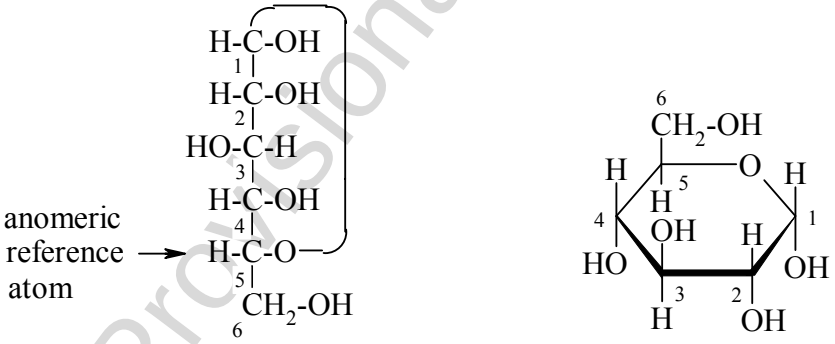


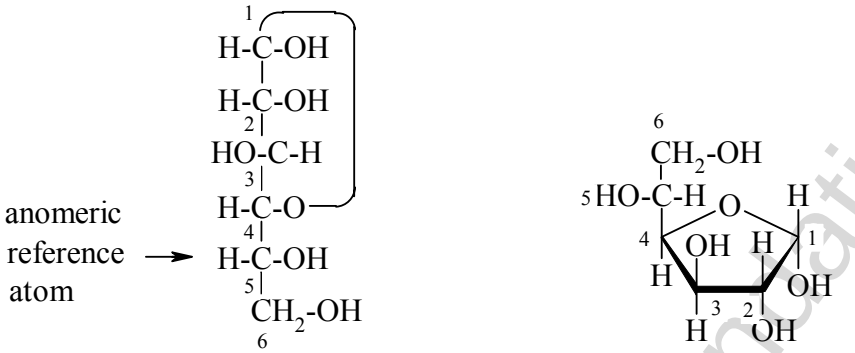
Fig. 1. Reorientation of Fischer projection to Haworth representation

The anomeric stereodescriptor 'α' or 'β', followed by a hyphen, is placed immediately before the configurational stereodescriptor 'D' or 'L' of the retained name.

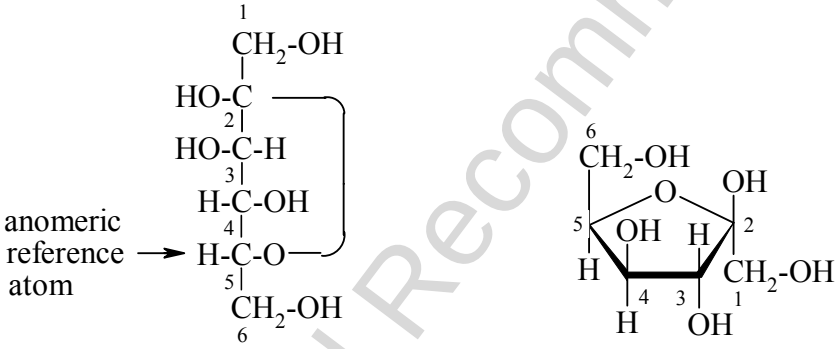
Examples:



α-D-glucopyranose (PIN)



α-D-glucofuranose (PIN)

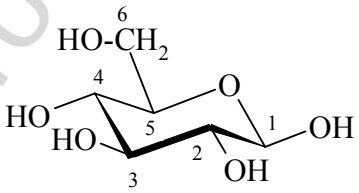


β-D-fructofuranose (PIN)

P-102.3.5 Conformation of monosaccharides

Pyranoses assume conformations that are not planar. For example β-D-glucopyranose assumes a chair conformation with characteristic substituent groups in equatorial conformation (hydrogen atoms are not shown):

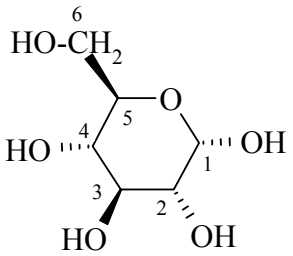
Example:



P-102.3.6 The Mills depiction

In this depiction, the main hemiacetal ring is drawn in the plane of the paper. Hashed wedges denote substituents below this plane, and solid wedges those above.

Example:



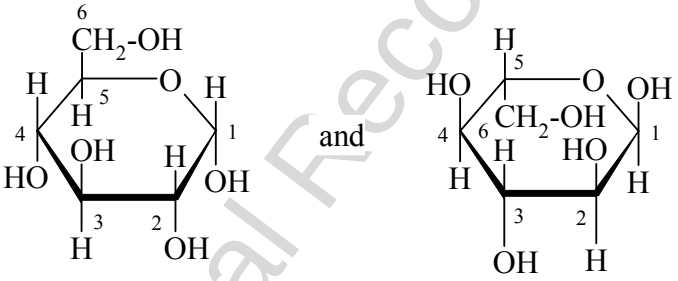
α-D-glucopyranose (PIN)

P-102.3.7 Stereodescriptors for denoting racemates and uncertain configurations

P-102.3.7.1 Stereodescriptors for denoting racemates

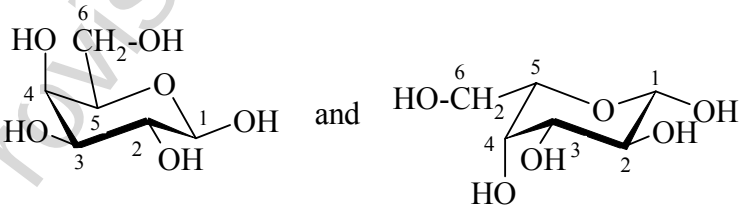
Racemates are indicated by the stereodescriptor 'DL'.

Examples:



D-configuration and L-configuration

α- DL-glucopyranose (PIN)



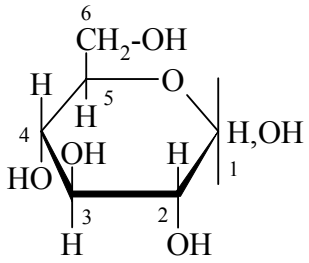
D-configuration and L-configuration

β- DL-galactopyranose (PIN)

P-102.3.7.2 Mixtures of anomers

When a mixture of anomers has to be described, the stereodescriptors 'α' and 'β' are placed at the front of the name; in Haworth representations, the symbols H and OH replace the formal bonds at the anomeric carbon atom.

Examples:



α,β-D-glucopyranose (PIN)

P-102.4 Choice of parent structure

In cases where more than one monosaccharide structure is embedded in a large molecule, a parent structure is chosen on the basis of the following criteria, applied in the order given until a decision is reached:

(a) the parent that includes the functional group most senior in the order of classes (see P-42). If there is a choice, it is made on the basis of the greatest number of occurrences of the most senior functional group. Thus aldaric acid > uronic acid/ketoaldonic acid/aldonic acid > dialdose > ketoaldose/aldose > diketose > ketose;

(b) the parent with the greatest number of carbon atoms in the chain, e.g. heptose rather than hexose;

(c) the parent with the name that comes first in an alphabetical listing based on the following:

(i) the trivial name or the configurational prefix(es) of the systematic name, e.g., glucose (PIN) rather than gulose; a '*gluco*' rather than a '*gulo*' derivative;

Example: D-glucitol (PIN) rather than L-gulitol (see P-102.5.5.5.1);

(ii) the configurational symbol D rather than L;

Example: 5-*O*-methyl- D-galactitol (PIN) rather than 2-*O*-methyl- L-galactitol (see P-102.5.5.5.2);

(iii) the anomeric stereodescriptor α rather than β;

Example: α-D-fructofuranose β-D-fructofuranose 1,2′:1′,2-dianhydride (PIN) and not β-D-fructofuranose α-D-fructofuranose 1,2′:1′,2-dianhydride (see P-102.5.5.7.2).

(d) the parent with the most substituent groups cited as prefixes (bridging substitution, for example 2,3-*O*-methylene is regarded as multiple substitution for this purpose); the prefixes 'deoxy' and 'anhydro' are detachable and alphabetized, thus regarded as substituent groups;

(e) the parent with the lowest locants for substituent prefixes;

Example: 2,3,5-tri-*O*-methyl-D-mannitol (PIN) rather than 2,4,5-tri-*O*-methyl-D-mannitol [see P-102.5.5.5.3(a)]

(f) the parent with the lowest locant for the first cited substuent.

Example: 2-*O*-acetyl-5-*O*-methyl-D-mannitol (PIN) rather than 5-*O*-acetyl-2-*O*-methyl-D-mannitol [see P-102.5.5.5.3 (b)].

P-102.5 Monosaccharides

P-102.5.1 Aldoses
P-102.5.2 Ketoses
P-102.5.3 Deoxy sugars
P-102.5.4 Amino sugars
P-102.5.5 Thio sugars
P-102.5.6 Substituted monosaccharides

P-102.5.1 Aldoses

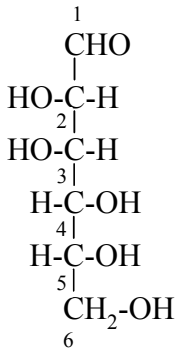
Names of aldoses are retained or substitutively formed. Retained names for aldoses with three to six carbon atoms are listed in Table 10.2. Except for glyceraldehyde, they are preferred IUPAC names. Names of aldoses having more than six carbon atoms are formed in two ways: by the procedures of systematic carbohydrate nomenclature, and by those of systematic substitutive nomenclature. Systematic substitutive names are preferred IUPAC names.

P-102.5.1.1 Systematic carbohydrate names

Systematic carbohydrate names of aldoses are formed from a stem name and a configurational prefix or prefixes. Stem names for the aldoses with three to ten carbon atoms are triose, tetrose, pentose, hexose, heptose, octose, nonose, and decose. The chain is numbered so that the carbonyl group receives the locant '1'.

P-102.5.1.1.1 The configuration of >CH-OH groups of the sugar is designated by the configurational prefix(es) listed in Table 10.2., such as '*glycero*', '*gluco*', '*manno*', etc. Each name is qualified by a 'D' or 'L' stereodescriptor, as defined in P-102.3.2.

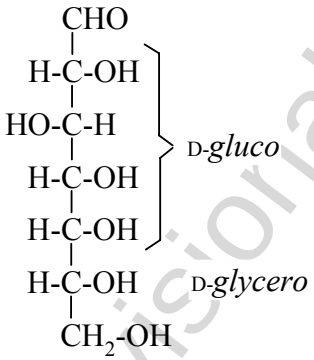
Example:



D-*manno*-hexose (systematic carbohydrate name)
D-mannose (PIN)

P-102.5.1.1.2 Aldoses composed of more than four chiral centers are named by adding two or more configurational prefixes (listed in Table 10.2) to the stem name. Prefixes are assigned in order to the chiral centers in groups of four, beginning with the group located next to the aldehydic group. The prefix relating to the group of carbon atoms farthest from the aldehydic group (which may contain fewer than four chiral centers) is cited first. These names are for use only in general nomenclature.

Examples:

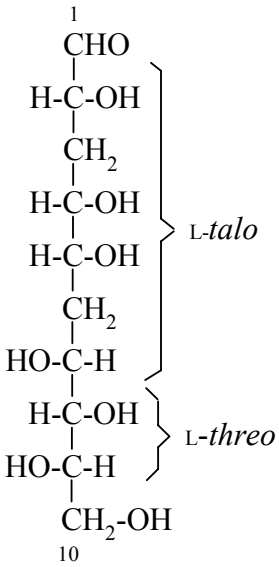


D-*glycero*- D-*gluco*-heptose
(not D-*gluco*- D-*glycero*-heptose)
(2*R*,3*S*,4*R*,5*R*,6*R*)-2,3,4,5,6-hexahydroxyheptanal (PIN)

P-102.5.1.1.3

When sequences of chiral centers are separated by nonchiral centers, the nonchiral centers are ignored, and the remaining set of chiral centers is assigned the appropriate configurational prefix (for four centers or less) or prefixes (for more than four centers).

Example:

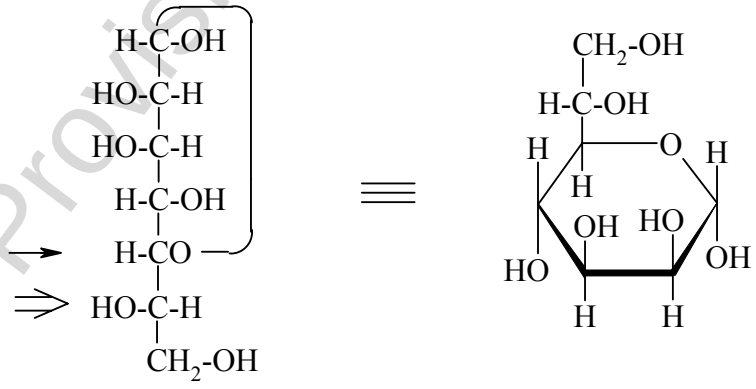


3,6-dideoxy-L-*threo*-L-*talo*-decose
(for deoxy sugars, see P-102.4.3)
(2*R*,4*S*,5*R*,7*R*,8*S*,9*S*)-2,4,5,7,8,9,10-heptahydroxydecanal (PIN)

P-102.5.1.1.4 Cyclic forms

For monosaccharides having more than six carbon atoms, the anomeric reference center is the highest numbered atom of the group of chiral centers next to the anomeric center that is involved in the heterocyclic ring and specified by a single configurational prefix. In the α-anomer, the exocyclic oxygen atom at the anomeric center is formally '*cis*', in the Fischer projection, to the oxygen atom attached to the anomeric reference atom; in the β-anomer these oxygen atoms are formally '*trans*'.

Example:



→ denotes the anomeric reference atom; ⇒ denotes the configurational atom

L-*glycero*-α-D-*manno*-heptopyranose
(2*S*,3*S*,4*S*,5S,6*R*)-6-[(1*S*)-1,2-dihydroxyethyl]oxane-2,3,4,5-tetrol (PIN)

P-102.5.2 Ketoses

P-102.5.2.1 Classification

Ketoses are classified as 2-ketoses, 3-ketoses, etc. according to the lowest locant for the position of the (potential) carbonyl group.

P-102.5.2.2 Retained names

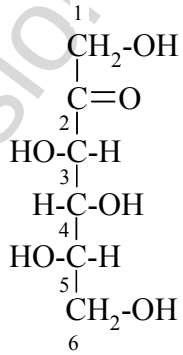
Retained names and structures are shown in Table 10.3. The retained names erythrulose, ribulose, xylulose, psicose, fructose, sorbose and tagatose are preferred IUPAC names for the 2-ketoses with four to six carbon atoms. Preferred IUPAC names for ketoses with more than six carbon atoms are formed by systematic substitutive nomenclature, the configuration being designated by the appropriate CIP stereodescriptors '*R*', '*S*', '*r*', *etc.*

P-102.5.2.3 Systematic carbohydrate names

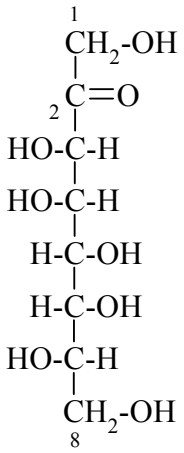
The systematic carbohydrate names of ketoses having four to six carbon atoms are formed from the stem name and the appropriate configurational prefix listed in Table 10.3. The stem names are formed from the corresponding aldoses stem names by replacing the ending 'ose' with 'ulose', preceded by the locant of the carbonyl group, e.g. 'pent-2-ulose' and 'hex-3-ulose'. The chain is numbered so that the carbonyl group receives the lowest possible locant. When the carbonyl group is in the middle of a chain with an odd number of carbon atoms, a choice between alternative names is made according to alphanumerical order.

For 2-ketoses, configurational prefixes are given in the same way as for aldoses. Retained names are preferred IUPAC names. Systematic carbohydrate names for 3-ketoses and ketoses with more than six carbon atoms are used only in general nomenclature. Preferred IUPAC names are systematic substitutive names.

Examples:



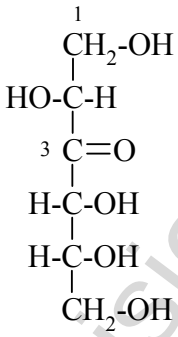
L-*xylo*-hex-2-ulose
L-sorbose (PIN)



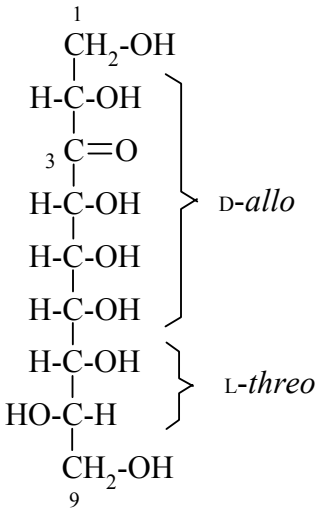
L-*glycero*-D-*manno*-oct-2-ulose
(3*S*,4*S*,5*R*,6*R*,7*S*)-1,3,4,5,6,7,8-heptahydroxyoctan-2-one (PIN)

For ketoses with the carbonyl group at C-3, or at a higher-numbered carbon atom, the carbonyl group is ignored and the set of chiral centers is given the appropriate prefix or prefixes according to Table 10.3.

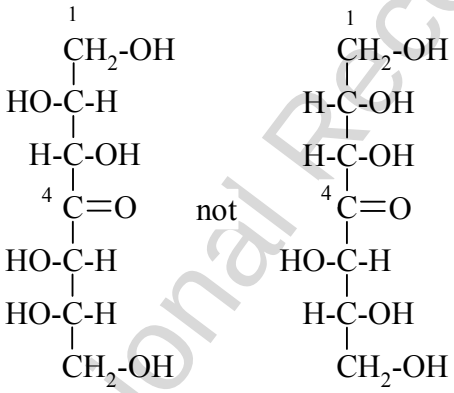
Examples:



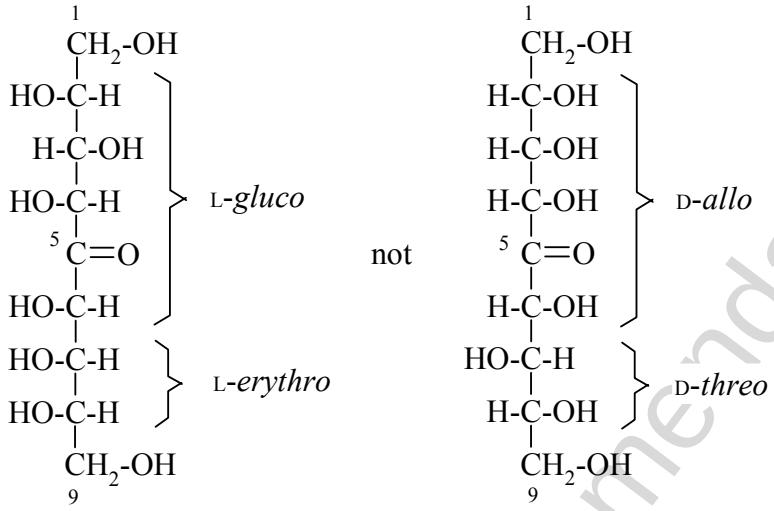
D-*arabino*-hex-3-ulose
(2*R*,4*R*,5*R*)-1,2,4,5,6-pentahydroxyhexan-3-one (PIN)



L-*threo*-D-*allo*-non-3-ulose
(2*S*,4*R*,5*R*,6*R*,7*R*,8*S*)-1,2,4,5,6,7,8-heptahydroxynonan-3-one



L-*gluco*-hept-4-ulose
(not D-*gulo*-hept-4-ulose; *gluco* is first in alphanumerical order see P-102.)
(2*R*,3*S*,5*S*,6*S*)-1,2,3,5,6,7-hexahydroxyheptan-4-one (PIN)
(not (2*S*,3*S*,5*S*,6*R*)-1,2,3,5,6,7-hexahydroxyheptan-4-one)
(when there is a choice, the *R* configuration is assigned to the lowest locant)



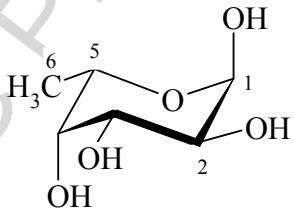
L-*erythro*-L-*gluco*-non-5-ulose
(not D-*threo*-D-*allo*-non-5-ulose; *erythro-gluco* is first in alphanumerical order)
(2*R*,3*S*,4*R*,6*S*,7*S*,8*S*)-1,2,3,4,6,7,8,9-octahydroxynonan-5-one (PIN)

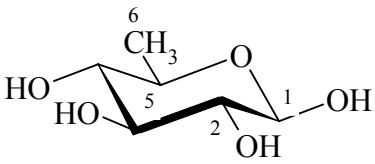
P-102.5.3 Deoxy sugars

P-102.5.3.1 The prefix 'deoxy' means the removal of an 'oxy' group, –O–, with rejoining of the hydrogen atom. In these recommendations, the prefix 'deoxy' is classified as detachable; i.e., it is alphabetized among the substituents arising from substitutive nomenclature. This is a change from the previous status (see R-0.1.8.4, ref. 2) that classified the prefix 'deoxy' among nondetachable prefixes (see also the prefix 'anhydro, that is now classified as detachable and alphabetized among all detachable prefixes).

P-102.5.3.2 Trivial names.

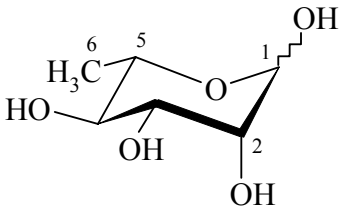
The following names are retained: fucose, quinovose and rhamnose. The corresponding structures are shown in the pyranose form. These names are used as preferred IUPAC names for the unmodified sugars. Systematic substitutive names are preferred IUPAC names for the formation of names of derivatives.





α-L-fucopyranose (PIN)
6-deoxy-α-L-galactopyranose

β-D-quinovopyranose (PIN)
6-deoxy-β-D-glucopyranose



L-rhamnopyranose (PIN)
6-deoxy-L-mannopyranose

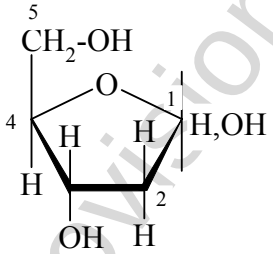
P-102.5.3.3 Carbohydrate names derived from retained names

Use of the prefix 'deoxy' in combination with a retained name other than glucose, mannose and galactose gives preferred IUPAC names when the deoxygenation does not involve the configuration at any chirality center, for example, 6-deoxy-D-allose. When the prefix 'deoxy' modifies a chirality center, a carbohydrate name is appropriate, but preferred IUPAC names are formed by substitutive nomenclature with CIP stereodescriptors (see P-102.5.3.4 for examples). As an exception, the combination of 'amino' and 'dexoy at the same position is allowed in preferred IUPAC names.

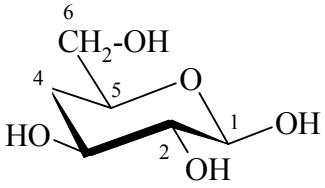
P-102.5.3.4 Systematic carbohydrate names

The systematic carbohydrate name consists of the prefix 'deoxy', preceded by the appropriate locant and followed by the stem name with such configurational prefixes as necessary to describe the chirality centers present in the deoxy compound. Configurational prefixes are cited in order commencing at the end furthest from C-1.

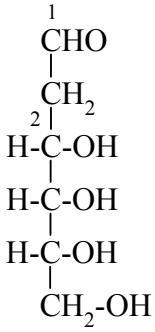
Examples:



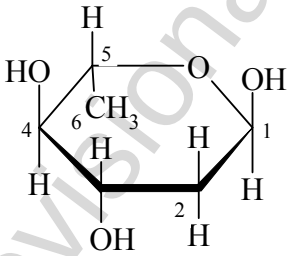
2-deoxy-D-*erythro*-pentofuranose
(often referred to as 2-deoxy-D-ribose)
(2*R*,3*S*,5ξ)-5-(hydroxymethyl)oxolane-3,5-diol (PIN)



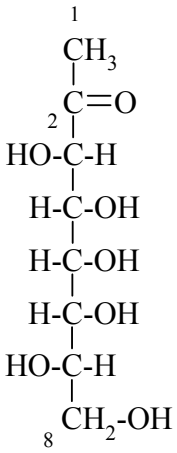
4-deoxy-β-D-*xylo*-hexopyranose
(not 4-deoxy-β-D-galactopyranose)
(2*R*,3*R*,4*S*,6*S*)-6-(hydroxymethyl)oxane-2,3,4-triol (PIN)



2-deoxy-D-*ribo*-hexose
(not 2-deoxy-D-allose)
(3*S*,4*S*,5*R*)-3,4,5,6-tetrahydroxyhexanal (PIN)



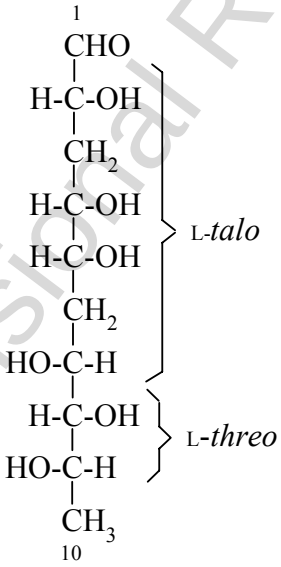
2,6-dideoxy-α-L-*arabino*-hexopyranose
(2*R*,4S,5*R*,6*S*)-6-methyloxane-2,4,5-triol (PIN)



1-deoxy-L-*glycero*-D-*altro*-oct-2-ulose
(3*S*,4*R*,5*R*,6*R*,7*S*)-3,4,5,6,7,8-hexahydroxyoctan-2-one (PIN)

When the –CH_2– group divides the chirality centers into two sets, it is ignored for the purpose of assigning the configurational prefix; the prefix(es) assigned should cover the entire sequence of chirality centers (see aldoses) (see P-102.5.1.1.3).

Example:



3,6,10-trideoxy-L-*threo*-L-*talo*-decose
(2*R*,4*S*,5*R*,7*R*,8*R*,9*S*)-2,4,5,7,8,9,10-heptahydroxydecanal (PIN)

P-102.5.4 Amino sugars and thiosugars and other chalcogen analogues

The replacement of a hydroxy group that is not an anomeric hydroxy group of a monosaccharide or a monosaccharide derivative by an amino group is envisaged as substitution of the appropriate hydrogen atom of the corresponding deoxy monosaccharide by an amino

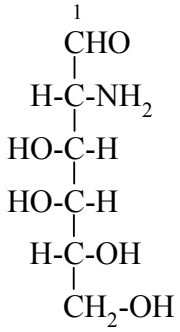
group. The configuration at the carbon atom carrying the amino group is expressed as that of an aldose, considering that the amino group has replaced a hydroxy group.

To the contrary, the replacement of a hydroxy group by a sulfanyl group is considered to be a functional replacement indicated by the prefix ‘thio’.

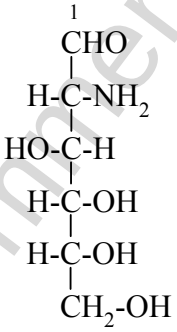
P-102.5.4.1 Amino sugars

P-102.5.4.1.1 Trivial names

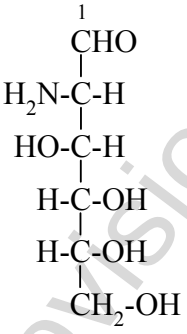
The following glycosamine names are retained but preferred IUPAC names are based on aldose names with amino and deoxy prefixes.



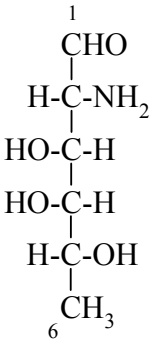
D-galactosamine
2-amino-2-deoxy-D-galactose (PIN)



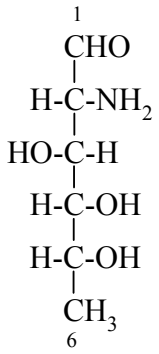
D-glucosamine
2-amino-2-deoxy-D-glucose (PIN)



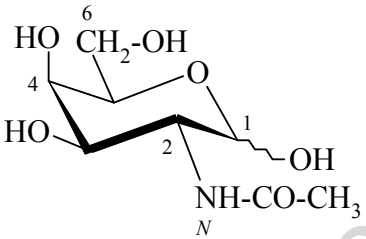
D-mannosamine
2-amino-2-deoxy-D-mannose (PIN)



D-fucosamine
2-amino-2,6-dideoxy-D-galactose (PIN)



D-quinovosamine
2-amino-2,6-dideoxy-D-glucose (PIN)

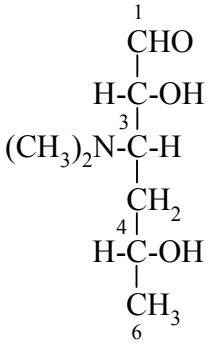


N-acetyl-D-galactosamine
N-acetamido-2-deoxy-D-galactopyranose (PIN)

P-102.5.4.1.2 Systematic carbohydrate names

Systematic carbohydrate names are formed, in two steps: in a first step a deoxy sugar is created by deoxygenation at the carbon atom where the amino group is to be introduced by substitution in a second step. Names of substituted amines are formed by using the name of the substituted amino group as a prefix.

Example:



3,4,6-trideoxy-3-(dimethylamino)-D-*xylo*-hexose
(2*R*,3*S*,5*R*)-3-(dimethylamino)-2,5-dihydroxyhexanal (PIN)

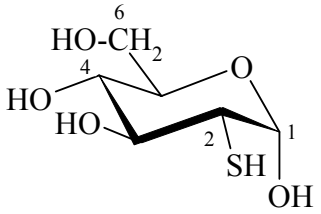
P-102.5.5 Thio sugars and other chalcogen analogues

The replacement of a hydroxy oxygen atom of an aldose or ketose, or of the oxygen atom of the carbonyl group of an acyclic aldose or ketose, by sulfur, selenium or tellurium is indicated by placing the prefix 'thio', 'seleno' or 'telluro', respectively, preceded by the appropriate locant, at the front of the systematic or trivial name of the aldose or ketose. In carbohydrate nomenclature, the prefixes 'thio', 'seleno' and 'telluro' are considered as detachable, alphabetized prefixes.

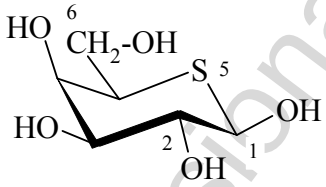
Replacement of the ring oxygen atom of the cyclic form of an aldose or ketose by sulfur, selenium, or tellurium is indicated in the same way, the number of the nonanomeric adjacent carbon atom of the ring being used as locant. In such a case, skeletal replacement expressed by 'a' replacement prefixes is not recommended.

Sulfoxides (and selenoxides or telluroxides) and sulfones (and selenones or tellurones) are named by functional class nomenclature (see P-63.6 for functional class names of sulfoxides and sulfones).

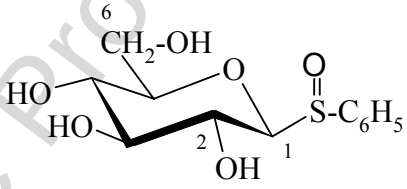
Examples:



2-thio-α-D-glucopyranose (PIN)



5-thio-β-D-galactopyranose (PIN)



β-D-glucopyranosyl phenyl sulfoxide
(for glycosyl groups, see P-102.6.1.1)
(2*S*,3*R*,4*S*,5*S*,6*R*)-2-(benzenesulfinyl)-6-(hydroxymethyl)oxane-3,4,5-triol (PIN)

P-102.5.6 Derivatives of monosaccharides

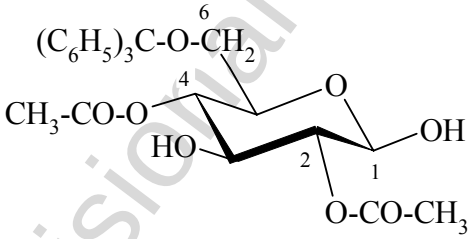
P-102.5.6.1 *O*-substitution
P-102.5.6.2 Glycosides
P-102.5.6.3 *C*-Substitution
P-102.5.6.4 *N*-substitution
P-102.5.6.5 Alditols
P-102.5.6.6 Carboxylic acids derived from monosaccharides
P-102.5.6.7 Anhydrides

P-102.5.6.1 *O*-Substitution

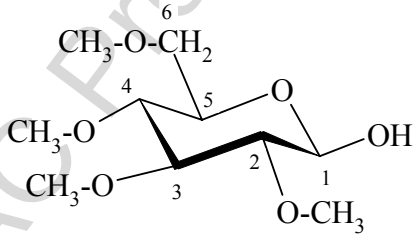
In order to maintain the integrity of structures and take advantage of retained names to imply the absolute configuration, *O*-substitution is allowed in carbohydrate nomenclature. Substituents replacing the hydrogen atom of an alcoholic hydroxy group of a monosaccharide or monosaccharide derivative are denoted as *O*-substituents. The substitution of an anomeric hydroxy group is discussed in P-102.5.5.2.2. The *O*-locant is not repeated for multiple substitutions by the same atom or group. Number locants are used as necessary to specify the positions of substituents; they are not required for compounds fully substituted by identical atoms or groups.

P-102.5.6.1.1 *O*-Acetyl and *O*-alkyl substitution. Names using acyl groups are preferred to names based on ester nomenclature ending in 'ate'.

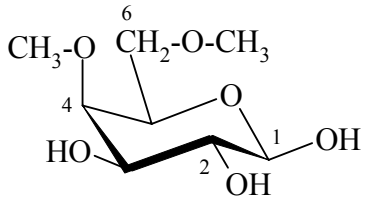
Examples:



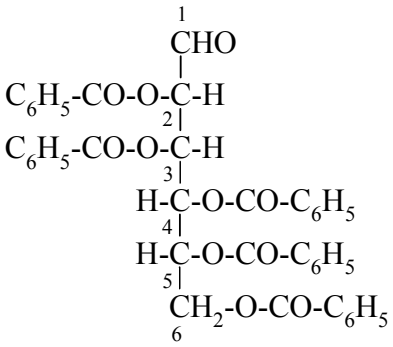
2,4-di-*O*-acetyl-6-*O*-trityl-β-D-glucopyranose (PIN)



2,3,4,6-tetra-*O*-methyl-β-D-glucopyranose (PIN)



4,6-di-*O*-methyl-β-D-galactoyranose (PIN)



2,3,4,5,6-penta-*O*-benzoyl-D-mannose (PIN)

P-102.5.6.1.2 Phosphoric acid esters

Esters of sugars with phosphoric acid are generally termed 'phosphates'. In biochemical usage, the term 'phosphate' indicates the phosphate residue regardless of the state of ionization or the counter ions present. Preferred IUPAC names must differentiate between a true phosphate, $–O\text{-}PO(O^–)_2$, and an acid phosphate, i.e., $–O\text{-}PO(OH)_2$, called a (dihydrogen phosphate). The prefixes 'phosphono', for $–PO(OH)_2$, and 'phosphonato', for $PO(O^–)_2$, are also used, to denote *O*-phosphonic acid derivatives.

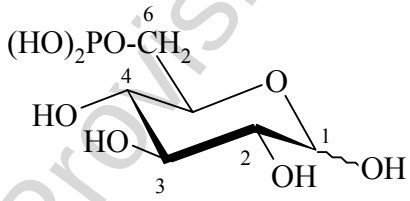
The term 'phospho' is used in place of 'phosphono' and 'phosphonato' in biochemical contexts.

When the sugar is esterified by two or more phosphate groups, the numerical terms 'bis', 'tris' are used, as 'bis(phosphate)', 'tris(phosphate)'.

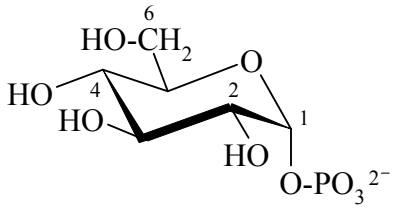
Phosphonates are treated in the same way as phosphates.

All phosphate esters names, including 1-phosphates, are preferred IUPAC names.

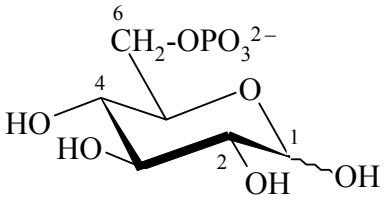
Examples:



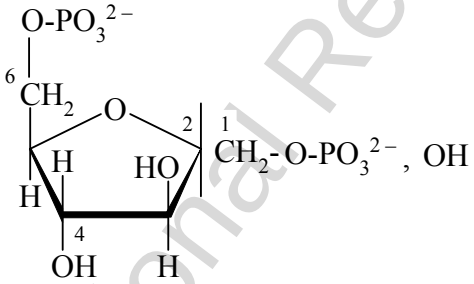
D-glucopyranose 6-(dihydrogen phosphate) (PIN)
6-*O*-phosphono-D-glucopyranose



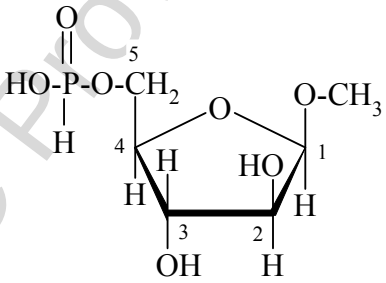
α-D-glucopyranosyl phosphate
α- D-glucopyranose 1-phosphate (PIN)



D-glucopyranose 6-phosphate (PIN)
6-*O*-phosphonato- D-glucopyranose



D-fructofuranose 1,6-bisphosphate (PIN)
1,6-di-*O*-phosphonato-D-fructofuranose

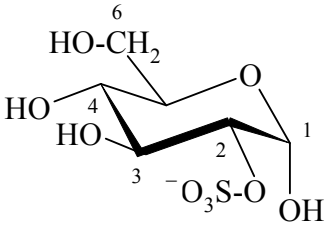


methyl β-D-ribofuranoside 5-(hydrogen phosphonate) (PIN)
methyl 5-deoxy- β-D-ribofuranosid-5-yl hydrogen phosphonate

P-102.5.6.1.2 Esters with sulfuric acid

Esters of sugars with sulfuric acid are named by adding the term 'sulfate' after the name of the sugar, with the appropriate locant. The prefixes 'sulfo' for $–SO_3H$, and 'sulfonato' for $SO_3^–$, can be used to denote *O*-derivatives.

Example:



α-D-glucopyranose 2-sulfate (PIN)
2-*O*-sulfonato-α-D-glucopyranose

P-102.5.6.2 Glycosides

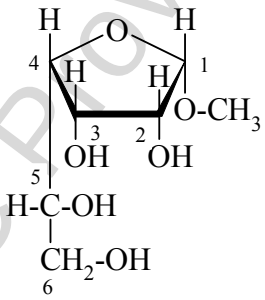
P-102.5.6.2.1 Definitions

Glycose is a less frequently used term for monosaccharide. Glycosides are mixed acetals (ketals) derived from cyclic forms of monosaccharides, having thus an *O*-substituted anomeric –OH group, such as –OR. See ref. 22 for a full discussion on the use of the term glycoside.

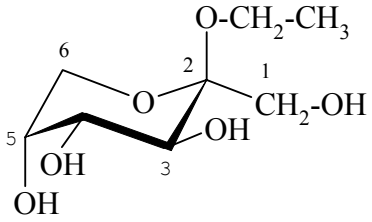
P-102.5.6.2.2 Names

Glycosides are named by using functional class nomenclature. The name of the class 'glycoside' is adapted to the name of each cyclic monosaccharide, by changing the letter 'e' at the end of the name to 'ide', for example glucopyranose becomes glucopyranoside, fructofuranose becomes fructofuranoside. The class name is preceded, as a separate word, by the name of the substituent group that is part of the acetal or ketal function.

Examples:



methyl α-D-gulopyranoside (PIN)



ethyl β-D-fructopyranoside (PIN)

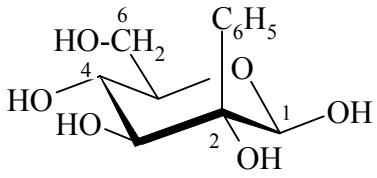
P-102.5.6.3 *C*-substitution

P-102.5.6.3.1 Substitution at a nonterminal carbon atom
P-102.5.6.3.2 Substitution replacing a hydroxy group
P-102.5.6.3.3 Substitution at a terminal carbon atom

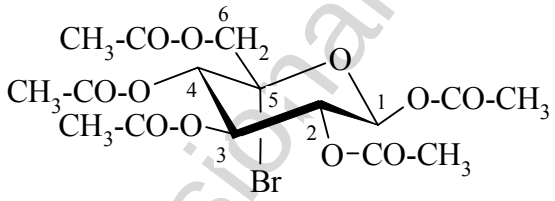
P-102.5.6.3.1 Substitution at a nonterminal carbon atom

The compound is named as a *C*-substituted monosaccharide. The group having priority in accordance with the CIP priority system is regarded as equivalent to –OH for assignment of configuration. Any ambiguity (e.g. at a carbon atom where ring formation occurs) is avoided by using the *R*,*S* system to specify the configuration at the modified chirality center. Preferred IUPAC names are substitutive names.

Example:



2-*C*-phenyl-β-D-glucopyranose
(2*R*,3*R*,4*S*,5*S*,6*R*)-6-(hydroxymethyl)-3-phenyloxane-2,3,4,5-tetrol (PIN)

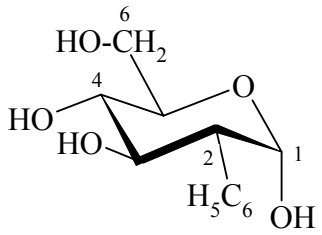


penta-*O*-acetyl-5-*C*-bromo-β-D-glucopyranose
(2*R*,3*R*,4*R*,5*S*,6*S*)-6-bromo-6-[(acetyloxy)methyl]oxane-2,3,4,5-tetrayl tetraacetate (PIN)

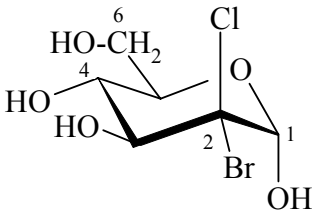
P-102.5.6.3.2 Substitution replacing a nonterminal hydroxy group

The compound is named as a substituted derivative of a deoxy sugar. The group replacing the –OH group determines the configuration. Any potential ambiguity must be dealt with by the use of the '*R*/*S*' system. The '*R*/*S*' system must be used to assign the preferred configuration of a chirality center twice substituted; this method is preferable to that establishing the configuration by making the substituent with high CIP priority equivalent to the –OH group.

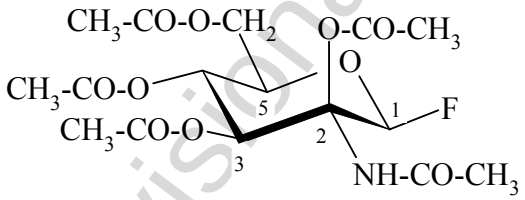
Example:



2-deoxy-2-phenyl-α-D-glucopyranose
2-deoxy-2-*C*-phenyl-α-D-glucopyranose
(2*R*)-2-deoxy-2-phenyl-α-D-*arabino*-hexopyranose
(1*S*,2*R*,3*R*,4*S*,5*R*)-6-(hydroxymethyl)-3-phenyloxane-1,3,4-triol (PIN)



(2*R*)-2-bromo-2-chloro-2-deoxy-α-D-*arabino*-hexopyranose
2-bromo-2-chloro-2-deoxy-α-D-glucopyranose
(2*S*,3*R*,4*S*,5*S*,6*R*)-3-bromo-3-chloro-6-(hydroxymethyl)oxane-2,3,4-triol (PIN)

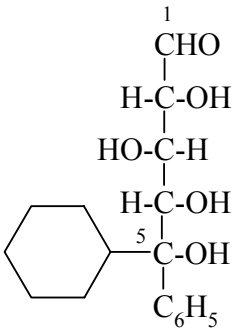


2-acetamido-2,3,4,6-tetra-*O*-acetyl-β-D-mannopyranosyl fluoride
(2*S*,3*S*,4*S*,5*S*,6*R*)-3-acetamido-6-[(acetyloxy)methyl]-2-fluorooxane-3,4,5-triyl triacetate (PIN)

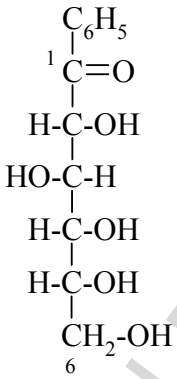
P-102.5.6.3.3 Substitution at a terminal carbon atom

Substitution at a terminal carbon atom of a carbohydrate chain creates a new chirality center; the configuration is indicated by the '*R*/*S*' system. Preferred IUPAC names are formed substitutively.

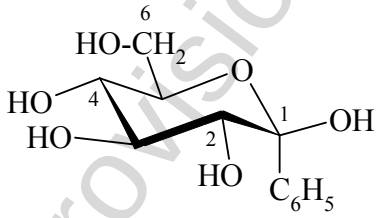
Examples:



(5*R*)-5-*C*-cyclohexyl-5-*C*-phenyl-D-xylose
(2*R*,3*S*,4*S*,5*R*)-5-cyclohexyl-2,3,4,5-tetrahydroxy-5-phenylpentanal (PIN)



1-phenyl-D-glucose
(2*R*,3*S*,4*R*,5*R*)-1-phenyl-2,3,4,5,6-pentahydroxyhexan-1-one (PIN)



1-*C*-phenyl-β-D-glucopyranose
(2*R*,3*R*,4*S*,5*S*,6*R*)-5-(hydroxymethyl)-2-phenyloxane-2,3,4,5-tetrol (PIN)

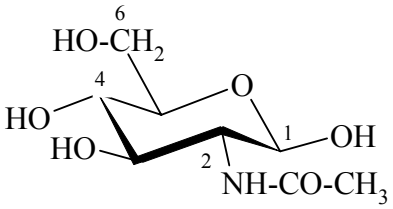
P-102.5.6.4 *N*-substitution

Substitution of the $-NH_2$ group of an amino sugar is dealt with in two different ways:

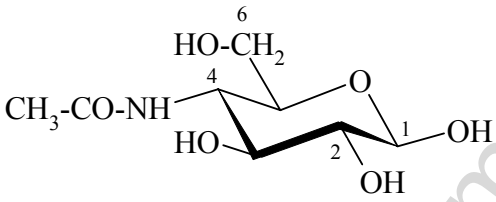
(1) The whole substituted amino group is designated as a prefix as in 2-acetamido (or 2-butylamino)-2-deoxy-D-glucose.

(2) If the amino sugar has a retained trivial name, the substitution is indicated by a prefix preceded by the capital italicized letter *N*.

Examples:



2-acetamido-2-deoxy-β-D-glucopyranose (PIN)
N-acetyl-β-D-glucosamine



4-acetamido-4-deoxy-β-D-glucopyranose (PIN)

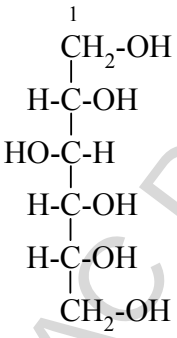
P-102.5.6.5 Alditols

Alditols are named by changing the ending 'ose' in the name of the corresponding aldose into 'itol'.

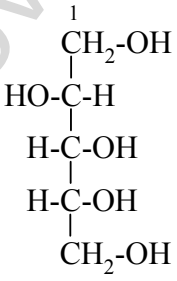
P-102.5.6.5.1 Choice of a parent structure

When the same alditol can be derived from either of two different aldoses, or from an aldose or a ketose, the recommended structure is derived from Rule P-102.3, with the exception of the retained names fucitol and rhamnitol.

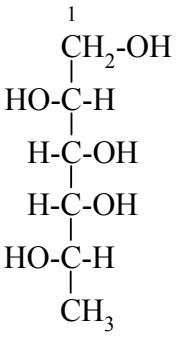
Examples:



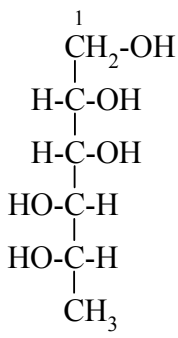
D-glucitol (PIN)
(not L-gulitol)



D-arabinitol (PIN)
(not D-lyxitol)



L-fucitol (PIN)
1-deoxy-D-galactitol

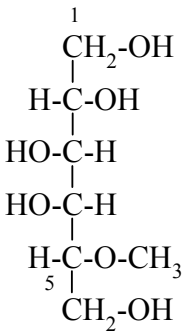


L-rhamnitol (PIN)
1-deoxy- L-mannitol

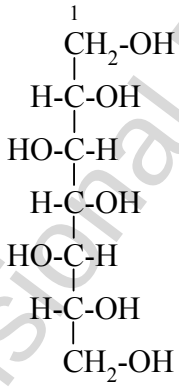
P-102.5.6.5.2 *meso*-Forms

The prefix '*meso*' must be included in the preferred names of erythritol, ribitol and galactitol. The stereodescriptor 'D' or 'L' must be given when a derivative of a '*meso*' form has become asymmetric by substitution. It is also necessary to use the stereodescriptor 'D' or 'L' in the case where there are more than four contiguous chirality centers.

Example:



5-*O*-methyl-D-galactitol (PIN)
(a 'D' configuration is senior to 'L')



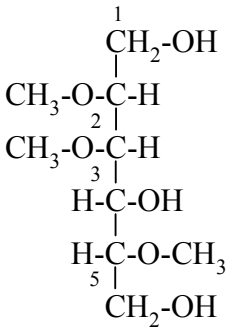
meso- D-*glycero*-L-*ido*-heptitol
(a 'D' configuration is senior to 'L')
(2S,3R,4r,5S,6R)-heptane-1,2,3,4,5,6,7-heptol (PIN)

P-102.5.6.5.3 Choice of parent structure for substituted alditols

The parent structure must have:

(a) the lowest locants for substituent prefixes in accordance with criterion (e) in Rule P-102.4.

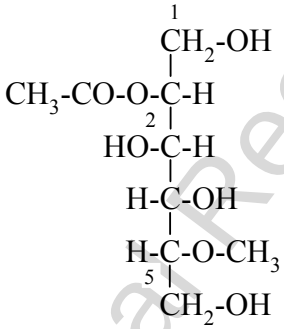
Example:



2,3,5-tri-*O*-methyl-D-mannitol (PIN)
(not 2,4,5-tri-*O*-methyl-D-mannitol)

(b) the lowest locant for the first cited substituent in alphanumerical order, in accordance with criterion (f) in Rule P-102.4.

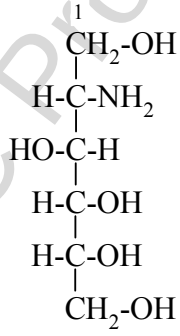
Example:



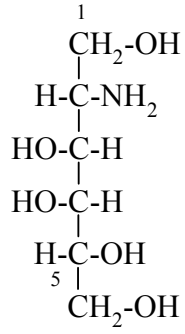
2-*O*-acetyl-5-*O*-methyl-D-mannitol (PIN)
(not 5-*O*-acetyl-2-*O*-methyl-D-mannitol)

P-102-5.6.5.4 Aminoalditols

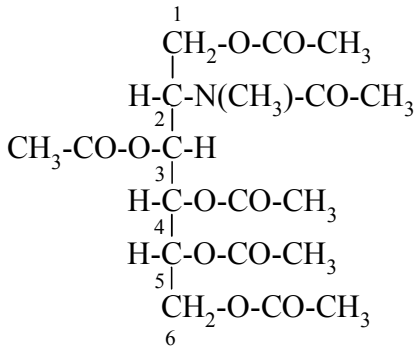
Alditols derived from galactosamine and glucosamine are aminoalditols. They have retained names, galactosaminitol and glucosaminitol, respectively.



D-glucosaminitol
2-amino-2-deoxy-D-glucitol (PIN)



D-galactosaminitol
2-amino-2-deoxy-D-galactitol (PIN)



1,3,4,5,6-penta-*O*-acetyl-2-deoxy-2-(*N*-methylacetamido)-D-glucitol (PIN)

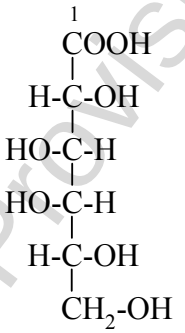
P-102.5.6.6 Monosaccharide carboxylic acids

P-102.5.6.6.1 Definitions

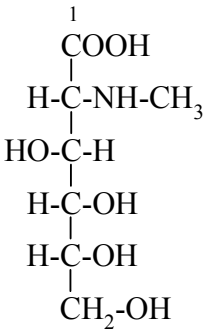
P-102.5.6.6.2 Aldonic acids.
P-102.5.6.6.3 Ketoaldonic acids.
P-102.5.6.6.4 Uronic acids.
P-102.5.6.6.5 Aldaric acids.

P-102.5.6.6.2 Aldonic acids

Aldonic acids are monocarboxylic acids formally derived from aldoses by oxidation of the aldehydic group to a carboxylic acid. Aldonic acids are divided into aldotrionic acids, aldotetronic acids, etc. according to the number of carbon atoms in the chain. The names of individual compounds are formed by changing the ending 'ose' of the retained or systematic name of the aldose to 'onic acid'. The locant 1 is assigned to the carboxy group.
Examples:



D-galactonic acid (PIN)

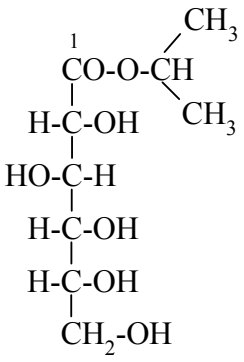


2-deoxy-2-(methylamino)-D-gluconic acid (PIN)

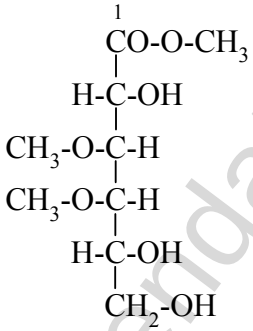
P-102.5.6.6.2.1 Derivatives of aldonic acids

Aldonic acids are treated as carboxylic acids having a retained name. They can form salts, esters, anhydrides, acyl groups and acid halides and pseudohalides, amides, hydrazides, nitriles and chalcogen analogues as described in Sections 65 and 66 for systematic nomenclature. Preferred IUPAC names are formed by using the described methodology for this type of name.

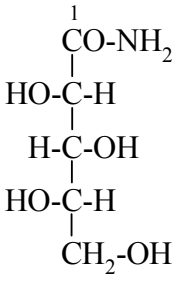
Examples:



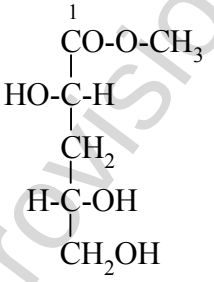
propan-2-yl D-gluconate (PIN)



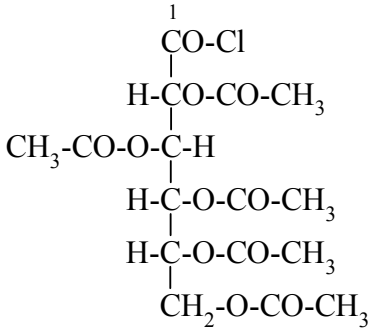
methyl 3,4-di-*O*-methyl-D-galactonate (PIN)



L-xylonamide (PIN)



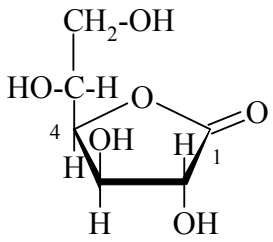
methyl 3-deoxy- D-*threo*-pentonate (PIN)



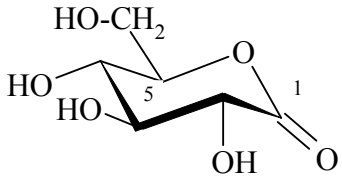
penta-*O*-acetyl- D-gluconoyl chloride (PIN)

Lactones and lactams are named by adapting Rules P-65.6.3.3.2 and P-66.1.4.1, respectively. Two locants are used before the lactone or lactam term: the first one is the locant 1 denoting the carboxy group position; the second locant denotes the position of attachment on the carbon chain. To name lactams, the amino group, $-NH_2$, must be generated and cited. The use of Greek letters to indicate the size of a lactone or lactam ring is not recommended. Preferred IUPAC names are formed substitutively on the basis of heterocyclic rings, with CIP stereodescriptors.

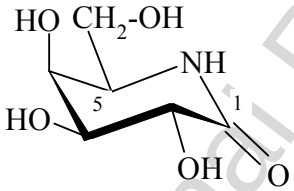
Examples:



D-glucono-1,4-lactone
(3R,4*R*,5*R*)- 4-[(1*R*)-1,2-dihydroxyethyl]-3,4-dihydroxyoxolan-2-one (PIN)



D-glucono-1,5-lactone
(2*R*,3*S*,4*S*,5*R*)-2,4-dihydroxy-5-(hydroxymethyl)oxolan-2-one (PIN)

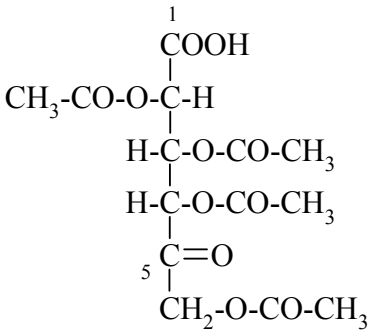


5-amino-5-deoxy- D-galactono-1,5-lactam
(3*R*,4*S*,5*S*,6*R*)-3,4,5-trihydroxy-6-(hydroxymethyl)piperidin-2-one (PIN)

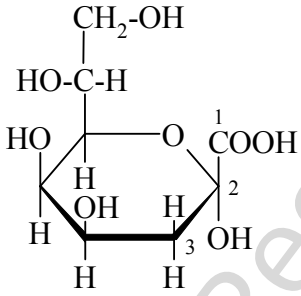
P-102.5.6.6.3 Ketoaldonic acids

Ketoaldonic acids are oxo carboxylic acids formally derived from aldonic acids by oxidation of a secondary –CHOH group to a carbonyl group. Names of individual ketoaldonic acids are formed by changing the ending 'ulose' in the name of the corresponding ketose to 'ulonic acid', preceded by the locant of the ketonic group. The numbering starts at the carboxy group.

Examples:



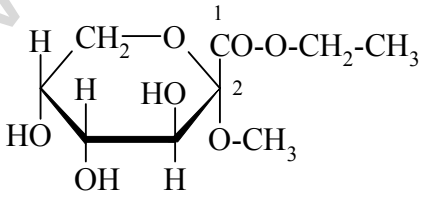
2,3,4,6-tetra-*O*-acetyl- D-*arabino*-hex-5-ulosonic acid
(1*S*,2*R*,3*S*)-2,3,4-tris(acetyloxy)-5-oxohexanoic acid (PIN)



3-deoxy-α-D-*manno*-oct-2-ulopyranosonic acid
(2*R*,4*R*,5*R*,6*R*)-6-[(1*R*)-1,2-dihydroxyethyl]-2,4,5-trihydroxyoxane-1-carboxylic acid (PIN)

Glycosides are named by changing the component 'pyranose' into 'pyranoside' in the name, to give '-ulopyranosidonic acid'. Names of derivatives of ketoaldonic acids are formed as described in P-102.5.6.6.2.1 for aldonic acids. When a glycoside is esterified, parentheses are used to isolate the glycosidic portion of the name.

Example:

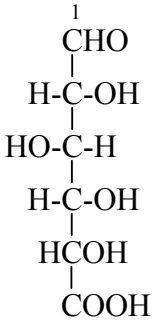


ethyl (methyl α-D-*arabino*-hex-2-ulopyranosid)onate
ethyl (2*R*,3*S*,4R,5*R*)-3,4,5-trihydroxy-2-methoxyoxane-2-carboxylate (PIN)

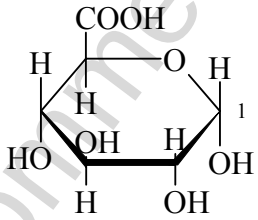
P-102.5.6.6.4 Uronic acids

P-102.5.6.6.4.1 Uronic acids are carboxylic acids formally derived from aldoses by oxidation of the terminal $–CH_2OH$ group to a carboxy group. Names of individual uronic acids are formed by changing the ending 'ose' in the retained or systematic name of the corresponding aldose to 'uronic acid'. The numbering of the aldose is kept intact; the locant '1' is still assigned to the (potential) aldehydic group.

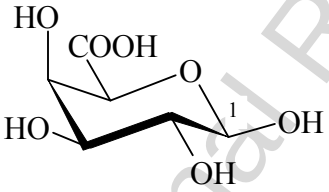
Examples:



D-glucuronic acid (PIN)



α-D-glucopyranuronic acid (PIN)

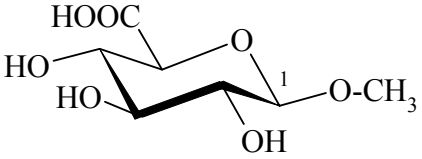


β-D-galactopyranuronic acid (PIN)

P-102.5.6.6.4.2 Glycosides of uronic acids

Names of glycosides of uronic acids are formed by changing the 'pyran' component in the name of the acid to 'pyranoside', with elision of the final letter 'e', to give 'pyranosiduronic acid'.

Example:

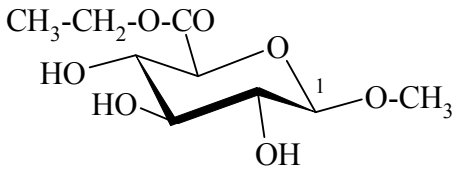


methyl β-D-glucopyranosiduronic acid (PIN)

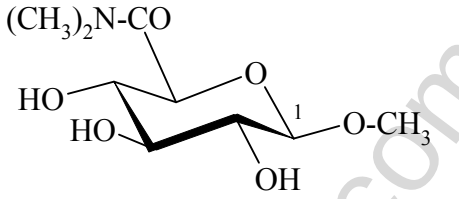
P-102.5.6.6.4.3 Derivatives of uronic acids

Names of derivatives are formed as indicated in P-102.5.6.6.2.1 and P-46.

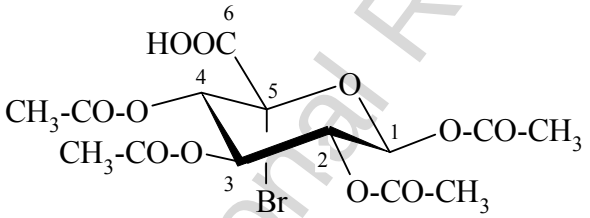
Examples:



ethyl (methyl β-D-glucopyranosid)uronate (PIN)



N,*N*-dimethyl(methyl β-D-glucopyranosid)uronamide (PIN)

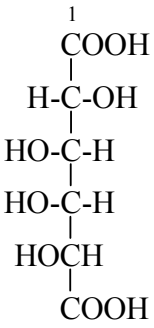


1,2,3,4-tetra-*O*-acetyl-5-bromo-β-L-idopyranuronic acid
(5*R*)-1,2,3,4-tetra-*O*-acetyl-5-bromo-α-D-*xylo*-hexopyranuronic acid
(2*R*,3*S*,4*R*,5*R*,6*R*)-2-bromo-3,4,5,6-tetra-(acetyloxy)oxane-2-carboxylic acid (PIN)

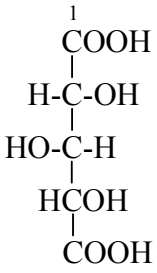
P-102.5.6.6.5 Aldaric acids

P-102.5.6.6.5.1 Aldaric acids are carboxylic acids formed by the oxidation of both terminal groups (–CHO and –CH_2OH) of aldoses to carboxy groups. Names of aldaric acids are formed by changing the 'ose' ending in retained or systematic names of parent aldoses to 'aric acid'. Choice of a parent structure is made in accordance with P-102.5.5.5.1. The stereodescriptor '*meso*' must be added for sake of clarity to the names of the appropriate aldaric acids.

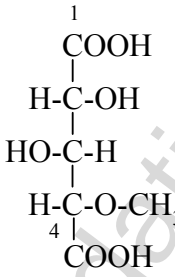
Examples:



L-altraric acid (PIN)
(not L-talaric acid)



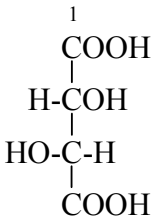
meso-xylaric acid (PIN)



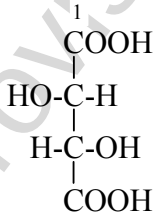
4-*O*-methyl-D-xylaric acid (PIN)
(not 2-*O*-methyl-L-xylaric acid)

P-102.5.6.6.5.2 Stereodescriptors for tartaric acids

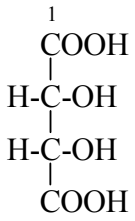
Tartaric acid is the retained name to describe the aldaric acids corresponding to the parent aldoses, erythrose and threose. '*R*' and '*S*' are preferred stereodescriptors for denoting the configuration of tartaric acid. Salts and esters are referred to as tartrates.



(2*R*,3*R*)-2,3-dihydroxybutanedioic acid (PIN)
(2*R*,3*R*)-tartaric acid
L-threaric acid
(+)-tartaric acid



(2*S*,3*S*)-2,3-dihydroxybutanedioic acid (PIN)
(2*S*,3*S*)-tartaric acid
D-threaric acid
(-)-tartaric acid

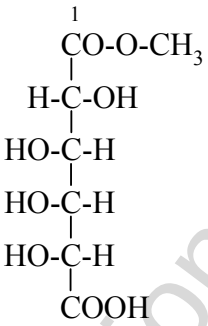


(2*R*,3*S*)-2,3-dihydroxybutanedioic acid (PIN)
(2*R*,3*S*)-tartaric acid
erythraric acid
meso-tartaric acid

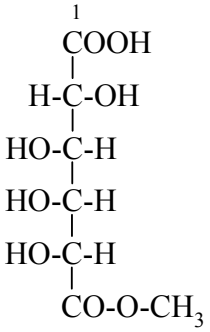
P-102.5.6.6.5.3 Derivatives of aldaric acids

Derivatives of aldaric acids formed by modifying the carboxy group (into esters, amides, hydrazides, nitriles, amic acids, etc.) are named by the methods described in P-102.5.6.6.2.1 and P-65.

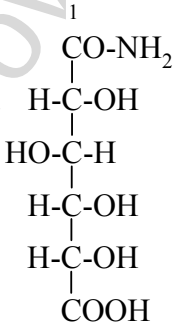
Examples:



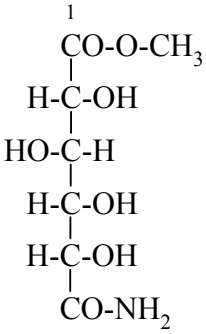
1-methyl hydrogen L-altarate (PIN)



6-methyl hydrogen L-altarate (PIN)



D-glucar-1-amic acid (PIN)



1-methyl D-glucar-6-amate (PIN)

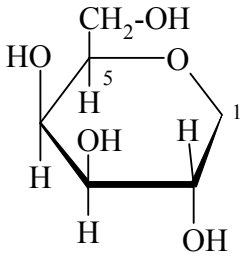
P-102.5.6.7 Anhydrides

Anhydrides are intramolecular or intermolecular derivatives of monosaccharides.

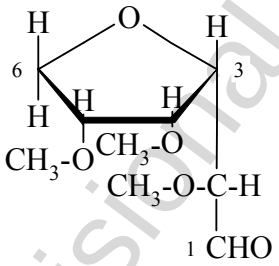
P-102.5.6.7.1 Intramolecular anhydrides

An intramolecular ether (commonly called intramolecular anhydride), formally arising by elimination of water from two hydroxyl groups of a single molecule of a monosaccharide (aldose, ketose) or monosaccharide derivative, is named by adding the detachable prefix 'anhydro', preceded by a pair of locants identifying the two hydroxy groups, to the name of the monosaccharide.

Examples:



1,5-anhydro-D-galactitol
(2*R*,3*R*,*4R*,5*S*)-6-(hydroxymethyl)oxane-3,4,5-triol (PIN)

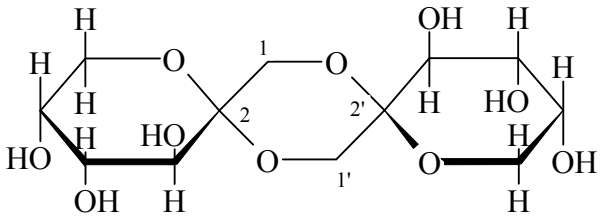


3,6-anhydro-2,4,5-tri-*O*-methyl- D-glucose
(2*R*)-2-[(2*S*,3*R*,4*R*)-dimethoxyoxolan-2-yl]-2-methoxyacetaldehyde (PIN)

P-102.5.6.7.2 Intermolecular anhydrides

The cyclic product of condensation of two monosaccharide molecules with elimination of two molecules of water (commonly called an intermolecular anhydride) is named by placing the term 'dianhydride' after the names of the two parent monosaccharides. When the two parents are different, the senior parent according to the selection criteria for selecting the parent structure (see P-102.4), is cited first. The position of each anhydride link is indicated by a pair of locants showing the position of the two hydroxy groups involved, the locants relating to one monosaccharide (in a mixed anhydride, the second monosaccharide named) are primed. The pair of locants immediately precedes the term 'dianhydride'.

Example:



α-D-fructopyranose β-D-fructopyranose 1,2′:1′,2-dianhydride
(α-D-fructopyranose is cited first; according to P-102.4 (c), α precedes β)
(3*R*,4*R*,5*S*,6*R*,9*S*,12*R*,13*S*,14*S*)-1,7,10,15-tetraoxadispiro[5.2.5^{9}.2^{6}]hexadecane-3,4,5,12,13,14-hexol (PIN)

P-102.6 Monosaccharides and derivatives as substituent groups

P-102.6.1 Glycosyl groups
P-102.6.2 Monosaccharides as substituent groups

P-102.6.1 Glycosyl groups

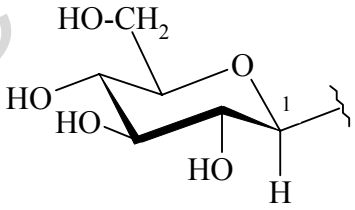
P-102.6.1.1 Glycosyl groups
P-102.6.1.2 *O*-Glycosyl compounds
P-102.6.1.3 *N*-Glycosylamines
P-102-6.1.4 *C*-Glycosyl compounds
P-102.6.1.5 Glycosyl halides, pseudohalides and esters
P-102.6.1.6 Substituent groups other than glycosyl groups

P-102.6.1.1 Glycosyl groups

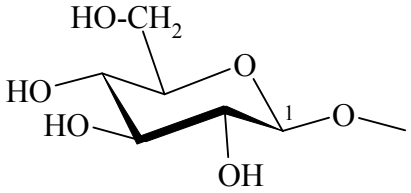
Substituent groups formed by removal of the anomeric hydroxy group from a cyclic monosaccharide are named by replacing the final letter ‘e’ of the monosaccharide name by ‘yl’. The term ‘glycosyl residue’ is used in the nomenclature of carbohydrates. Terms of this nature are widely used in naming glycosides, when they are not the parent structures, and oligosaccharides.

No locant is added to the name of the substituent to indicate the position of the free valence. A sinuous line denotes the free valence, as recommended for cyclic substituent groups in systematic nomenclature.

Examples:



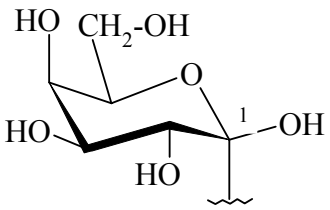
β-D-glucopyranosyl (PIN)
(the hydrogen atom at position 1 is shown)



β-D-glucopyranosyloxy (PIN)

When the free valence is formed at carbon '1' by subtraction of a hydrogen atom, the substituent group is named as a glycosyl group but the presence of the hydroxy group is denoted by substitution at carbon '1'. In this case, the stereodescriptor 'α' or 'β' refers to the free valence, not to the –OH group.

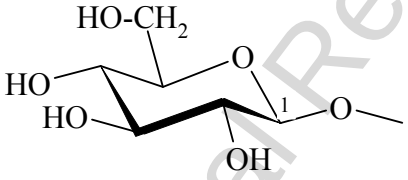
Example:



1-hydroxy-α-D-galactopyranosyl (PIN)

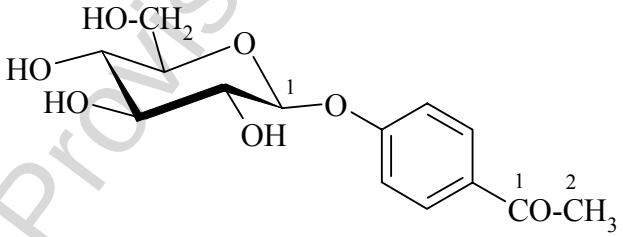
P-102.6.1.2 *O*-Glycosyl compounds

The substituent group formed by removal of a hydrogen atom from the anomeric –OH group is considered as a compound substituent group formed by the 'glycosyl' group and an 'oxy' group.

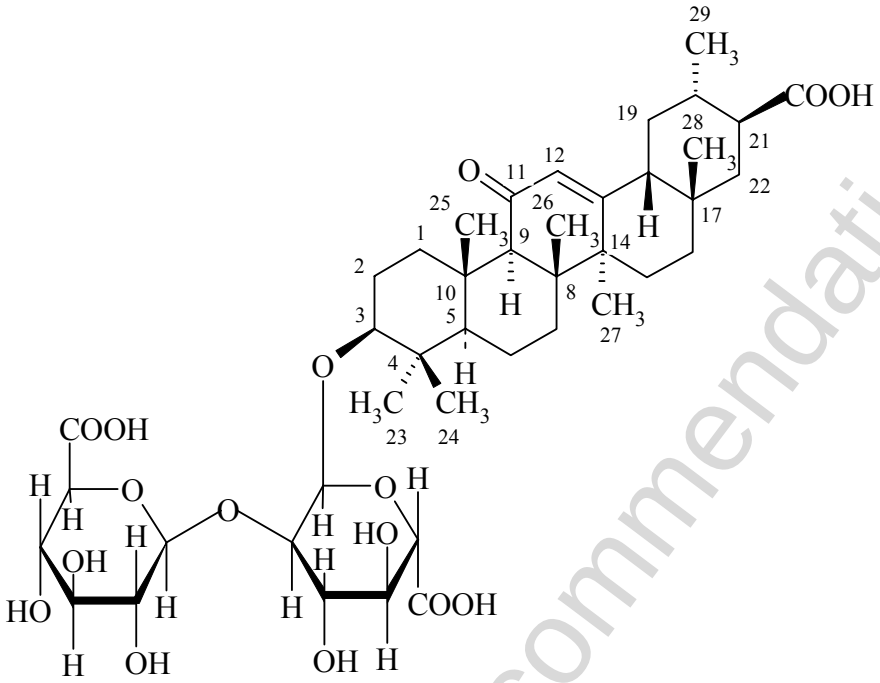


β-D-glucopyranosyloxy (PIN)

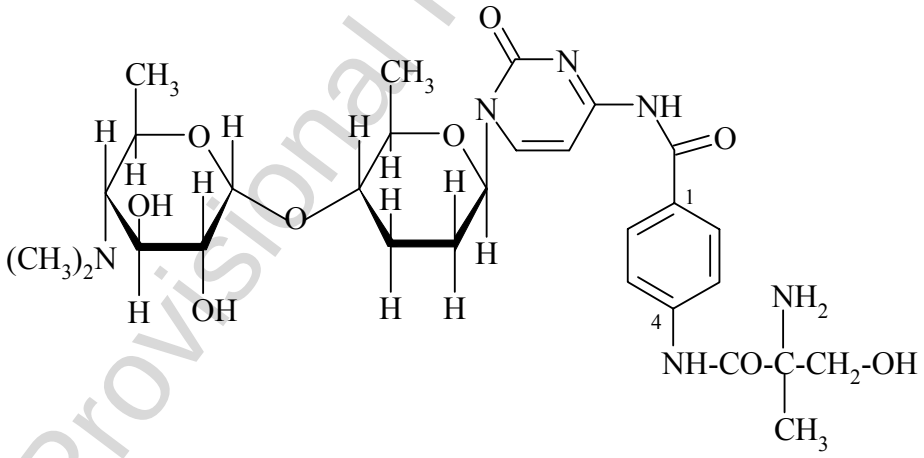
Examples:



1-[4-(β-D-glucopyranosyloxy)phenyl]ethan-1-one (PIN)
(not 4′-(β-D-glucopyranosyloxy)acetophenone; acetophenone is no longer recommended as a preferred IUPAC name)
(not 4-acetylphenyl β-D-glucopyranoside; a ketone is senior to a hydroxy compound)



21β-carboxy-11-oxo-30-norolean-12-en-3β-yl (2-*O*-β-D-glucopyranosyluronic acid)-α-D-glucopyranosiduronic acid (PIN)



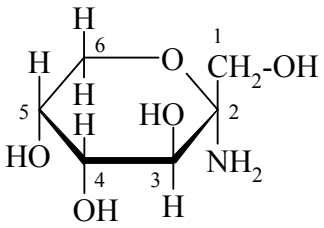
4-[(*RS*)-2-amino-3-hydroxy-2-methylpropanamido]-*N*-(1-{5-[(4,6-dideoxy-4-(dimethylamino)-α-D-glucopyranosyl)oxy]-1-[(2*R*,5*S*,6*R*)-6-methyloxan-2-yl]}-2-oxo-1,2-dihydropyrimidin-4-yl)benzamide (PIN)

Guide to name construction: the principal function is an amide; the cyclic amide, benzamide, is senior to the acyclic amide, propanamide)

P-102.6.1.3 *N*-Glycosyl compounds (glycosylamines)

N-Glycosyl derivatives are conveniently named as glycosylamines.

Example:

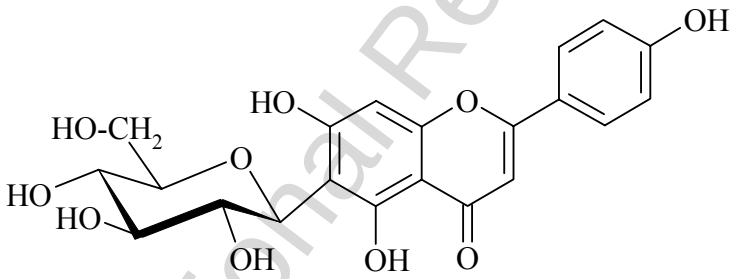


α-D-fructopyranosylamine (PIN)

P-102.6.1.4 *C*-glycosyl compounds

Compounds arising formally from the elimination of water from the glycosidic hydroxy group and a hydrogen atom bound to a carbon atom (thus creating a C-C bond) are named using the appropriate glycosyl group.

Examples:

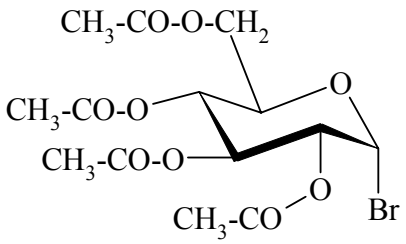


6-(β-D-glucopyranosyl)-5,7-dihydroxy-2-(4-hydroxyphenyl)-4*H*-chromen-4-one (PIN)
6-(β-D-glucopyranosyl)-4′,5,7-dihydroxyflavone

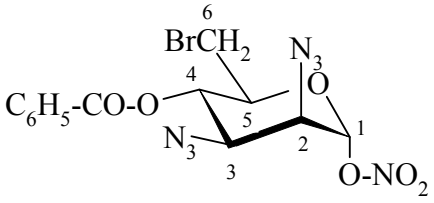
P-102.6.1.5 Glycosyl halides, pseudohalides and esters

Glycosyl halides and pseudohalides are named by using functional class nomenclature, by adding, as a separate word, the class name ‘chloride’, ‘isocyanate’, etc. to the name of the appropriate glycosyl group. Esters in position 1 of oxoacids other than *P* and *S* acids are treated like glycosyl halides and pseudohalides.

Example:



2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl bromide (PIN)

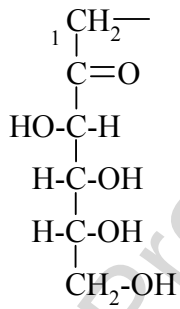


2,3-diazido-4-*O*-benzoyl-6-bromo-2,3,6-trideoxy-α-D-mannopyranosyl nitrate (PIN)

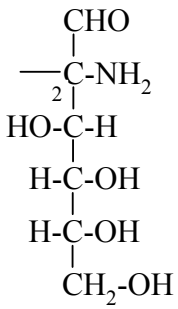
P-102.6.1.6 Substituent groups other than glycosyl groups

A hydrogen atom may be removed from any position of a monosaccharide other than C-1. This formation of a free valence is denoted by the suffix 'yl', but a locant is necessary to indicate the position of the free valence and to distinguish such a name from that of glycosyl substituents for which the locant 1 is omitted. These prefixes can be formed by replacing the final letter 'e' of the systematic or trivial name of a monosaccharide by *n*-*C*-yl, *n*-*O*-yl. The symbol '*C*' is omitted when the free valence is derived from a position at which hydrogen atoms only are attached.

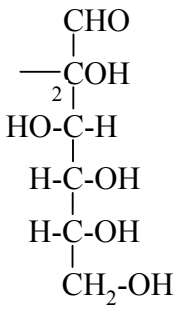
Examples:



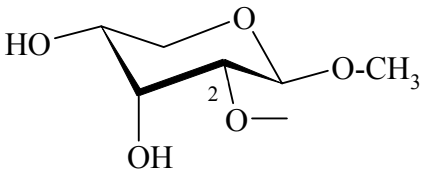
1-deoxy-D-fructos-1-yl (PIN)



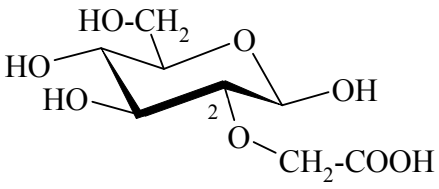
2-amino-2-deoxy-D-glucos-2-*C*-yl (PIN)



D-glucos-2-*C*-yl (PIN)



methyl β-D-ribopyranosid-2-*O*-yl (PIN)



2-(β-D-glucopyranos-2-*O*-yl)acetic acid (PIN)
(not 2-*O*-carboxymethyl β-D-glucopyranose; this name is not constructed in conformity with P-102.4.3(a) for selecting a parent structure; a carboxylic acid is senior to a hydroxy compound).

P-102.7 Disaccharides and oligosaccharides

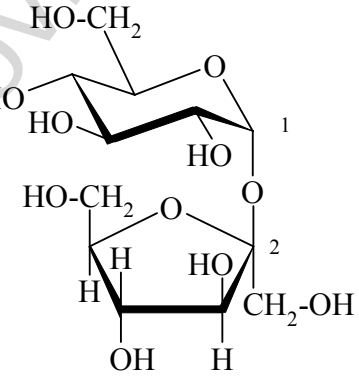
Names of disaccharides and oligosaccharides are formed by the principles, rules and conventions described above for monosaccharides. As names may become very long and cumbersome, full systematic names have been replaced by a system based on symbols, such as Glc for glucose. A brief description of the extended form, the condensed and the short form is given in this Section.

P-102.7.1 Disaccharides
P-102.7.2 Oligosaccharides

P-102.7.1 Disaccharides

P-102.7.1.1 Disaccharides without a free hemiacetal group

Disaccharides which can be regarded as formed by elimination of one molecule of water from two glycosidic (anomeric) hydroxyl groups, are named as glycosyl glycosides. The parent (cited as the 'glycoside') is chosen in accordance with criteria described in P-102.4. Both anomeric descriptors must be cited in the name.
Example:



β-D-fructofuranosyl α-D-glucopyranoside (PIN)
(not α-D-glucopyranosyl β-D-fructofuranoside; gluco precedes fructo in the alphabetical order)
sucrose (trivial name)

P-102.7.1.2 Disaccharides with a free hemiacetal group

Disaccharides which can be regarded as formed by elimination of one molecule of water from one glycosidic (anomeric) hydroxyl group and one alcoholic hydroxyl group, are named as glycosylglycoses. Locants and anomeric descriptors must be cited in the full name.

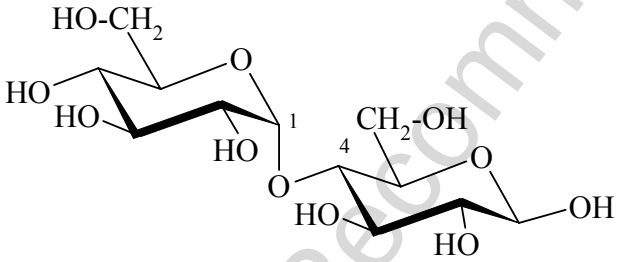
There are two established methods for citing locants:

(1) in parentheses between the components with an arrow going from the locant of the glycosyl component to that of the glycose component;

(2) at the front of the glycosyl component.

Method (1) leads to preferred IUPAC names.

Example:



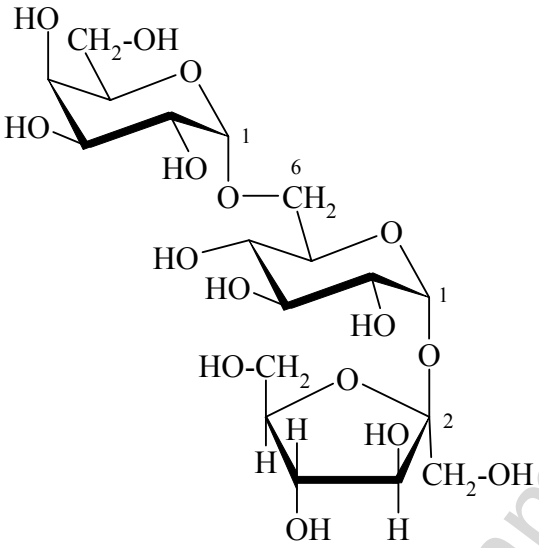
(1) α-D-glucopyranosyl-(1→4)-β-D-glucopyranose (PIN)
(2) 4-*O*-α-D-glucopyranosyl-β-D-glucopyranose
β-maltose (trivial name; not β-D-maltose)

P-102.7.2 Oligosaccharides

P-102.7.2.1 Oligosaccharides without a free hemiacetal group

A trisaccharide, for example, is named as a glycosylglycosyl glycoside or glycosyl glycosylglycoside as required. A choice between the two residues linked through their anomeric positions for citation as the 'glycoside' portion can be made on the basis of P-102.4. Alternatively, a sequential (end-to-end) naming approach may be used, regardless of P-102.4. The name is formed by the preferred method for naming disaccharides.

Example:

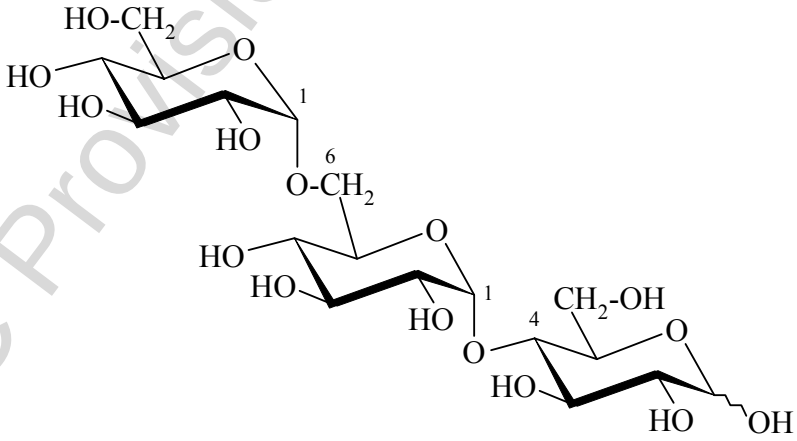


β-D-fructofuranosyl α-D-galactopyranosyl-(1→6)-α-D-glucopyranoside (PIN)
(glucose, not fructose, is selected as the 'glycoside')
α-D-galactopyranosyl-(1→6)-α-D-glucopyranoside β-D-fructofuranoside
(sequential method)
raffinose (trivial name)

P-102.7.2.2 Oligosaccharides with a free hemiacetal group

An oligosaccharide of this type is name as a glycosyl[glycosyl]$_n$glycose, the 'glycose' portion being the parent. The conventional depiction has the 'glycose' portion on the right. Names are formed as described in P-102.7.2.1.

Example:



α-D-glucoyranosyl-(1→6)-α-D-glucopyranosyl-(1→4)-D-glucopyranose (PIN)
panose (trivial name)

P-103 Amino acids and peptides

P-103.0 Introduction

This Section describes the nomenclature of amino acids that constitute the building blocks of peptides and proteins. They are functional parents having retained names listed in Table 10.3. Less common amino acids also have retained names. The nomenclature of amino acids is composed of two types of names: names based on retained names for functional parents, with a limited capacity of functionalization and substitution, and systematic substitutive names for all other compounds.

The nomenclature of these amino acids is described in the document entitled 'Nomenclature and Symbolism for Amino Acids and Peptides' (ref. 23). In this Section, the nomenclature of these amino acids is restricted to their derivatives outside the field of peptides and proteins and their use in the construction of preferred IUPAC names.

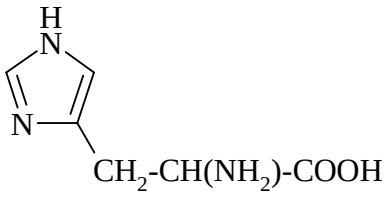
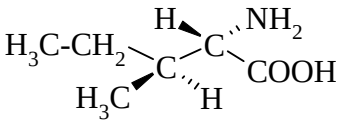
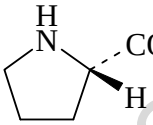
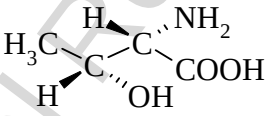
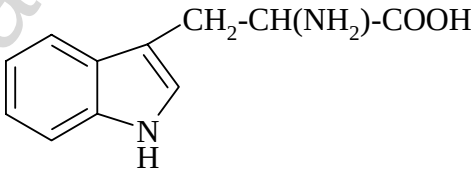
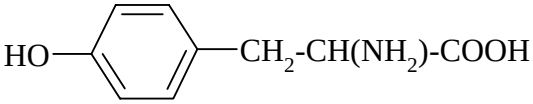
P-103.1 Nomenclature based on retained names
P-103.2 Substitutive names of amino acids
P-103.3 Nomenclature of peptides

P-103.1 Nomenclature based on retained names (see Tables 3 and 4)

P-103.1.1 Configuration of α-amino carboxylic acids
P-103.1.2 Functional modifications of α-amino carboxylic acids
P-103.3 Derivatives of α-amino carboxylic acids
P-103.4 α-Amino carboxylic acids as substituent groups

Table 10.3. Retained names of α-amino acids

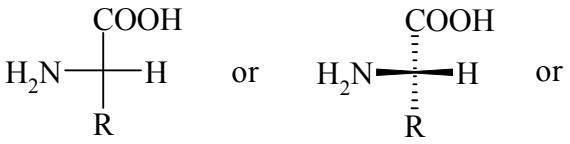
Retained name	*Symbols*		*Formula*
alanine (PIN)	Ala	A	CH_3-CH(NH_2)-COOH
arginine (PIN)	Arg	R	H_2N-C(=NH)-NH-[CH_2]$_3$-CH(NH_2)-COOH
asparagine (PIN)	Asn	N	H_2N-CO-CH_2-CH(NH_2)-COOH
aspartic acid (PIN)	Asp	D	HOOC-CH_2-CH(NH_2)-COOH
cysteine (PIN)	Cys	C	HS-CH_2-CH(NH_2)-COOH
glutamine (PIN)	Gln	Q	H_2N-CO-[CH_2]$_2$-CH(NH_2)-COOH
glutamic acid(PIN)	Glu	E	HOOC-[CH_2]$_2$-CH(NH_2)-COOH
glycine (PIN)	Gly	G	H_2N-CH_2-COOH

histidine (PIN)	His	H	
isoleucine* (PIN)	Ile	I	
leucine (PIN)	Leu	L	$(CH_3)_2CH\text{-}CH(NH_2)\text{-}COOH$
lysine (PIN)	Lys	K	$H_2N\text{-}[CH_2]_4\text{-}CH(NH_2)\text{-}COOH$
methionine (PIN)	Met	M	$CH_3\text{-}S\text{-}[CH_2]_2\text{-}CH(NH_2)\text{-}COOH$
phenylalanine (PIN)	Phe	F	$C_6H_5\text{-}CH_2\text{-}CH(NH_2)\text{-}COOH$
proline* (PIN)	Pro	P	
serine (PIN)	Ser	S	$HO\text{-}CH_2\text{-}CH(NH_2)\text{-}COOH$
threonine* (PIN)	Thr	T	
tryptophan (PIN)	Trp	W	
tyrosine (PIN)	Tyr	T	
valine (PIN)	Val	V	$(CH_3)_2CH\text{-}CH(NH_2)\text{-}COOH$
unspecified amino acid	Xaa	X	

* L-forms shown

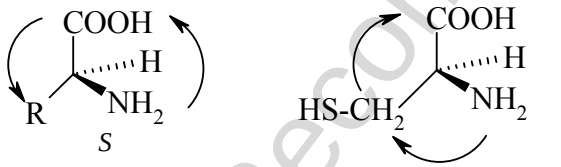
P-103.1.1 Configuration of α-amino carboxylic acids

The absolute configuration at the α-carbon atom of the α-amino carboxylic acids is designated by the stereodescriptor '*D*' or '*L*' to indicate a formal relationship to '*D*' or '*L*' glyceraldehyde. The stereodescriptor 'ξ' (Greek letter xi) indicates unknown configuration.



Three tridimensional representations of the '*L*' configuration
(The '*L*' configuration of isoleucine, proline and threonine is shown in Table 10.3)

The '*L*' configuration corresponds to the '*S*' configuration in the CIP system, except for cysteine that has the '*R*' configuration..



P-103.1.2 Functional modifications of α-amino carboxylic acids

Retained names are use to form preferred IUPAC names of salts and esters, and those of alkyl, aryl and acyl derivatives substituted on *N*, *O*, and *S* atoms.

The carboxy group –COOH can be transformed into various characteristic groups such as a hydroxymethyl group, $–CH_2-OH$, or an aldehyde group, –CHO. Some names derived from retained amino acid names are recommended to be used for naming acyl groups, amides, alcohols, aldehydes, and even ketones, in the context of peptide and protein nomenclature. They can be used in general nomenclature, but not as preferred IUPAC names that must be systematically constructed as indicated in P-103.2.

P-103.1.2.1 Ionization of characteristic groups

Names of zwitterions, anions and cations are derived from the rules expressed in Chapter 7.

Examples:

$NH_3^+-CH_2-COO^-$	2-azaniumylacetate (PIN) glycine zwitterion
$H_2N-CH_2-COO^-$	glycinate (PIN) glycine anion

NH_3^+-CH_2-COOH	glycinium (PIN) glycine cation

Specific names are recommended for the monoanion and the dianion of aspartic acid and glutamic acid:

$^-$OOC-CH_2-CH_2-CH(NH_3^+)-COO^-	hydrogen glutamate (PIN) glutamate(1-)
$^-$OOC-CH_2-CH_2-CH(NH_3^+)-COO^- Na^+	sodium hydrogen glutamate (PIN) sodium glutamate(1-)
$^-$OOC-CH_2-CH_2-CH(NH_2)-COO^-	glutamate (PIN) (by definition glutamate = 2−) glutamate(2-) (see 3AA-6, ref. 23) glutamic acid dianion
$^-$OOC-CH_2-CH_2-CH(NH_3^+)-COO^- $2Na^+$	disodium glutamate (PIN)
NH_3^+-$[CH_2]_4$-CH(NH_3^+)-COO^-	lysinium (PIN) (by definition "-ium" = 1+) lysinium(1+) (see 3AA, ref. 23) lysine monocation
NH_3^+-$[CH_2]_4$-CH(NH_3^+)-COO^- Cl^-	lysinium chloride (PIN) (by definition "-ium" = 1+) lysinium(1+) chloride (see 3AA, ref. 23) lysine monohydrochloride)
NH_3^+-$[CH_2]_4$-CH(NH_3^+)-COOH	lysinediium (PIN) lysinium(2+) (see 3AA, ref. 23)

P-103.1.2.2 Acyl groups

Preferred IUPAC names of acyl groups , H_2N-CHR-CO–, are formed by changing the ending 'ine' (or 'an' in tryptophan) into 'yl', for example alanyl, valyl, tryptophyl. 'Cysteinyl' is used instead of 'cysteyl'; 'cystyl' is derived from 'cystine'.

The following names are used to name the acyl groups derived from dicarboxylic amino acids and their corresponding amides.

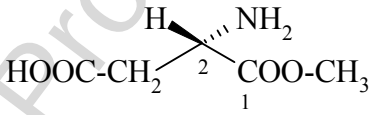
HOOC-CH_2-CH(NH_2)-CO–	α-aspartyl (PIN) aspart-1-yl
–CO-CH_2-CH(NH_2)-COOH	β-aspartyl (PIN) aspart-4-yl
–CO-CH_2-CH(NH_2)-CO–	aspartoyl
HOOC-CH_2-CH_2-CH(NH_2)-CO–	α-glutamyl (PIN) glutam-1-yl
–CO-CH_2-CH_2-CH(NH_2)-COOH	γ-glutamyl glutam-5-yl (PIN)
–CO-CH_2-CH_2-CH(NH_2)-CO–	glutamoyl
H_2N-CO-CH_2-CH(NH_2)-CO–	asparaginyl
H_2N-CO-CH_2-CH_2-CH(NH_2)-CO–	glutaminyl

P-103.1.2.3 Esters

Preferred IUPAC names of esters, R-CO-OR′, are formed by the general method of using the 'ate' ending obtained by replacing the 'ic acid' ending or the final letter 'e' of the retained name (or adding the ending 'ate' to the name tryptophan) and the name of the substituent group R′.

Example:

methyl L-alaninate (PIN)
alanine methyl ester

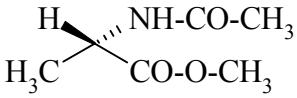


1-methyl L-aspartate (PIN)

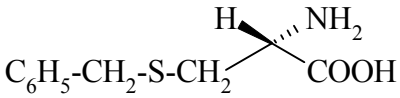
P-103.1.2.4 Nitrogen, oxygen and sulfur substituted amino acids

Retained names are used to indicate nitrogen, oxygen and sulfur substitution. Locants *N*, *O* and *S*, as appropriate, indicate the location of the substitution. For lysine, locants N^2 and N^6 are used to denote the two amino groups located at positions 2 and 6, respectively.

Examples:



methyl *N*-acetyl-L-alaninate (PIN)

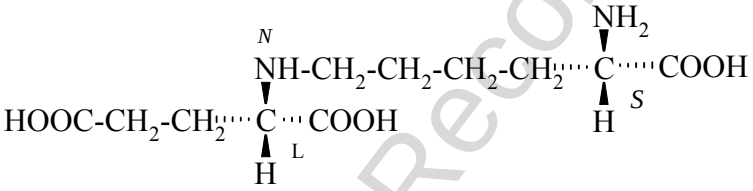


S-benzyl-L-cysteine (PIN)

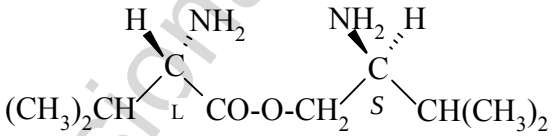
P-103.1.2.5 Addition of chirality centers

By substitution, added atoms or groups may add new chirality centers. Preferred IUPAC names are formed by using '*R*' and '*S*' stereodescriptors for all new chirality centers.

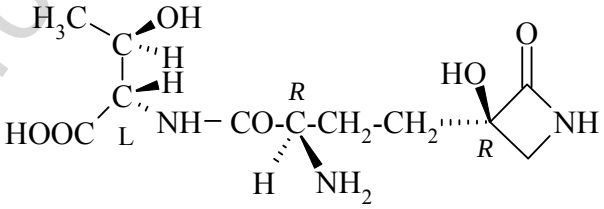
Examples:



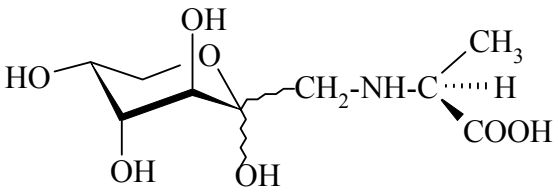
N-[(2*S*)-(5-amino-5-carboxypentyl)]-L-glutamic acid (PIN)



(2*S*)-2-amino-3-methylbutyl L-valinate (PIN)



N-{(2*R*)-2-amino-4-[(3*R*)-3-hydroxy-2-oxoazetidin-3-yl]butanoyl}-L-threonine (PIN)

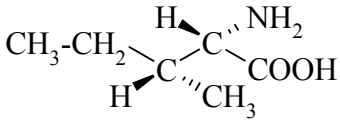
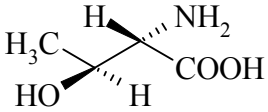


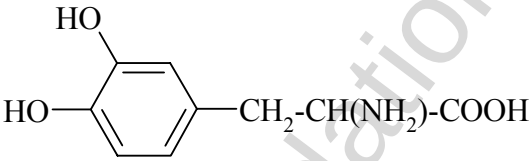
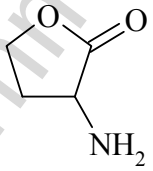
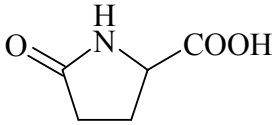
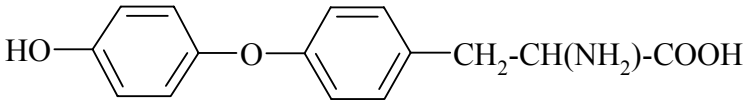
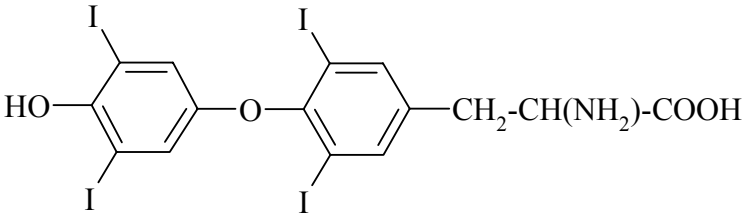
N-(1-deoxy-D-fructopyranos-1-yl)-L-alanine (PIN)

P-103.1.2.6 Other retained names

Several other trivial names are used in addition to those listed in Table 10.3. A few of them and their symbols are described in Table 10.4. They are used as preferred IUPAC names; they can be functionalized and substituted as shown in sections P-1.2. Stereodescriptors D, L and DL are used as required. The L configuration of alloisoleucine and allothreonine is shown. The publication 'Nomenclature and symbolism for amino acids and peptides' (ref. 23) must be consulted for the complete description of the use of less common amino acids.

Table 10.4 Amino acids with trivial names (other than those listed in Table 10.3)

Retained name	*Symbol*	*Structure*
β-alanine (PIN)	(βAla)	$H_2N\text{-}CH_2\text{-}CH_2\text{-}COOH$
alloisoleucine (PIN)	aIle	
allothreonine (PIN)	aThr	
allysine (PIN)	---	$HCO\text{-}[CH_2]_3\text{-}CH(NH_2)\text{-}COOH$
citrulline (PIN)	Cit	$NH_2\text{-}CO\text{-}NH\text{-}[CH_2]_3\text{-}CH(NH_2)\text{-}COOH$
cystathionine (PIN)	Al | HCy	$CH_2\text{-}CH(NH_2)\text{-}COOH$ | $S\text{-}[CH_2]_3\text{-}CH(NH_2)\text{-}COOH$
cysteic acid (PIN)	Cya	$^{-}O_3S\text{-}CH_2\text{-}CH(NH_2)\text{-}COOH$

cystine (PIN)	Cys | Cys	S-CH_2-CH(NH_2)-COOH | S-CH_2-CH(NH_2)-COOH
dopa (PIN)	----	
homocysteine (PIN)	Hcy	HS-CH_2-CH_2-CH(NH_2)-COOH
homoserine (PIN)	Hse	HO-CH_2-CH_2-CH(NH_2)-COOH
homoserine lactone (PIN)	Hsl	
lanthionine (PIN)	Ala | Cys	CH_2-CH(NH_2)-COOH | S-CH_2-CH(NH_2)-COOH
ornithine (PIN)	Orn	H_2N-[CH_2]$_3$-CH(NH_2)-COOH
5-oxoproline (PIN)	Glp	
sarcosine(PIN)	Sar	CH_3-NH-CH(NH_2)-COOH
thyronine (PIN)	-----	
thyroxine (PIN)	Thx	

P-103.2 Substitutive names of amino acids and derivatives

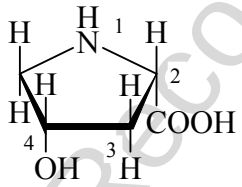
Names of derivatives of amino acids described in P-103.1 other than salts, esters and formed by *N*, *O* or *S* substitution are formed substitutively. Preferred names are constructed in accordance with principles, rules and conventions recommended for the formation of systematic substitutive names.

P-103.2.1 Carbon substitution
P-103.2.2 Substituent groups
P-103.2.3 Amides, anilides and hydrazides
P-103.2.4 Aldehydes, alcohols and ketones

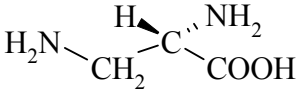
P-103.2.1 Carbon substitution

Preferred IUPAC names of amino acids substituted on carbon atoms are systematic substitutive names. CIP stereodescriptors are used.

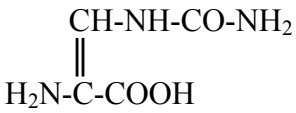
Examples:



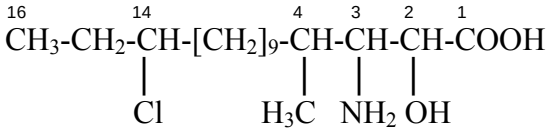
pyrrolidine-2-carboxylic acid (PIN)
(2*S*,4*S*)-4-hydroxyproline (PIN)
(4*S*)-4-hydroxy-L-proline
cis-4-hydroxy-L-proline



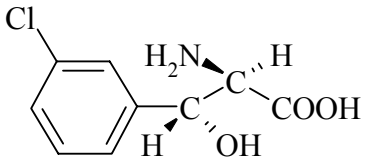
(2*S*)-2,3-diaminopropanoic acid (PIN)
L-2,3-diaminopropanoic acid
(2*S*)-2-amino-β-alanine (see ref. 49)
3-amino-L-alanine



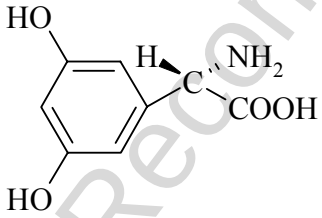
(2*E*)-3-(carbamoylamino)-2-aminoprop-2-enoic acid (PIN)
(2*E*)-2,3-didehydro-3-ureido-L-alanine (see ref. 49)



(2ξ,3ξ,4ξ,14ξ)-3-amino-14-chloro-2-hydroxy-4-methylhexadecanoic acid (PIN)
3-[(1ξ,14ξ)-11-chloro-1-methyltridecyl]-2-hydroxy-β-alanine



(2*R*,3*R*)-2-amino-3-(3-chlorophenyl)-3-hydroxypropanoic acid (PIN)
3-chloro-(*R*)-β-hydroxy-D-tyrosine



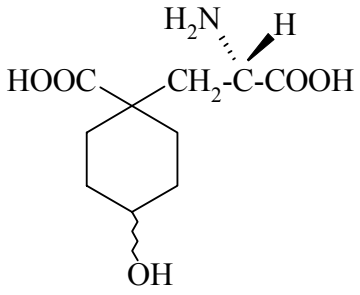
(2*S*)-2-amino-3-(3,5-dihydroxyphenyl)acetic acid (PIN)
L-(3,5-dihydroxyphenyl)glycine

P-103.2.2 Substituent groups

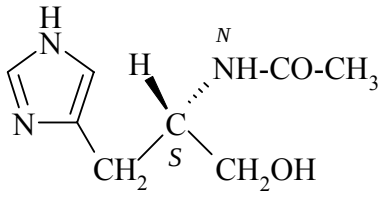
P-103.2.1.1 Substituent groups with the free valence on a carbon atom

Preferred IUPAC names are formed by the rules, principles and conventions of substitutive nomenclature when α-amino carboxylic acids must be cited as substituent groups in presence of characteristic groups having seniority for citation as suffix.

Examples:



1-[(2*S*)-2-amino-2-carboxyethyl]-4ξ-hydroxycyclohexane-1-carboxylic acid (PIN)



N-[(2*S*)-1-hydroxy-3-(1*H*-imidazol-4-yl)propan-2-yl]acetamide (PIN)

P-103.2.1.2 Substituent groups with the free valence on a nitrogen atom

Preferred IUPAC names are formed by the rules, principles and conventions of substitutive nomenclature when α-amino carboxylic acids must be cited as substituent groups in presence of characteristic groups having seniority for citation as suffix.

Substituent groups derived by subtracting a hydrogen atom from the amino group of an amino acid may also named by changing the ending 'e' into 'o' in names of appropriate amino acids, by adding the letter 'o' to tryptophan and by constructing the names asparto and glutamo, from aspartic acid and glutamic acid, respectively (see 3AA-7, ref. 23).

Examples:

–HN-CH_2-COOH	glycino (carboxymethyl)amino (PIN)

When there is more than one nitrogen atom in the amino acid, the use of a locant of the form N^x with a superscript is recommended.

Examples:

–HN-$[CH_2]_4$-$CH(NH_2)$-COOH

N^6-lysino
(4-amino-5-carboxypentyl)amino (PIN)

–HN-C(=NH)-NH-$[CH_2]_3$-$CH(NH_2)$-COOH

N^{ω}-arginino
N'-[(3-amino-4-carboxybutyl)amino]carbamimidamido (PIN) (see P-66.1.5.4.2)

–HN-CO-$[CH_2]_2$-$CH(NH_2)$-COOH

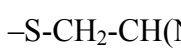
N^5-glutamino
(4-amino-4-carboxy)butanamido (PIN)

P-103.2.1.3 Substituent groups with the free valence on an oxygen or sulfur atom

Preferred IUPAC names are formed by the rules, principles and conventions of substitutive nomenclature when α-amino carboxylic acids must be cited as substituent groups in presence of characteristic groups having seniority for citation as suffix.

Substituent groups formed by subtraction of a hydrogen atom from an oxygen or sulfur atom may also be named by changing the final letter 'e', when appropriate, into '*x*-yl', *x* being the locant of the atom from which the hydrogen atom has been subtracted, for example cystein-*S*-yl, threonin-O^3-yl, alanin-3-yl, or by adding '*x*-yl' to aspartic, glutamic and tryptophan, for example aspartic-2-yl, tryptophan-2-yl.

Example:



–S-CH_2-CH(NH_2)-COOH

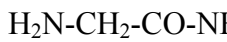
cystein-*S*-yl
(2-amino-2-carboxyethyl)sulfanyl (PIN)

P-103.2.3 Amides, anilides, hydrazides and analogous derivatives

Preferred IUPAC names of amides, anilides, hydrazides and analogous derivatives derived from amino acids are systematic substitutive names.

Names of amides derived from amino acids are formed by changing the final letter 'e' in the names of amino acids, when appropriate, into 'amide' or adding the term 'amide' to the name tryptophan.

Examples:



H_2N-CH_2-CO-NH_2 2-aminoacetamide (PIN)
glycinamide

Note that the 4-amide of aspartic acid has its own name, asparagine, and the 5-amide of glutamic acid is glutamine. Their 1-amides are named aspartic 1-amide and glutamic 1-amide, respectively.

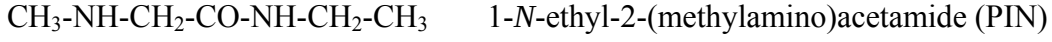
Preferred IUPAC names of anilides are formed by *N* substitution of the amide group by a phenyl group or substituted phenyl group. The ending 'anilide', in place of 'amide', may be used in general nomenclature.

Example:

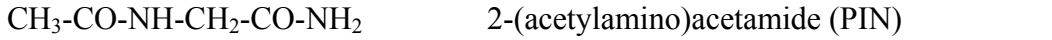
H_2N-CH_2-CO-NH-C_6H_5 2-amino-1-*N*-phenylacetamide (PIN)
glycinanilide

Substitution on nitrogen atoms in amides of amino acids is expressed by the methods described for amides (P-66.1.1.3) and amines (P-62.2.1.1.2).

Examples:



CH_3-NH-CH_2-CO-NH-CH_2-CH_3 1-*N*-ethyl-2-(methylamino)acetamide (PIN)

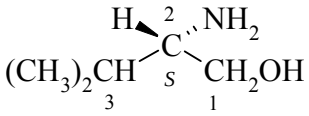


CH_3-CO-NH-CH_2-CO-NH_2 2-(acetylamino)acetamide (PIN)

P-103.2.4 Alcohols, aldehydes, and ketones

Preferred IUPAC names of alcohols, aldehydes and ketones corresponding to amino acids with retained trivial names are formed systematically by using the principles, rules and conventions of substitutive nomenclature. The endings 'ol' and 'al' added to the retained names, with elision of the final letter 'e', may be used in general nomenclature. Stereoisomers are denoted by '*R*' and '*S*' stereodescriptors.

Examples:



(*S*)-2-amino-3-methylbutan-1-ol (PIN)
valinol

$(CH_3)_2CH\text{-}CH_2\text{-}CH(NH_2)\text{-}CHO$

2-amino-4-methylpentanal (PIN)
leucinal

$H_2N\text{-}CH_2\text{-}CO\text{-}CH_2Cl$

1-amino-3-chloropropan-2-one (PIN)

P-103.3 Nomenclature of peptides

Nomenclature of peptides is highly specialized and well documented in reference 19.

P-103.3.1 Names of peptides

To name peptides, the names of acyl groups ending in 'yl' (see P-103.2.6) are used. Thus if the amino acids glycine, $H_2N\text{-}CH_2\text{-}COOH$, and alanine, $H_2N\text{-}CH(CH_3)\text{-}COOH$, condense so that glycine acylates alanine, the dipeptide formed, $H_2N\text{-}CH_2\text{-}CO\text{-}NH\text{-}CH(CH_3)\text{-}COOH$, is named glycylalanine. If they condense in the reverse order, the product $H_2N\text{-}CH(CH_3)\text{-}CO\text{-}NH\text{-}CH_2\text{-}COOH$ is named alanylglycine. Higher peptides ane named similarly, e.g. alanylleucyltryptophan.

P-103.3.2 Symbols of peptides

The peptide glycylglycylclycine is symbolized Gly-Gly-Gly. This involves modifying the symbols Gly for glycine, $H_2N\text{-}CH_2\text{-}COOH$, by adding hyphens to it, in three ways:

(a) Gly- = $H_2N\text{-}CH_2\text{-}CO–$

(b) -Gly = $–HN\text{-}CH_2\text{-}COOH$

(c) -Gly- = $–HN\text{-}CH_2\text{-}CO–$

Thus the hyphen, which represents the peptide bond, removes a –OH group from the –COOH group of the amino acid when written on the right of the symbol, and a hydrogen atom, when written on the left of the symbol.

P-103.3.3 Indication of configuration in peptides

The stereodescriptor 'L' is not indicated in the symbols of peptides. To the contrary, stereodescriptor 'D' is indicated at the front of the acyl group or name of each component having that configuration. When present, symbols for less common amino acids are always preceded by the stereodescriptor 'D' or 'L'.

Example:

Leu-D-Glu-L-aThr-D-Val-Leu (the symbol aThr is for allothreonine)

P-103.3.4 Cyclic peptides

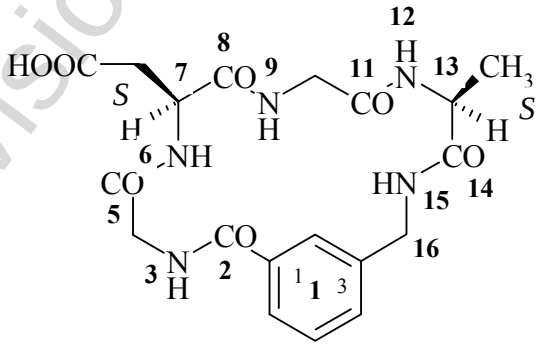
Names of cyclic peptides are composed of the name of the peptide preceded by the prefix 'cyclo'. The symbolic representation, placed in parentheses, is also preceded by the prefix 'cyclo'. Stereodescriptors 'D' are cited in peptides composed entirely of amino acid residues with regular peptide bonds (eupeptide linkage) (see ref. 49). Synthetic cyclopeptides require stereodescriptors '*R*' and '*S*' when symbols of amino acid residues are not used.

Preferred IUPAC names are systematic substitutive names.

Examples:

cyclo-(leucyl-D-phenylalanyl-prolyl-valyl-ornithyl-leucyl-D-phenylalanyl-prolyl-valyl-ornithyl-)

cyclo-(Leu-D-Phe-Pro-Val-Orn-Leu-D-Phe-Pro-Val-Orn-)



2-[(7*S*,13*S*)-2,5,8,11,14-pentoxo-3,6,9,12,15-pentaaza-1(1,3)-benzenacyclohexadecaphan-7-yl]acetic acid (PIN)

The prefixes 'endo' and 'des' used in peptide nomenclature deserve a special mention, as 'endo' is used in a different context and 'des' with a different meaning than in the nomenclature of organic compounds.

P-103.3.5 The prefix 'endo'.

In peptide nomenclature, the prefix 'endo' (nonitalic) is used to denote the insertion of an amino acid residue in a well identified position in the peptide. For example, the name endo-4a-tyrosine-angiotensin II means that the the amino acid residue 'tyrosyl' has been inserted between the positions 4 and 5 in the structure of tyrosine. The prefix 'endo' is not to be confused with the recommended stereodescriptor '*endo*' (written in italics) described in P-92.2.1.2.

P-103.3.6 The prefix 'des'

The subtractive prefix 'des', in peptide nomenclature, is used to denote the removal of an amino acid residue from any position in a peptide structure. For example, the name des-7-proline-oxytocin means that the amino-acid residue 'prolyl', located in position 7 of the peptide oxytocin, has been removed. In the modification of parent structures described in Section P-101, the prefix 'des' is used to indicate the removal of a terminal ring in steroids with the addition of the appropriate number of hydrogen atoms at each junction with the adjacent ring (see P-101.3.6).

P-104 Cyclitols

P-104.1 Definitions

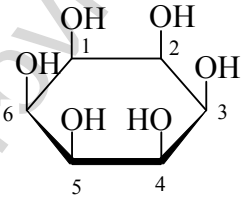
Cyclitols are cycloalkanes in which three or more ring atoms are each substituted with one hydroxyl group. Inositols, cyclohexane-1,2,3,4,5,6-hexols, are a specific group of cyclitols.

Preferred IUPAC names are retained names of inositols and *O*-alkyl, aryl and acyl derivatives.

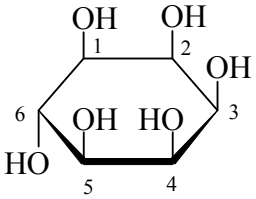
P-104.2 Name construction

Various methods are recommended for naming cyclitols.

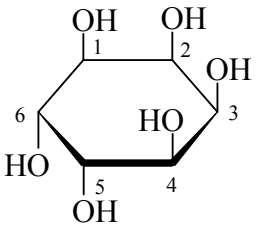
P-104.2.1 Stereoisomeric inositols are described by adding italicized prefixes at the front of the name 'inositol'. Positional numbers described in P-104.2.3 are shown in parentheses. Names denoted by the prefixes are preferred IUPAC names.



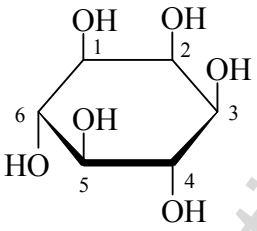
cis-inositol (PIN)
(1,2,3,4,5,6/0-)



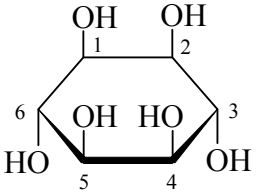
epi-inositol (PIN)
(1,2,3,4,5/6-)



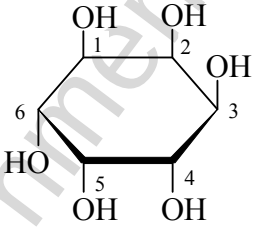
allo-inositol (PIN)
(1,2,3,4/5,6-)



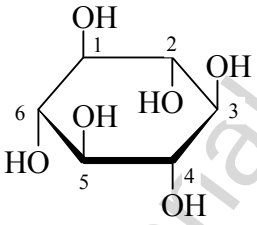
myo-inositol (PIN)
(1,2,3,5/4,6-)



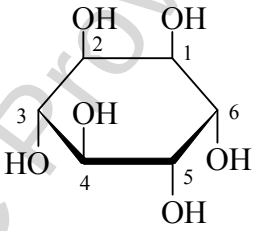
muco-inositol (PIN)
(1,2,4,5/3,6-)



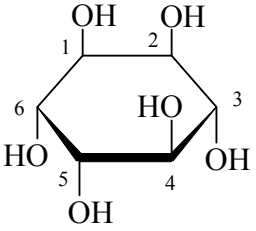
neo-inositol (PIN)
(1,2,3/4,5,6-)



scyllo-inositol (PIN)
(1,3,5/2,4,6-)



1D-*chiro*-inositol (PIN)
(1,2,4/3,5,6-)
(formerly D-*chiro*-inositol
or (+)-inositol)



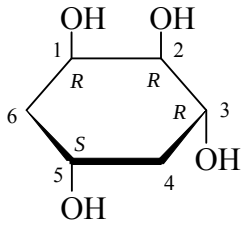
1L-*chiro*-inositol (PIN)
(1,2,4/3,5,6-)
(formerly L-*chiro*-inositol
or (−)-inositol)

The absolute configuration is denoted by 'D' and 'L' are determined in the following way. For the planar ring representation where the hydroxy group numbered 1 is above the plane of the ring,

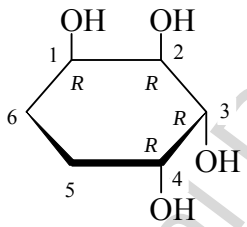
the configuration 'L' corresponds to a clockwise numbering, and the configuration 'D' corresponds to an anticlockwise numbering, as illustrated by the two enantiomeric *chiro*-inositols.

P-104.2.2 Cyclitols, with the exception of inositols, are named systematically based on the parent hydride cyclohexane by using the CIP method and its sequence rules for describing stereoisomers. This method is preferred to the method of positional numbers described in P-104.2.3.

Examples:



(1*R*,2*R*,3*R*,5*S*)-cyclohexane-1,2,3,5-tetrol (PIN)



(1*R*,2*R*,3*R*,4*R*)-cyclohexane-1,2,3,4-tetrol (PIN)

P-104.2.3 Locants are assigned to hydroxy groups in cyclitols, and thus the direction of numbering is described, with reference to the steric relations and nature of the substituents attached to the ring. The substituents lying above the plane of the ring constitute a set, and those lying below another set. Lowest locants are related to one set of the substituents according to the following criteria, which are applied successively until a decision is reached.

(a) to the substituents considered as a numerical series, without regard to configuration;

(b) if one set of the substituents is more numerous than the other, to the more numerous;

(c) if the set are equally numerous and one of them can be denoted by lower numbers, to that set;

(d) to substituents other than unmodified hydroxyl groups;

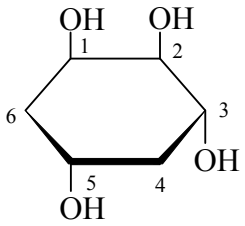
(e) to the substituent first cited in the alphanumerical order;

(f) to those designations that lead to an L rather than a D configuration, as determined by method (1) above (applies to meso compounds only)

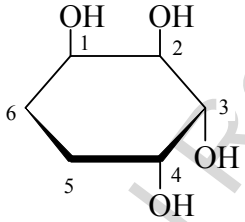
(g) the positional numbers are described by means of the numerator consisting of the set of substituents with the lowest locants, arranged in ascending order. The denominator is the other set. The fraction is denoted by the solidus (/) symbol.

Preferred IUPAC names are formed by the method in P-104.4.2.

Examples:



1L-1,2/3,5-cyclohexanetetrol
(1*R*,2*R*,3*R*,5*S*)-cyclohexane-1,2,3,5-tetrol (PIN)



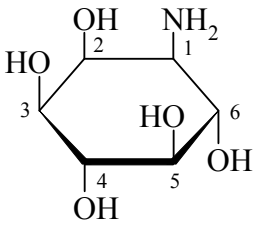
1L-1,2/3,4-cyclohexanetetrol
(1*R*,2*R*,3*R*,4*R*)-cyclohexane-1,2,3,4-tetrol (PIN)

P-104.3 Derivatives of cyclitols

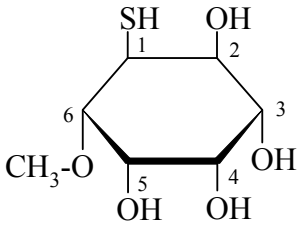
P-104.3.1 Derivatives of inositols

Inositols are modified in the same way as carbohydrates to generate preferred IUPAC names for their derivatives. There is no limit to *O*-substitution by alkyl (aryl) and acyl groups (see P-102.5.6.1). Hydroxy groups can be exchanged for amino groups using the 'deoxy' operation (see P-102.5.6.4). When characteristic groups that are senior to hydroxyl groups are put in the place of a hydroxyl group, fully substitutive names must be constructed. The numbering of the inositol remains unchanged and the configuration is expressed by an 'L' or 'D' stereodescriptor.

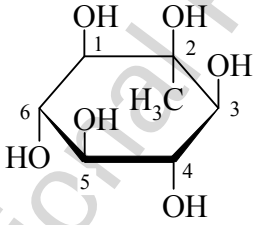
Examples:



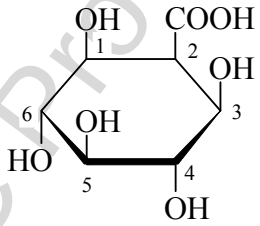
1L-1-amino-1-deoxy-*myo*-inositol (PIN)



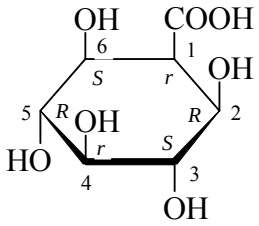
1L-1-deoxy-6-*O*-methyl-1-sulfanyl-*chiro*-inositol
(not 1L-6-*O*-methyl-1-thio-*chiro*-inositol)
(1*S*,2*S*,3*R*,4*S*,5*S*,6*S*)-6-methoxy-1-sulfanylcyclohexane-2,3,4,5-tetrol (PIN)



2-methyl-*myo*-inositol
(1*R*,2*S*,3*S*,4*R*,5*R*,6*S*)-2-methylcyclohexane-1,2,3,4,5,6-hexol (PIN)



I



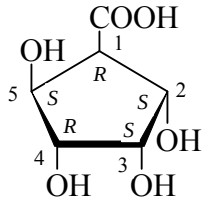
II

I 2-carboxy-2-deoxy-*myo*-inositol
II (1*r*,2*R*,3*S*,4*r*,5*R*,6*S*)-2,3,4,5,6-pentahydroxycyclohexane-1-carboxylic acid (PIN)

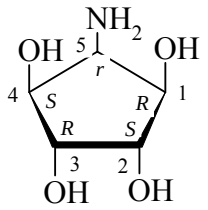
P-104.3.2 Derivatives of cyclitols other than inositols

Preferred names of derivatives of cyclitols other than inositols are all constructed by applying the principles, rules and conventions of substitutive nomenclature described in Chapters 1 to 9.

Examples:



(1*R*,2*S*,3*S*,4*R*,5*S*)-2,3,4,5-tetrahydroxycyclopentane-1-carboxylic acid (PIN)

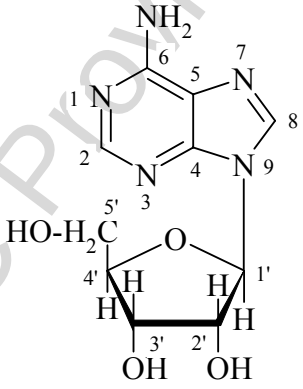


(1*R*,2*S*,3*R*,4*S*,5*r*)-5-aminocyclopentane-1,2,3,4-tetrol (PIN)

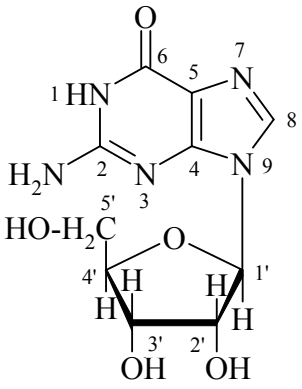
P-105 Nucleosides

Nucleosides are ribosyl or deoxyribosyl derivatives of the pyrimidine or purine bases adenine, guanine, xanthine, thymine, cytosine, and uracil, which are all retained names and are all preferred IUPAC names.

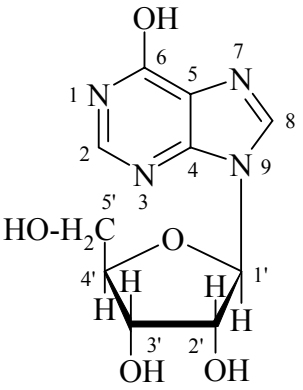
P-105.1 The following names are retained and recommended as preferred IUPAC names.



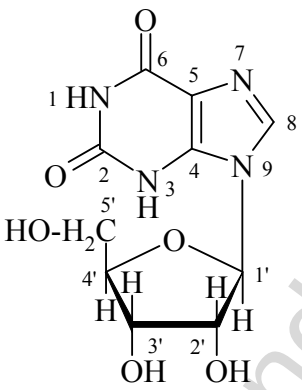
adenosine (PIN)



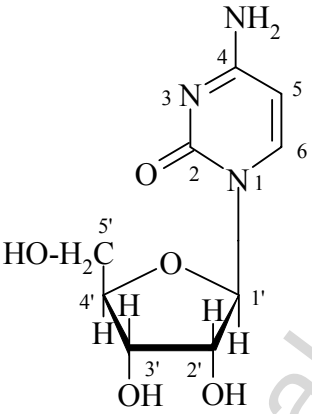
guanosine (PIN)



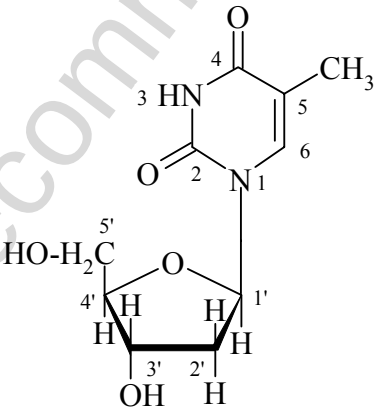
inosine (PIN)



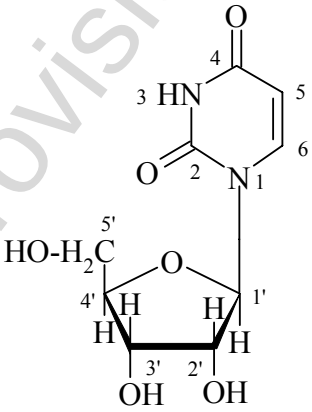
xanthosine (PIN)



cytidine (PIN)



thymidine (PIN)

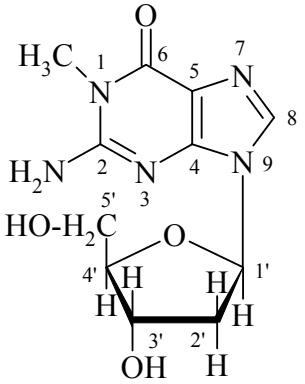


uridine (PIN)

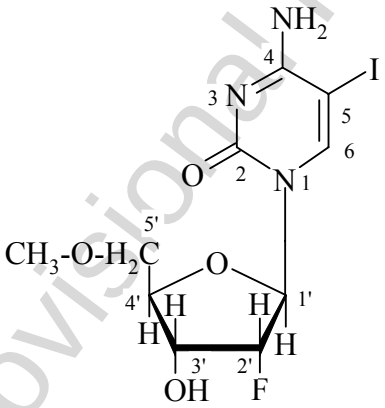
P-105.2 Substitution on nucleosides

P-105.2.1 Nucleosides having retained names can be fully substituted on the purine or pyrimidine ring. Replacement of oxo groups of nucleosides is described by functional replacement prefixes. The ribofuranosyl component may be modified as prescribed for carbohydrates (see P-102.5). 2′- and 3′- deoxyribose modificationsof the ribose component are allowed in preferred IUPAC names.

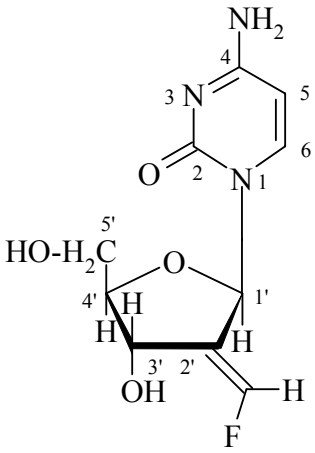
Examples:



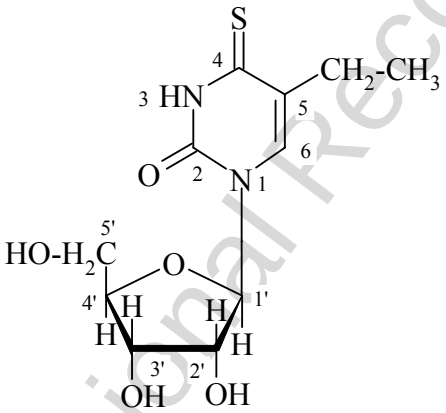
2′-deoxy-1-methylguanosine (PIN)



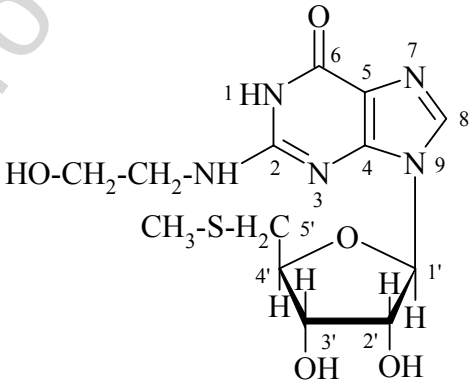
2′-deoxy-2′-fluoro-5-iodo-5′-*O*-methylcytidine
4-amino-1-[(2R,3R,4R,5R)-3-fluoro-4-hydroxy-5(hydroxymethyl)oxolan-2-yl]-5-iodopyrimidin-2(1*H*)-one (PIN)



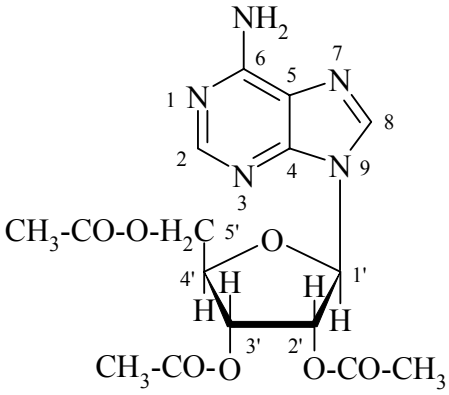
(2′*E*)-2′-deoxy-2′-(fluoromethylidene)cytidine (PIN)
4-amino-1-(2R,3*E*,4S,5R)-4-hydroxy-5(hydroxymethyl)-2-(fluoromethylidene)oxolan-2-yl]pyrimidin-2(1*H*)-one (PIN)



5-ethyl-4-thiouridine (PIN)



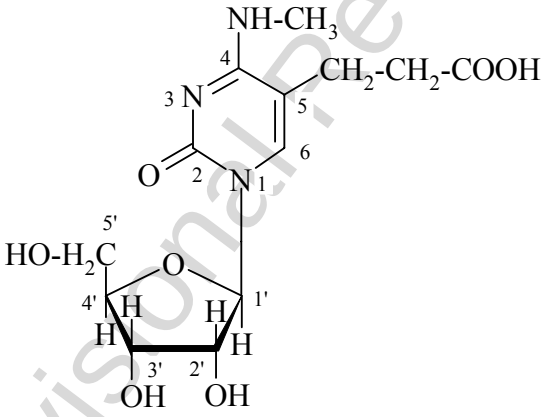
N-(2-hydroxyethyl)-5′-*S*-methyl-5′-thioguanosine (PIN)



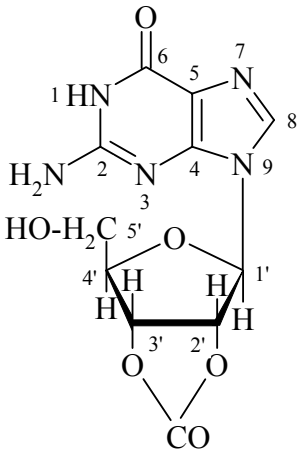
2′,3′,5′-tri-*O*-acetylguanosine (PIN)
guanosine 2′,3′,5′-triacetate

P-105.2.2 In the presence of a characteristic group higher than (pseudo) ketone, normal substitutive nomenclature principles are applied.

Example:



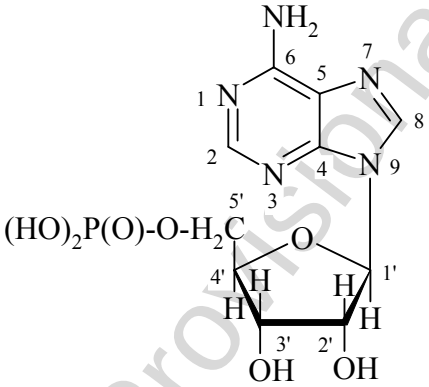
3-[4-(methylamino)-2-oxo-1-β-D-ribofuranosyl-1,2-dihydropyrimidin-5-yl]propanoic acid (PIN)



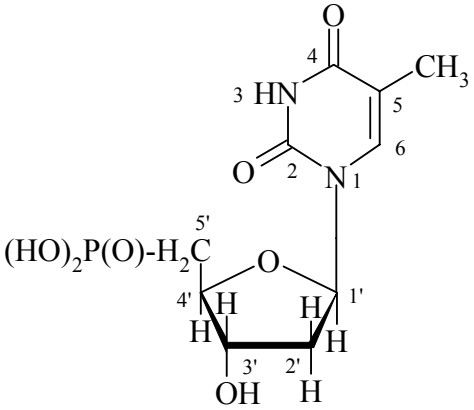
2′,3′-dideoxyguanosine-2′,3′-diyl carbonate (PIN, see P-101.7.4)
adenosine cyclic-2′,3′-carbonate

P-106 Nucleotides

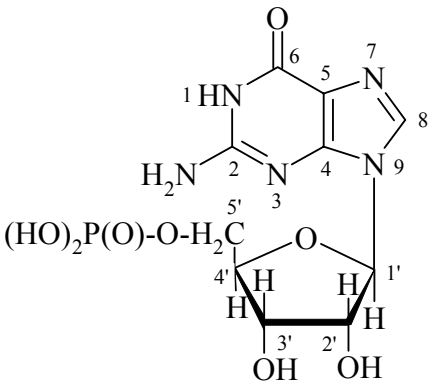
P-106.1 The following traditional names are retained to be used as preferred IUPAC names for esters of nucleosides with phosphoric acid. The primed locant of the ribosyl component is cited to locate the position of the phosphate group.



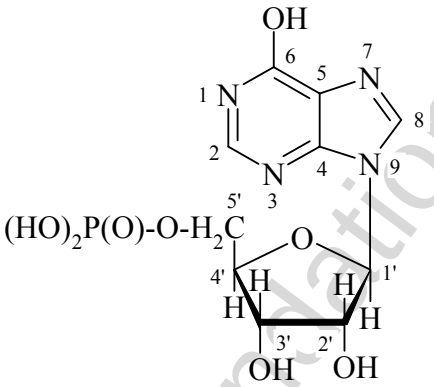
5′-adenylic acid (PIN)



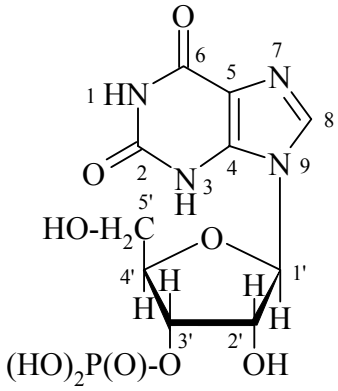
5′-thymidylic acid (PIN)



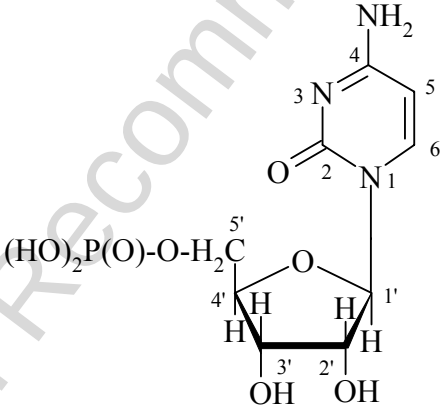
5′-guanylic acid (PIN)



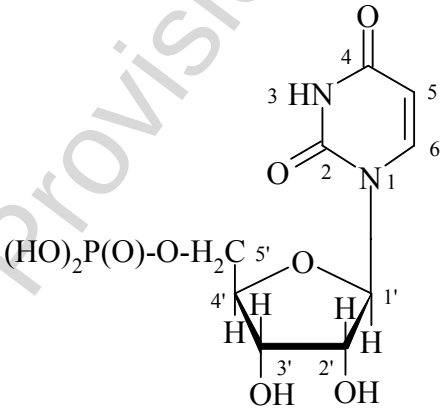
5′-inosinic acid (PIN)



3′-xanthylic acid (PIN)



5′-cytidylic acid (PIN)

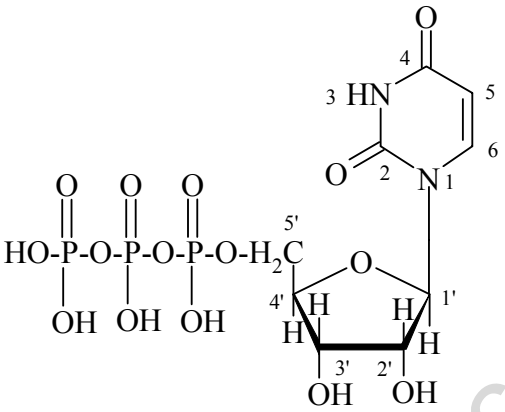


5′-uridylic acid (PIN)

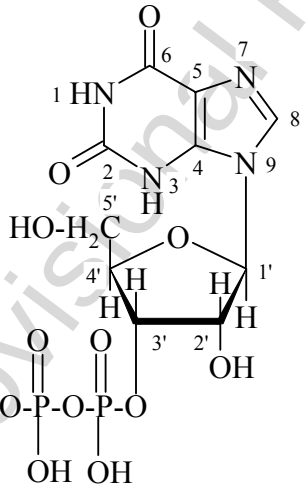
P-106.2 Nucleotide diphosphates and triphosphates

Diphosphate, triphosphate, etc. esters of nucleosides are named by citing a phrase such as diphosphate, after the name of the nucleoside. Preferred IUPAC names must indicate the presence of the hydrogen atoms on the diphosphate, triphosphate, etc. component of the molecule. Parentheses are used to avoid ambiguity.

Examples:



uridine 5′-(tetrahydrogen triphosphate) (PIN)

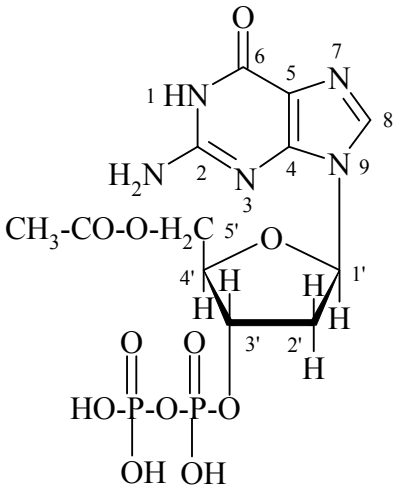


xanthine 3′-(trihydrogen phosphate) (PIN)

P-106.3 Derivatives of nucleotides

P-106.3.1. Derivatives of nucleotides having retained names are named in the same manner as the corresponding nucleoside, i.e., they can be fully substituted on the purine or pyrimidine ring and the ribofuranosyl component may be modified as prescribed for carbohydrates (see P-102.5). 2′- and 3′- deoxyribose modificationsof the ribose component are allowed in preferred IUPAC names.

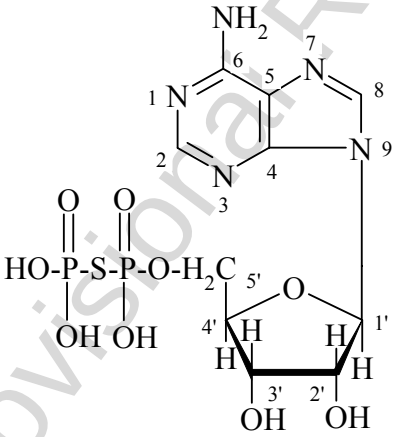
Examples:



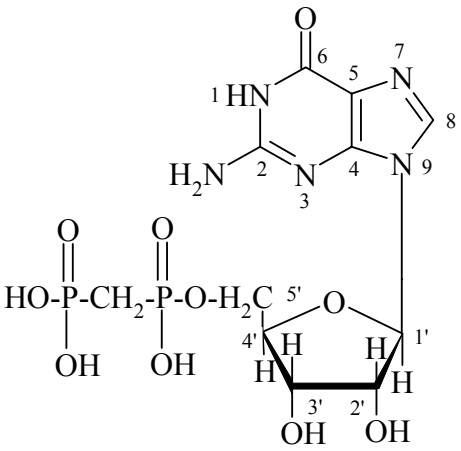
2′-deoxy-5′-*O*-acetylguanosine 3′-(trihydrogen diphosphate) (PIN)

P-106.3.2 Analogues of nucleoside di- and polyphosphates can be named by the functional replacement techniques applicable to di- and polyphosphoric acids (see P-67.2).

Examples:



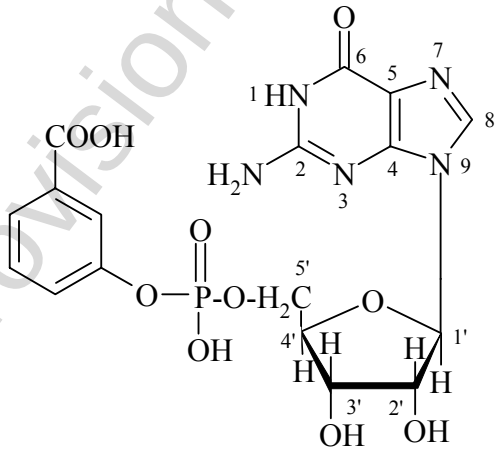
adenosine 5′-(trihydrogen 2-thiodiphosphate) (PIN)



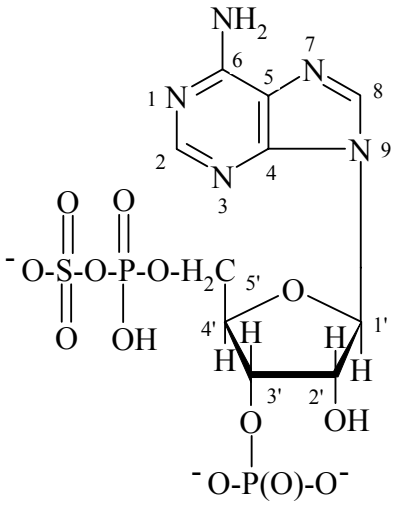
guanosine 5′-(trihydrogen methylenediphosphonate) (PIN)
guanosine 5′-(trihydrogen 2-carbadiphosphate)
(this name preserves the integrity of the
nucleotide name; see P-101.4.3)

P-106.3.3 In the presence of a characteristic group higher than the phosphoric acid residue, normal substitutive nomenclature principles may be applied. Substitutive prefix names may be derived from the traditional names for the nucleotide monophosphates by replacing the 'ic acid' ending with 'yl', for example, adenylyl and cytidylyl. Note that the substituent prefix name from inosinic acid is an exception; it is named inosinylyl so that the ending is like the other substituent prefix names derived from the nucleotide monophosphates.

Examples:



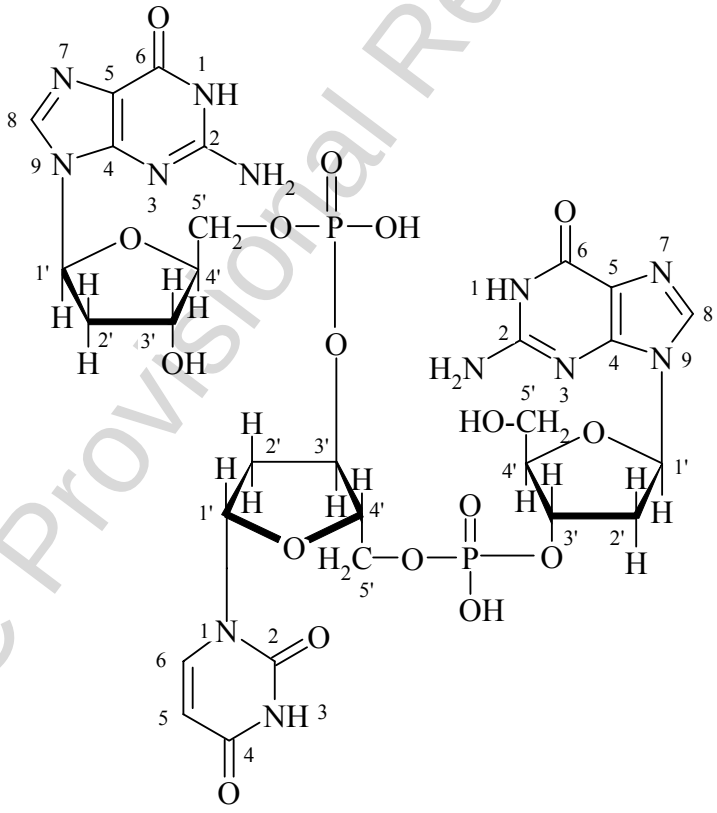
3-(5′-guanylyloxy)benzoic acid (PIN)



3′-*O*-phosphonato-5′-adenylyl sulfate (PIN)
3′-phospho-5′-adenylyl sulfate

P-106.3.4 Oligonucleotides are named using the prefix names derived from the traditional names for the nucleotides.

Example:



2′-deoxyguanylyl-(3′→5′)-2′-deoxyuridinylyl-(3′→5′)-2′-deoxyguanosine (PIN)

P-107 Lipids

P-107.1 Definitions

'Lipids' is a loosely defined term for substances of biological origin that are soluble in nonpolar solvents. They consist of saponifiable lipids, such as 'glycerides' (fats and oils) and 'phospholipids', as well as nonsaponifiable lipids, specifically 'steroids'.

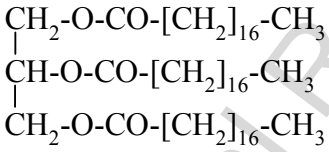
The nomenclature of glycerides, phospholipids and glycolipids has been published in 1976 (ref. 46); the nomenclature of glycolipids was revised in 1997 (ref. 47).

Preferred IUPAC names for individual substances are formed by using substitutive nomenclature with CIP stereodescriptors.

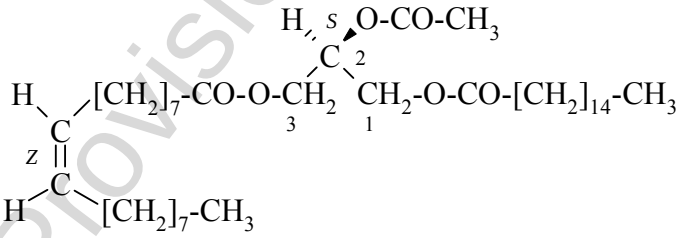
P-107.2 Glycerides

Glycerides are esters of glycerol (propane-1,2,3-triol) with fatty acids. They are by long established custom subdivided into triglycerides, 1,2- or 1,3-diglycerides, and 1- or 2-monoglycerides, according to the number and position of acyl groups. The recommended method for naming individual glycerides is mono-, di- or tri-*O*-acylglycerol. The name glycerol is allowed in general nomenclature to name organic compounds; it is however the preferred name in the field of natural products.

Examples:



tri-*O*-octadecanoylglycerol
propane-1,2,3-triyl trioctadecanoate (PIN)



(2*S*)-2-*O*-acetyl-1-*O*-hexadecanoyl-3-*O*-(9*Z*)-octadec-9-enoylglycerol
(numbering shown)
(2*S*)-2-*O*-acetyl-1-*O*-oleoyl-3-*O*-palmitoylglycerol
(2*S*)-propane-1,2,3-triyl 2-acetate-1-hexadecanoate-3-[(9*Z*)-octadec-9-enoate] (PIN)

P-107.3 Phospholipids

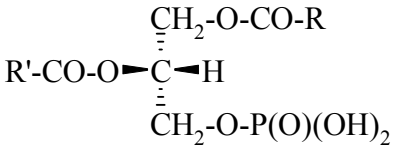
Phospholipids are lipids containing phosphoric acid as mono- or diesters, including 'phosphatidic acids' and 'phosphoglycerides'.

Phosphatidic acids are derivatives of glycerol in which one hydroxyl group, commonly but not necessarily primary, is esterified with phosphoric acid, and the other two hydroxyl groups are esterified with fatty acids.

Phosphoglycerides are phosphoric diesters, esters of phosphatidic acids, generally having a polar head group (–OH or $–NH_2$) on the esterified alcohol which typically is 2-aminoethanol (not ethanolamine), choline, glycerol, inositol, serine. The term includes 'lecithins' and 'cephalins'.

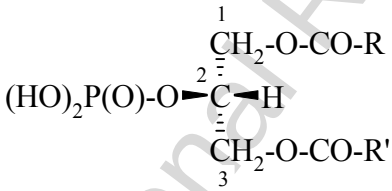
P-107.3.1 Phosphatidic acids

P-107-3.1.1 Phosphatidic acids have the following generic structure:



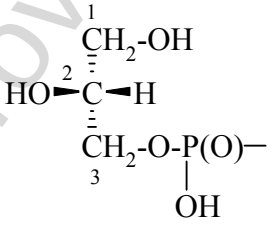
3-*sn*-phosphatidic acid
(for a discussion and examples of the symbol '*sn*' see P-107.3.1.2)

In general, the 3-*sn*-phosphatic acids are simply called phosphatidic acids.



2-phosphatidic acid

The name of the monovalent acyl group:

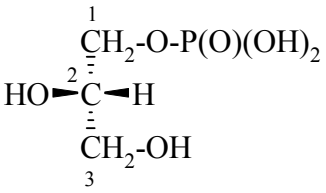


is 'phosphatidyl', a retained name used in general nomenclature.

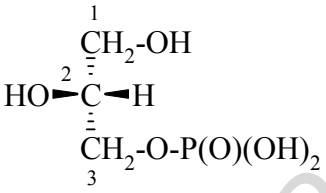
P-107.3.1.2 Configuration of phosphatidic acids

In order to designate the configuration of glycerol derivatives, the carbon atoms of glycerol are numbered stereospecifically. The carbon atom that appears on top of that Fischer projection that shows a vertical carbon chain with the hydroxy group at carbon 2 to the left is designated as C-1.

To differentiate such numbering from conventional numbering, which conveys no steric information, the stereodescriptor '*sn*' (for stereospecifically numbered) is used. This descriptor is written in lower case italic letters, even at the beginning of a sentence, immediately preceeding the glycerol term, from which it is separated by a hyphen. The stereodescriptor '*rac*' is used to describe racemates and the stereodescriptor '*X*' may be used if the configuration of the compound is unknown or unspecified.



sn-glycerol 1-phosphate
(2*S*)-1,2-dihydroxypropyl dihydrogen phosphate (PIN)

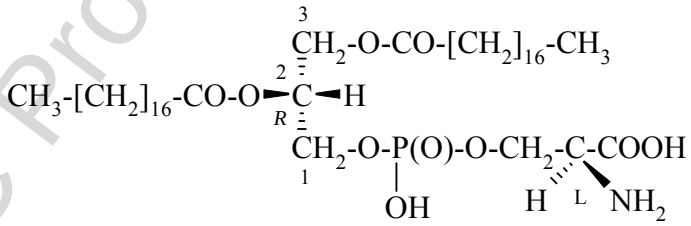


sn-glycerol 3-phosphate
(2*R*)-1,2-dihydroxypropyl dihydrogen phosphate (PIN)

P-107.3.1.3 Phosphatidylserines

The term 'phosphatidylserines' is used to describe the acyl derivatives of phosphatidic acids whose phosphorus acid component is esterified with the amino acid 'serine', usually L-serine. Preferred names of specific compounds are formed in accordance with the principles, rules and conventions of substitutive nomenclature.

Example:

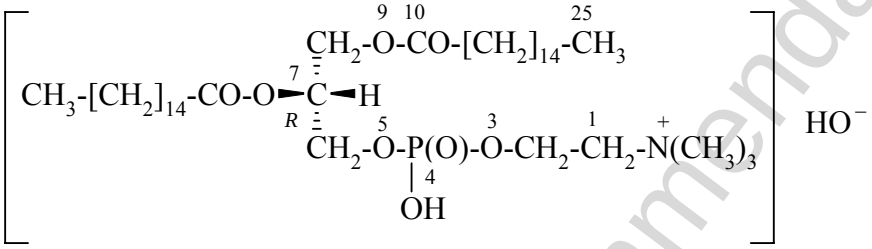


{[(2*R*)-2,3-bis(octadecanoyloxy)propoxy]hydroxyphosphoryl}-L-serine (PIN)
(2*R*)-2-*O*,3-*O*-bis(octadecanoyl)phosphatidyl-L-serine (traditional name)

P-107.3.1.4 Phosphatidylcholines

The term 'phosphatidylcholines' is used to describe the acyl derivatives of phosphatidic acids whose phosphorus acid component is esterified with choline. Preferred IUPAC names of specific compounds are formed in accordance with the principles, rules and conventions of substitutive nomenclature.

Example:

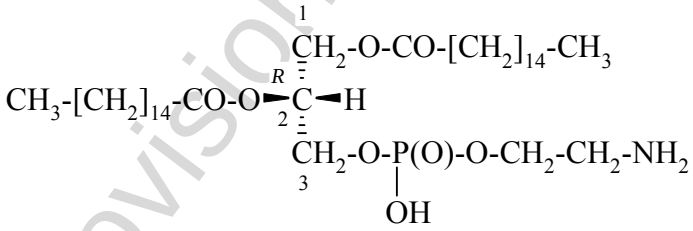


(7*R*)-4-hydroxy-*N*,*N*,*N*-trimethyl-7-(hexadecanoyloxy)-4,10-dioxo-3,5,9-trioxa-4λ^5-phosphapentacosanaminium hydroxide (PIN)

P-107.3.1.5 Phosphatidylethanolamine

The term 'phosphatidylethanolamines' (more correctly 'phosphatidyl(amino)ethanols) is used to describe the acyl derivatives of phosphatidic acids whose phosphorus acid component is esterified with 2-aminoethanol. Preferred IUPAC names of specific compounds are formed in accordance with the principles, rules and conventions of substitutive nomenclature.

Example:

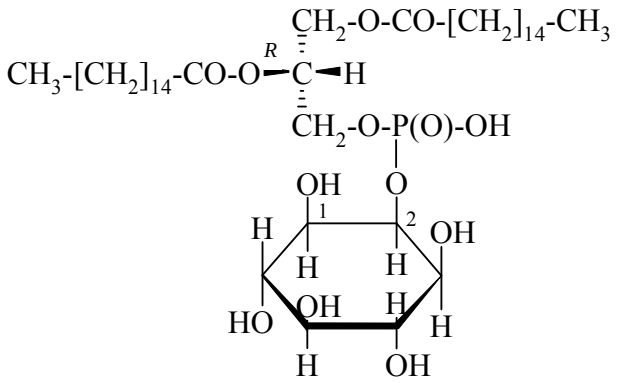


(2*R*)-3-{[(2-aminoethoxy)hydroxyphosphoryl]oxy}propane-1,2-diyl dihexadecanoate (PIN)

P-107.3.2.6 Phosphatidylinositols

The term 'phosphatidylinositols' is used to describe the acyl derivatives of phosphatidic acids whose phosphorus acid component is esterified with an inositol molecule. Preferred IUPAC names of specific compounds are formed in accordance with the principles, rules and conventions of substitutive nomenclature.

Example:



(2*R*)-3-[({[(1*r*,2*R*,3*S*,4*r*,5*R*,6*S*)-(2,3,4,5,6-pentahydroxycyclohexyl]oxy}
hydroxyphosphoryl)oxy]propane-1,2-diyl dihexadecanoate (PIN)
2-*O*-{[(2*R*)-2,3-bis(hexadecanoyloxy)propoxy]hydroxyphosphoryl}-*myo*-inositol

P-107.4 Glycolipids

P-107.4.1 Definitions

The term 'glycolipid' designates any compound containing one or more monosaccharide residues bound by a glycosidic linkage to a hydrophobic moiety such as an acyl glycerol, a sphingoid (a long chain aliphatic amino alcohol), a ceramide (an *N*-acyl-sphingoid) or a prenylphosphate.

Glycoglycerolipids are glycolipids containing one or more glycerol residues.

Glycosphingolipids designate lipids containing at least one monosaccharide residue and either a sphingoid or a ceramide.

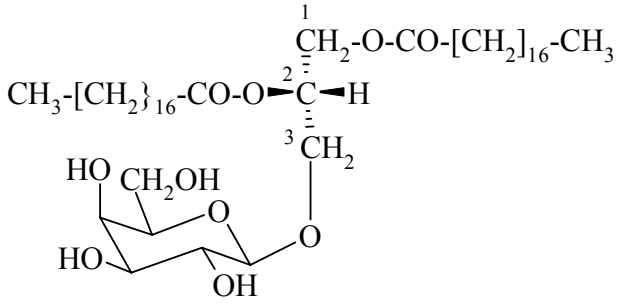
The term 'glycophosphatidylinositol' designates glycolipids which contain saccharides glycosidically linked to the inositol moiety of phosphatidylinositols.

Specific compounds are named systematically.

P-107.4.2 Glycoglycerolipids

Specific compounds are named on the basis of parent glycerol, whose configuration is specifically numbered as indicated in P-107.3.2.1.

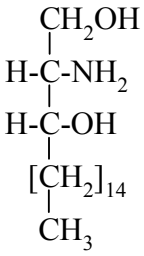
Example:



3-*O*-β-D-galactopyranosyl-1,2-di-*O*-octadecanoyl-*sn*-glycerol
(2*R*)-3-(*O*-β-D-galactopyranosyl)propane-1,2-diyl dioctadecanoate

P-107.4.3 Glycosphingolipids

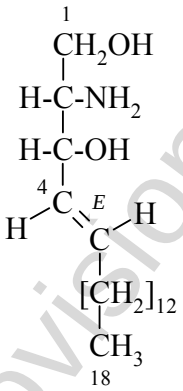
P-107.4.3.1 Names are formed by using the retained name 'sphinganine' for the aliphatic amino alcohol having the described absolute configuration. The retained name 'sphinganine' is preferred to the systematic name (2*S*,3*R*)-2-aminooctadecane-1,3-diol.



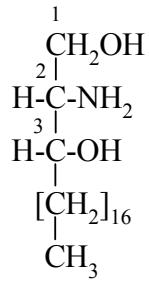
sphinganine (PIN)
(2*S*,3*R*)-2-aminooctadecane-1,3-diol

The retained name sphinganine is used to generate the names of unsaturated derivatives. Other derivatives, such as hydroxy, oxo and amino derivatives, as well as isomers with different chain length or other diastereoisomers are named systematically in accordance with the principles, rules and conventions of substitutive nomenclature.

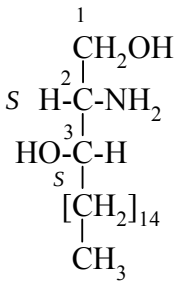
Examples:



(4*E*)- sphing-4-enine
(2S,3R,4*E*)-2-aminooctadec-4-ene-1,3-diol (PIN)



(2*R*,3*R*)-2-aminoicosane-1,3-diol (PIN)
icosasphinganine

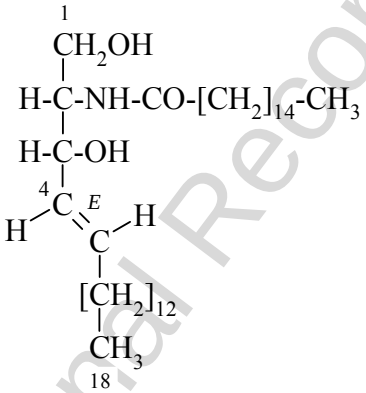


(2*S*,3*S*)-2-aminooctadecane-1,3-diol (PIN)

P-107.4.3.2 Ceramides

Ceramides are *N*-acylsphingoids.

Example:

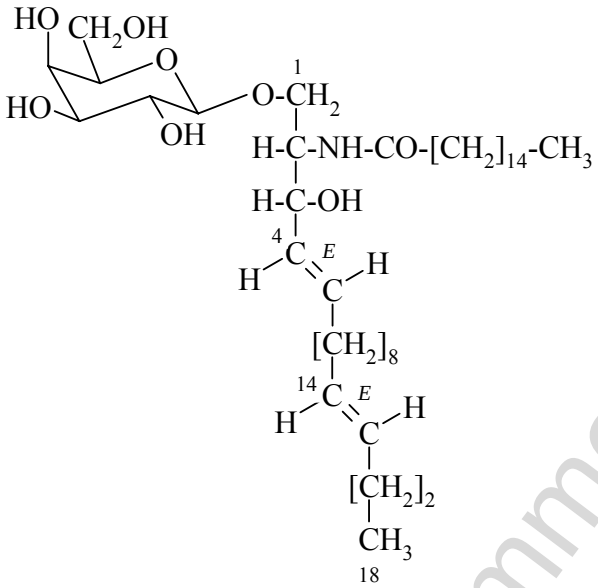


(4*E*)- *N*-hexadecanoylsphing-4-enine
N-[(2S,3R,4*E*)-1,3-dihydroxyoctadec-4-en-2-yl]hexadecanamide (PIN)

P-107.4.3.4 Neutral glycosphingolipids

A neutral glycosphingolipid is a carbohydrate containing derivative of a sphingoid or ceramide. It is understood that the carbohydrate residue is attached by a glycosidic linkage to 1-*O*-. Preferred systematic names must include all locants.

Example:



(4*E*,14*E*)-*N*-hexadecanoylsphinga-4,14-dienine (PIN)
N-[(4*R*,4*E*,14*E*)-1-(β-D-galactopyranosyloxy)-3-hydroxyoctadeca-4,14-dien-2-yl]hecadecanamide (PIN)